THE DARK SIDE OF WHITE LIES: HOW GENDER-BIASED FEEDBACK CONTRIBUTES TO THE GLASS CEILING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Lily E. Jampol

August 2014

THE DARK SIDE OF WHITE LIES: HOW GENDER-BIASED FEEDBACK CONTRIBUTES TO THE GLASS CEILING

Lily E. Jampol, Ph. D.
Cornell University 2014

Although explicit endorsement of sexism has declined, covert barriers to the advancement of women in the workplace remain. In this dissertation, I argue that women are given benevolent but less accurate feedback about their performance and that this communication bias is accentuated by stereotypical beliefs about women. I additionally argue that this gender feedback bias has significant consequences for women and for workplace productivity and I discuss possible interventions that could be developed to ameliorate the bias. Although intentions may be benevolent, unequal feedback may undermine success in the workplace if women are not receiving accurate information needed to improve performance.

In study 1a, people upwardly distorted evaluations of female authors' work when giving feedback but not to male authors. The results suggest that people tell more white lies to women about their performance. Study 1b demonstrated that this elevated feedback was not just "shifting standards" in evaluation of men and women but white lies. In study 2a, people were more likely to assume an employee was female if the employee had been given inaccurate feedback. In study 2b, people justify their assumptions about gender using language related to stereotypes about women's emotionality. In Study 3, participants who held stereotypical views about women were more likely to tell white lies to women. Study 4 demonstrates that one of the

potential problems with a gender bias during feedback is a mismatch between behavior and women's actual desires regarding biased feedback. Study 5 explores the possible performance consequences of being told white lies. And finally, in Study 6, I attempt to design and evaluate a possible intervention for use in the workplace in order to attenuate the gender feedback bias. These findings are interpreted using a benevolent sexism framework and the workplace implications for a feedback bias for women are discussed. Collectively, the studies suggest that white lies during performance feedback to women may contribute to the maintenance of the "glass ceiling."

Key words: gender performance lying feedback sexism language

BIOGRAPHICAL SKETCH

Lily Jampol grew up in a B&B and coffee farm in Costa Rica, and moved to the United States in 2002 to pursue a higher education degree. She graduated with a BA in 2006 from Mount Holyoke College in South Hadley, MA where she majored in Politics and minored in French. Her senior thesis was on the topic of American Political Philosophy and she used the food industry as a lens through which to examine the political implications of Americans' concepts of freedom and happiness. After she graduated in 2007, Lily entered a special non-degree program at Harvard University, during which she took courses in the disciplines of psychology and economics and volunteered in different psychological laboratories. After completing the program in 2008, Lily was hired as a researcher at Princeton University to examine the link between cognitive fatigue and poverty in the psychology department. In 2009 Lily began her graduate school program in the psychology department at Cornell University where she studied judgment and decision making and cognitive biases and graduated in 2014 after completing her PhD.

This dissertation is dedicated to my family: Teri, Glenn, Olivia, & Jeremy

ACKNOWLEDGMENTS

I have several people to whom I owe my gratitude, and without whom this dissertation would never have happened. I would first like to thank my advisors for seeing me through five years of learning, struggling, developing as an academic, and inspiring me the whole time. I am very fortunate to have had all of them on my committee. In particular, I am very grateful to Dr. Vivian Zayas, who shared my enthusiasm for this research that is so important to me, and who approached the challenge of collaborating on work outside of her immediate specialty with excitement instead of apprehension. Her ideas and contributions were integral to this dissertation and research. I also want to thank her for her guidance concerning both academic and non-academic life-changes and for always being accessible and supportive.

I owe my gratitude to Dr. Tom Gilovich as the chair of my committee and for supporting me in my endeavor to do all kinds of different research during my five years at Cornell, and for guiding me through devastating research failures and exciting successes alike. I am also particularly grateful for both his flexibility and continuous support in difficult decisions and life changes, including moving across an ocean while still in graduate school. I would also like to thank Dr. David Pizarro for inspiring me and challenging me to think deeply about very interesting questions, and popular culture, and advising me on how to do interesting and creative research. Finally, I would like to thank Dr. Tim DeVoogd for providing a unique perspective in regards to my work, and to Dr. David field for reading my dissertation and for his role as an incredibly effective Director of Graduate Studies.

I am in deep gratitude to my family, some of whom directly contributed to this

dissertation. I would not have become a confident academic and remained a sane person throughout this process if it were not for my husband and confidant, Dr. Jeremy Skipper, who has supported me, loved me, and challenged me and who above all inspires me to continue to reach past what I think are my limits. My sister and best friend, Olivia Jampol, also deserves special thanks, for without her and Brady Hiatt's comments and their late night copy-editing, this dissertation would not have been near the quality I had desired it to be. Finally, I owe my gratitude to my parents, Glenn and Teresa Jampol for being the people who made me the idealist that I am, who inspire me to do good in the world, and who constantly remind me that despite setbacks, I always have something of which to be proud.

I would also like to mention my friends who were integral to my experience in graduate school, and who will always be the best group of people I can think of spending time with and the smartest, funniest, most engaging colleagues. Thank you Amit Kumar, Shai Davidai, Yardenne Greenspan, Erik Helzer, Jun Fukukura, and especially Chelsea Helion, my partner in crime, house, office, and dissertation pain. You are all loved and appreciated.

I must also thank both the National Science Foundation and the American Psychological Association for the generous funding that was granted to me. Thanks to the NSF Graduate Research Fellowship for funding three years of my degree, and to the Geis Memorial Award for its recognition of my work on sexism and for the funds that allow me to do my research and also to live more comfortably than I could have.

I would also like to thank all of my research assistants, Catherine Schrage, Di Wu, Natasha Lawson, Jonathan Westman, Joy Jiang, and Micah Turner who have

helped me collect data, and the participants who volunteered for my studies. I would also like to thank Pamela Cunningham for resolving every logistical difficulty I have had. Finally, much thanks to Dr. Ana Guinote for her support and guidance concerning academia in the UK, and to University College London for generously giving me space to work and an academic home in London.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

How do Covert Biases Contribute to Inequality in the Workplace?

While considerable progress toward gender equality has been made in the workplace over the past several decades, inequalities remain. For example, there is a gender leadership gap, particularly in top positions, where women continue to languish at approximately 17% of the overall workforce (e.g., the 2012 Catalyst Census) and make up only 3.2% of CEOs and 15.7% of director positions despite representing more than 51.5% of managers. Yet reports indicate that companies with more women in senior management outperform those with less gender diversity (2004 Catalyst report). It is not just businesses that suffer from inequality; research has shown that scientists and professors hold many of the same biases, hiring (hypothetical) female research assistants less often and for less pay (Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012), giving awards to female academics less often than male academics (Wold & Wennerås, 1997), and ignoring email discussion requests from women and minorities more often, especially in high paying disciplines (Milkman, Akinola, & Chugh, 2014).

The picture that emerges is that underneath even the most educated and egalitarian motives, biases not only exist but are also difficult to suppress and eradicate, despite good intentions. Thus, while much of the overt sexism that plagued the workplace has disappeared (though not entirely; Willness, Steel & Lee, 2007), this

manuscript addresses the following pressing question: what are the less explicit factors that may be contributing to gender inequality? Additionally, how do they remain undetected, how are they maintained, and how are they perpetuated?

Gender discrimination in the workplace is likely to play out in multiple ways, and maintained by a diverse range of singular and routine actions, and institutional structures and cultures. For example, much of the research investigating the barriers to advancement in the workplace for women has focused on institutional factors, such as child-rearing decisions in a system not set up for motherhood (Williams, Manvell, & Bornstein, 2006). However, not all barriers that women face today are so overt; indeed, they may occur even when all explicit signs indicate gender equality. Instead of overt factors, it may be that covert and implicit biases, those that may go undetected or even celebrated as routine or empathic (Barreto & Ellemers, 2005), are the most pernicious of biases and may particularly contribute to the continuation of inequality.

Specifically, I propose that one covert factor contributing to the maintenance of the "glass ceiling" may be the way in which women are given feedback regarding their performance. I propose that because telling people the truth can be difficult and stereotyping comes easily to us, benevolent dishonesty during performance feedback both draws from and undermines gender equality in the workplace.

I argue that there are two reasons why performance feedback may particularly elicit gender-biased communication. First of all, giving feedback is difficult because it is colored by an anxiety-producing tension between the more immediate goals of preserving social bonds and the future goals of enhancing workplace productivity. Second, stereotypes about the emotionality of women combined with the anxiety of

giving critical feedback may lead to the decision to tell white lies more often to women about their performance than to their male counterparts.

First I will review the literature on language and communication in the workplace and explain why performance feedback is likely to be a vehicle for bias. In the next section I will review the literature on ambivalent sexism and stereotypes about women and how these factors are likely to play out in situations of anxiety. Finally, I will discuss the difference between "shifting standards" of evaluation of women (Biernat & Manis, 1991) and white lies told to women.

Language as a Vehicle for Bias: Why are White Lies Likely During Performance Feedback?

Biases can easily be spread through everyday language. Language is easily manipulated and can be used to convey several different meanings at once (Linell, 1998; Gibbs, 1999) depending on the cultural norms and motives of the speaker (Krauss & Chiu, 1998). For example, people may change their speech patterns, the tone of their voice, and the content conveyed based on slight influences in their environment, the target of the conversation, the context of the conversation, and the social relationship that colors that particular conversation (Clark & Murphy, 1982; Graumann, 1989). There are also unspoken rules of communication, and mutual, yet implicit understandings of meaning (e.g. 'we should get coffee some time' usually means, 'I am just trying to end the conversation nicely') (Grice, 1957, 1969, 1975). Sometimes, however, motives may not be entirely clear to the agents involved in communication. People can communicate one thing while intending another, which can often lead to misunderstandings (e.g., Goffman, 1967, 1981; Grice, 1975;

Schwarz, 1996).

Performance feedback is one of the most ubiquitous and important forms of communication in the workplace, as it is essential to maintaining both the productivity of the employees and of the business (Hillman, Schwandt, & Bartz, 1990). Yet it is also one of the most difficult types of communication to both impart and receive, especially if the feedback is critical (Belschak & Hartog, 2008). People report disliking having to give negative feedback (Fisher, 1979), and supervisors have been shown to avoid, distort, and delay giving critical feedback to their employees (Benedict & Levine, 1988; Ilgen & Knowlton, 1980), especially when the employees are performing poorly and most need feedback (Larson, 1986). This effect was coined the "mum effect" (Tesser & Rosen, 1975).

When in a situation of social anxiety, such as giving direct feedback to someone with whom they often interact, a person may be more likely to manipulate language to reduce the impact of negative information, or even distort information itself in order to protect that person's feelings (DePaulo & Bell, 1996; Stevens & Fiske, 1995). This upward distortion or delay of negative information may reduce the target's pain while simultaneously reducing the supervisor's guilt (Charness & Dufwenberg, 2006; Fisher, 1979). Inaccurate information as a result of this prosocial distortion of language and information may be construed as 'white lies': altruistic lies that are told in order to protect another person's feelings (either consciously or unconsciously).[1]

White lies have been differentiated from blatant lies in that they are often told

[1] Of course, there are other types of white lies as well (including selfish white lies), but the argument in

in order to avoid a negative social interaction (DePaulo, 2004, 1996). However, Erat and Gneezy (2001) differentiate between a *pareto* white lie (one in which both parties benefit), and an *altruistic* white lie (where only the receiver of the lie benefits). While many people are uncomfortable telling the former, which can often be self-centered, most people report that telling an altruistic white lie is acceptable if it truly benefits someone else. If there is good reason to tell a lie (e.g., to protect someone) then people report feeling justified in that "it was the right thing to do" (Cwiden et al., 1984). Thus, white lies are generally accepted, and while people still report that they find all types of lies immoral, they have no trouble justifying white lies after telling them themselves (Turner et al., 1970) and especially to people they like (Bell & DePaulo, 1996); In fact, people report that about 70% of all the lies they tell are white lies (DePaulo, 1996; DePaulo & Kashy, 1998).

Competition Between Implicit Social Goals And Explicit Achievement Goals

People regularly pursue multiple goals simultaneously and can be influenced by different personal needs and social situations, which are often in competition with each other (Deci & Ryan, 2000), even if those goals and influences are outside of their awareness (for a review see: Custers & Aarts, 2003; Moskovitz, Li, & Kirk, 2004). The motivation to pursue these goals is facilitated by the affective responses that accompany them, even when the affect is implicit (Aarts, Custers, & Veltkamp, 2008; Fergusen & Bargh, 2004). These affective components (and individual personalities) can motivate people to take on different regulatory foci in order to enhance the pleasure or reduce the pain associated with particular goals (Higgins, 1997, 1998). That is, people can be motivated to either avoid negative consequences or approach

perceived positive outcomes, and these goals can be at odds with each other.

When offering critical performance feedback, the long-term goal is usually to impart information to an employee such that he or she can use that information to improve his or her work. However, in the immediate present, avoiding potential emotional reactions from the employee can be a particularly salient goal (Freitas, Gollwitzer, & Trope, 2004). People make short-term affect regulation a priority over other self-regulatory goals when feeling distress (Tice, Bratslavsky & Baumeister, 2001), just as they tend to do when giving critical feedback (Fisher, 1979). White lies are more likely to be told in the workplace when giving critical feedback because the motive behind giving feedback and the process of giving feedback are often at odds with one another. That is, while people who are giving feedback have a focal goal to impact performance and productivity by being truthful (DePaulo & Kashy, 1996), this motivation can be undermined by the more implicit alternate goal of being nice (Brief & Motowidlo, 1986; Shah & Kruglanski, 2002). While this goal of being nice is present in most white lie motivations (including romantic relationships), it may be especially pronounced in the workplace. There, productivity can depend on individual improvement, making truthfulness imperative, whereas in other relationships truthfulness (e.g. "you look fat in that dress") may not be essential or necessary for some greater good.

The tenuous state of accurate information delivery may be accentuated if the target is more deserving of empathy and if the relationship with the target is not fleeting. Humans are instinctively motivated to preserve social bonds and normally have empathic responses to other's pain (at least within their own social group see

Batson, 1998). Thus a prosocial norm of compassion can be especially strong if the communicator is likely to interact in the future with the target, which is often the case in most feedback situations (Baumeister & Leary, 1995; Tyler & Feldman, 2004). This may be particularly likely for individuals who elicit empathy and for groups of people especially if they are stereotypically perceived as needing protection.

Women As Targets Of Empathic Behavior

Women are particularly likely to be recipients of empathic but inaccurate information. Because it is particularly easy to tell white lies when one can use an 'empathic' or protecting reason to justify the lie, targets that elicit more pitying feelings may be more likely to receive such protective behavior. Historically, women have been the targets of empathic behavior and several gender biases, both explicit and implicit, which makes them perfect targets for white lies in the workplace, especially if the person giving the feedback holds stereotypical and therefore sexist views about women.

One of the most widespread stereotypes people hold about women is that they are more emotional than men (Fabes & Martin, 1991; Fischer, 1993; Grossman & Wood, 1993; Hess et al., 2000; Plant, Hyde, Keltner, & Devine, 2000). On the one hand, women are negatively perceived as being more emotionally unstable and volatile and worse at responding to negative information and over-reacting to emotional slights than are men (Timmers, Fischer, & Manstead, 2003; Barrett et al., 1998; Diener et al., 1985). On the other hand, they are also perceived positively as more empathic and socially attuned to other's emotions (Babchuck et al., 1985; Hall, 1978). These stereotypes are both widespread and appear to be learned and

incorporated into American culture from a very young age (Birnbaum & Chemelski, 1984).

While emotional intensity and volatility are typically perceived as feminine characteristics, anger is an example of a heightened emotion from which women are exempt. Conversely, men are perceived as more likely to express more dominance and aggression (Hochschild, 1983). That is, because of androgenizing hormones males may react to negative events (especially ones that challenge dominance) with stronger aggressive emotional reactions, and experience and exhibit emotions like anger (O'Connor, Archer, & Wu, 2004).

However it is unclear whether males and females are innately different emotionally, and it is difficult to separate experience from behavior; Reports of emotion, and the empirical evidence on gender differences in emotional responding is varied (Bradley, Codispoti, Sabatinelli, & Lang, 2001; Fujita, Diener, & Sandvik, 1991; Grossman & Wood, 1993; Labouvie-Vief, Lumley, Jain, & Heinze, 2003; Seidlitz & Diener, 1998). Additionally, many of these studies have been criticized for using self-report measures where answers may be affected by culturally imbued and gendered beliefs and may be less than reliable (Grossman & Wood, 1993) and men may be less likely to report feeling strong stereotypically feminine emotions like sadness because of learned culture-appropriate behavior (Belk & Snell, 1986).

However, these stereotypical and widespread beliefs that women are more emotional may affect how they are given feedback. In particular, if people are motivated to tell white lies out of empathy and compassion (Lupoli, Jampol, & Oveis, *in preparation*) then these beliefs about heightened emotionality may lead to more

protective behavior of the recipient's feelings or reactions. This protective behavior may translate into upward distortion of negative information in the form of white lies given more often to female recipients than male recipients.

Competence and Warmth: Ambivalent Sexism

Another common stereotype about women is that they are less competent than men (particularly in male-dominated domains, such as business or science), making them more likely to elicit empathy. Fiske, Cuddy, and Glick (2002) have detailed in their stereotype content model that the way people are perceived usually falls into one quadrant of two primary dimensions, *competence* and *warmth*. Stereotypically, status predicts high competence and competition predicts low warmth thus business executives are usually seen as highly competent, but cold, while elderly people are seen as incompetent but warm. The most common feeling toward those who fall low on competence and high on warmth is *pity.* Women are typically (or stereotypically) perceived as being warmer than men yet less competent, and invoking of a certain protective quality, almost in the way one would perceive and treat a baby.

This *benevolent* sexism, where women are seen as needing to be protected and cherished, is a paternalistic ideology and is unlike hostile sexism. Hostile sexists are more likely to believe that women are not as smart, and more likely to believe that women use their sexuality to entrap men. Benevolent sexism, on the other hand, does involve stereotyping women in culturally traditional and restricted roles, but also tends to elicit behaviors and attitudes that are primarily positive for the perceiver, such as helping behaviors and intimacy seeking (Cikara et al., 2009; Glick & Fiske, 1996). Both types of sexism, however, are predictive of gender inequality (Glick & Fiske,

2001) and while not intentionally disenfranchising, may contribute to paternalistic feelings toward women, especially those who are subordinates (in a workplace for example) (e.g., Vescio & Gervais, 2005).

Stereotyping may also be more common in situations where there is an asymmetrical power distribution and in most feedback situations, the direction of feedback is downward from a person in a higher power position toward a subordinate. Though inconclusive, there is evidence that powerful people show more implicit prejudice to out-groups (Fiske & Morling 1996; Guinote, Reese, & Wilkinson, 2009), pay less attention to individuating information about subordinates (Guinote & Phillips 2010), attend more to stereotype consistent information (Fiske, 1993), rely on more automatic cognition (Keltner et al., 2003), and are better liars (Carney et al., 2014).

Additionally, stereotyping may be particularly exacerbated by anxiety (e.g., Bodenhausen, 1993; Smith, Ingram, & Brehm, 1983; Wilder, 1993). Given that person-to-person feedback is an anxiety-provoking situation, people in power may be even more likely to act based on preconceived assumptions about women. These assumptions may even lead people to undertake otherwise unacceptable dishonest behavior.

Past & Present Research: Gender-Biased Language

There is a rich literature examining language, communication, and gender bias, and several researchers have demonstrated that people tend to speak or write about and evaluate women differently than men. Women are written about with more interpersonal warmth in recommendations while simultaneously evaluated differently than men on quantitative scales (Biernat, Tocci, & Williams, 2011). People also

withhold resources such as payment or promotions from women while simultaneously praising them verbally for their work (Vescio & Gervais, 2005). Additionally, there is a tendency to evaluate women using more gender-stereotypic criteria (e.g., "organized" rather than "competent," Guillemin, Holmstrom & Garvin, 1979).

This important research on gender-biased semantics has paved the way to understanding biased language in the workplace *about* women, but as of yet, researchers have not examined how objectively inaccurate communication *to* women could contribute to gender inequality in the workplace. In the present research, I differentiate the overarching proposal from past work by examining white lies in the *content* of information, and in how women are communicated *to* rather than communicated about.

The dominant theory concerning evaluation and women assumes that women are spoken about differently, because they are being evaluated on completely different scales and terms than men. The "shifting standards model" (Biernat, Manis, & Nelson, 1991; Biernat & Vescio, 2002; McCabe & Brannon, 2004) proposes that relative comparisons, and not objective ones, affect judgments. In particular, evaluation is subjective and depends on one's prior experience with or stereotypes of particular group members, not just the individual's objective performance and, unlike prior models of evaluation, does not assume that assessments are stable across time (Biernat & Manis, 2007).

Biernat and Manis' theory of "shifting standards" in evaluations of women suggests that the same evaluative traits are both interpreted and used in different ways for women and men. That is, when someone is said to have done a "good" job, the

standards for "good" are both relative (compared to other women) and qualitatively different (a woman's "good" may be equivalent to a man's "mediocre"). Additionally, stereotypes prompt lower reference points for these standards and thus lower minimum standards or entry criteria, such as being hired for a job (Biernat & Feugen, 2001). However, Biernat & Kobrynowicz (1997) found that while women were evaluated in terms of lower *minimum* competency standards than men (e.g., women do not have to be as competent as men in order to be hired), they were also evaluated with respect to higher ability standards (once hired) and had to work harder to prove they were competent.

This work has influenced how researchers understand stereotyping, evaluation and feedback for different categories of groups. However, the theory does not have anything to say about evaluations of men and women following a gender-blind, objective evaluation. Additionally, the work has implied that these differing evaluations and gender biases in feedback are primarily carried out through qualitative manipulation and interpretation of language. Though not in complete conflict with this theory, I argue that people miscommunicate to women on objectively quantitative terms and not just with qualitative language, thus suggesting that "shifting standards" of evaluation may be less at play than the choice to tell white lies to women based on stereotypes about female emotionality.

Summary of Literature and Argument

Past work has found that feedback can be emotionally difficult for both the person imparting negative information and the recipient of the feedback. Because this creates a situation of anxiety, one's focal goal of staying truthful while improving an

employee's productivity can be undermined by the alternate goal of avoiding the emotional distress of the recipient which in turn may lead to telling white lies about performance. White lies, however, may be more likely to be told to women because of stereotypes about their elevated emotionality and because of a benevolent yet misguided belief that women need protection from harm. Specifically, I predict that women will be told more white lies (not just evaluated differently) about their performance than men because people's stereotypical beliefs about women's emotionality are made salient by the difficulty for most people to give critical feedback.

Research Questions and Overview of the Present Research

There are several research questions that need to be answered in order to test and evaluate this hypothesis. Chapter 2 will address the question "Is there a gender bias in performance feedback and is it due to shifting standards?" and Studies 1a and 1b will test the prediction that women will be told more white lies than men during feedback, not just evaluated according to different standards.

Chapter 3 will address the question "Are people aware of the bias in themselves and others?" and Studies 2a and 2b will test the prediction that although people do not recognize gender bias in themselves, they do recognize it in others.

Chapter 4 will address the question "what are the mechanisms driving this bias?" and Study 3 will test the prediction that beliefs in stereotypes about women (sexism) will exacerbate telling white lies to women.

Chapter 5 will address the question "what are the problems with gender-biased feedback?" and Study 4 will test the prediction that women do not want to be told

white lies about their performance more than men, while Study 5 will test the hypothesis that white lies will result in negative motivational and performance consequences.

Finally, Chapter 6 will address the question of "how can we attenuate the effect and level the playing field?" and Study 6 will test the prediction that changing people's regulatory focus by motivating them to think about the their focal goal will attenuate the gender bias during feedback.

Following the study descriptions, in Chapter 7 the implications and future applications of the findings will be discussed, along with the limitations of the current work.[2]

[2] Please note that all work done is in collaboration with Dr. Vivian Zayas of Cornell University.

CHAPTER 2

What is the Performance Feedback Accuracy Bias?

In order to test our general hypothesis that women are more often the targets of white lies, through two studies, I will attempt to establish the phenomenon. In Study 1a, I will test the hypothesis that women are told more white lies than men. In Study 1b, I will demonstrate that this effect is likely not due to 'shifting standards' of evaluation but to actual lying on objective quantitative measures.

Study 1a

Are Women Told More White Lies about their Performance than Men?

The primary objective of Study 1a was to identify a gender bias in the accuracy of performance feedback and test the prediction that people would tell more white lies to women about their performance than men, despite having made lower objective evaluations of their work. As specified in the introduction, person-to-person feedback can be difficult as the person giving the feedback must fight the more immediate social urge to not hurt the recipient's feelings while simultaneously focusing on the broader goal of telling the truth so that the recipient may improve.

Thus, in this study, I attempted to simulate as much as possible the social

anxiety of giving someone feedback, while being able to control for several variables (such as attractiveness, interpersonal relationships, and physical behavior). I tested the main hypothesis in a laboratory setting where I could manipulate the gender of the recipient, and I used an online-chat paradigm in order to achieve the social connection and pressures associated with giving feedback.

Because I wanted to realistically simulate the physical and social process of giving someone feedback without having them come face-to-face with another person, participants took part in validated computer-generated chat simulation (modeled after North & Fiske, 2013) that evokes the feeling of a real-life feedback scenario. I also chose 'good writing' as the evaluation domain since academic achievement is of particular importance to college students at Cornell University, where this study took place. Participants evaluated essays (which were pre-tested for quality equivalence) of two other (computer generated) male and female targets and submitted their objective evaluations privately to the researcher. The participants then gave second evaluations directly to both targets as feedback. Objective quality of the essays was measured using the participant's quantitative assessments of both targets and white lies were measured comparing their quantitative feedback to these original objective evaluations. I hypothesized that participants would give upwardly distorted feedback to the female targets but not the male targets.

Another important factor in this study is our use of *quantitative* evaluations and feedback where previous studies examining feedback use *qualitative* measures. In this study, I predicted that participants would actively change their quantitative scores when giving them directly to the recipients, thus indicating that they were not just

verbally *sugar-coating* the truth, but actually bending qualitative information.

Method

Participants

Sixty-eight Cornell University undergraduate students were recruited through the internal social sciences recruitment system (SUSAN) in exchange for two course credits or $10 for one hour. Only two participants were excluded from the analysis because one was a former research assistant in the lab and also knew the researcher, and the other's data were lost during a system failure, leaving a total of 66 participants (39 female, 26 male, one answered 'other'). The mean age of participants was 20 years old, and of the participants who reported their ethnicity, 49% were Caucasian, 28% were Asian, 6% were African-American, 3% were South-East Asian/Indian, 1% were Hispanic, and 12% identified as 'other.' 37% of participants reported practicing religion, 76% of whom practiced Christianity, 8% Judaism, 8% Islam, 4% Buddhism, and 4% Hinduism.

Procedure.

Participants were brought into the laboratory room and seated in front of a computer, at which point they were told that they were going to be participating with two other students at the same time. They were informed that the experiment they were participating in was a joint project with Ithaca College intending to study how well people are able to perform "off-the-cuff" writing and to gauge their suitability for graduate school. They were also told that the other two students would be participating

remotely and that they would interact through an online chat session. In fact, the chat session was rigged such that responses from the "others" were pre-programmed using Qualtrics software to look real (Javascript paradigm modified from North & Fiske, 2013). This Ithaca College ruse was established in order to (a) explain why the other two students were not in the same room (e.g. Forsythe, Horowitz, Savin, & Sefton, 1994) and (b) increase the power asymmetry of the situation as Cornell students typically see themselves as superior to their neighboring Ithaca College students.

Participants were told that they would be taking part in a "writing project" where they would be randomly selected to be either the "writer" or the "judge" of an exemplary essay on why they should get into grad school, and that the best example would be selected and showcased as an example in a document for students prepping to write their admittance essays. After they logged into the computer experiment and selected "assign role," the computer automatically assigned each participant to the role of "judge." Participants were then asked to sign in to the "chat room" to establish the feeling of psychological connection and a social atmosphere.

Participants were then told that while waiting for the others to write and submit their essays, they should take five minutes and read through a set of criteria for good writing and the criteria page was pre-programmed to stay up on the screen for exactly five minutes. This way, all participants were on the same page detailing what to look for in an essay, while allowing for a believable amount of time to pass in which one might write an essay.

After the criteria for reading/writing phase had finished, participants were informed that the other two students had submitted their essays, and that the

participant should introduce him/herself by typing his or her initials into the chat server. The "initials only" rule was employed in order to keep any identifying information out of the first part of the experiment. In order to make the session as realistic as possible, there was a 'loading period' of two seconds between each interaction (programmed via JavaScript), and then the other two participants introduced themselves through their initials ("SB," and "AM"), in order to not reveal identifying information in the first phase of the experiment.

Essays. The two pre-programmed essays used were selected based on a previous pilot study that tested several different short essays for quality by a random sample of 78 participants from the Amazon Mechanical Turk (MTurk) participant database (Full data in Study 1b). The two essays selected were rated as equally bad writing on a scale of 0% (*very bad*) to 100% (*very good*), (M = 30%), equally likely to have been written by a male or female, on average identified to be between the age of 18 and 25, and similar enough in content and structure so that one did not stand out from the other. The essays were counterbalanced such that each appeared for SB and AM an equal number of times, and participants read them in counterbalanced order. After finishing their objective ratings of each of the essays, they submitted their answers and were told to move on to the next phase.

Phase 1: objective essay evaluations. Participants were then asked to objectively evaluate the writing of each of the targets on three different quantitative scales matching the original writing criteria rubric they had observed.

1. ***General Quality***. Participants were asked to indicate on a scale of *not good at all* (0) to *excellent* (100) their assessment of the overall quality of the essays.

2. ***Specific Criteria***. Modified from the Writing Assessment and Evaluation Rubrics Guide (McGraw-Hill Staff, 2000), participants were asked to evaluate both of the targets on the Focus, Organization, Logic, Support, and Mechanics of their writing, on a scale of *not at all (*0) to *completely (*100).

Writing Criteria Rubric

Focus *- How well do they stick to their topic (assigned: Graduate program)?*

Logic *- How good is their reasoning? Are their statements accurate? Do they connect their points so there is a clear flow from one idea to the next? Do they go beyond simple claims to develop or discuss the points they raise?*

Organization *- How well have they organized their presentation? Is there an introduction, a body, and a conclusion. Do they introduce the topic so that it is clear what they are writing about and what they intend to say about that topic? Do they offer their reader a closing section that pulls together their main arguments without simply repeating them?*

Support *- Do they offer specific examples to support the claims they make? Mention individuals, terms, specific passages, titles, etc. Do they indicate the sources of their information by using quotations marks?*

Mechanics *- Have they proofread their work and corrected all spelling or typographical errors? Have they written in complete sentences following standard rules of grammar?*

3. ***Willingness to recommend***. Participants were additionally asked how likely they would be to recommend the writing of the target, on a scale of *not at all (*0) to *completely (*100).

Writer characteristics measure. In order to asses whether or not participants would see female targets as possessing more stereotypically feminine characteristics (high in warmth, low in competence) after giving inaccurate feedback, they were asked to evaluate the writer of each essay on different traits, based on their reading of the essays. The traits that were used, stereotypically feminine and masculine traits, are those often used in research on sexism and are elaborated in the stereotype content

model (Fiske, Cuddy, & Glick, 2002). These included several items within the quadrant of *warmth* (warmth, agreeableness, dominance) and several within the quadrant of *competence* (competence, confidence, optimism about future, confidence, intelligence). The participants rated each of the target essays on each of the items on a scale of 0 to 100. I additionally hypothesized that participants would upwardly distort their perception of characteristically feminine traits for the female targets, thus revealing the traits that were implicated in their decision to lie.

Phase 2 (feedback & gender manipulation). After the participants had submitted their answers, they were then told that, as part of the writing project, writers would get a chance to improve their writing and that critical feedback would be helpful to them. It was explained that the student writers would have a chance to rewrite and resubmit their essays for consideration if they so chose, and if they made it into the top selection that they would win a small prize. This explanation was given to most closely mimic a real-world workplace or mentoring situation and to motivate the person in the evaluative position to give truthful and helpful feedback to the participants. A manipulation check was performed after the experiment was finished to make sure that this was sufficient motivation for the Cornell student participants (who value themselves as intelligent and good writers) to give helpful critique to the underperforming confederates.

Participants were then asked to re-enter the chat room, but this time to introduce themselves using their full given name, what year they were in college, and what from what school. This introduction not only served to increase the social and psychological connection between the participants but also allowed the researchers to

introduce the gender of the confederates via the subtle manipulation of first names. Thus, the other users introduced themselves as Sarah (instead of SB) and Andrew (instead of AB).[3] Additionally, they introduced themselves (slightly distinctly, and in a counterbalanced way) as freshman at Ithaca College. Participants were then asked to once again fill out the exact same quantitative evaluations as they had in the objective Time 1 (specific criteria and general quality) and submit them directly to each writer, Sarah and Andrew, respectively, over the computer-programmed chat server.

After they had submitted the feedback, participants were once again asked to rate each of the targets on the trait characteristics that they had done before the feedback phase, however this time around, they knew the gender of the targets. This was important to get a gendered measure to evaluate against their original 'objective' trait evaluations.

White lie awareness measure. After the completion of the main part of the experiment, participants were asked explicitly to answer "Yes," "No," or "Maybe" whether they had been objective when giving their feedback directly to the other students overall. Additionally, participants were asked to report what percentage of the truth they had told each of the targets, Andrew and Sarah, to examine whether there was any awareness of personal bias. Finally, participants were asked to indicate on a scale of 0%, *easy* to 100 %, *difficult* how difficult it had been giving feedback to Sarah and Andrew. These questions were added to determine whether any bias was due to deliberate lying and not just because participants' evaluations had actually changed

[3] The targets' names were chosen from the top 20 baby names of the 1980's and 1990's so as to match as much as possible popularity and familiarity and, thus reducing the possibility that participants would draw extraneous conclusions from the names themselves.

after discovering gender (e.g. *shifting standards*; Biernat et al., 2011). That is, it is possible that participants, once they learned the gender of the targets, would reevaluate the essays according to some gendered standard, and give their new, but still honest feedback to the targets. On the other hand, if participants admit to lying overall, or to one gender over the other, this would lend support to the hypothesis that participants tell more upwardly distorted information to women more than men.

Participants additionally answered a series of demographic questions, including their own gender, ethnicity, religiosity, political orientation, age, and a hypothesis check.[4]

Results

Preliminary analyses

Before analyzing the data for the main hypotheses, a repeated measure t-test was conducted to make sure that there was no quality difference between the ratings for Essay 1 and Essay 2 at Time 1 (objective evaluation phase), and I found that although the essays had been pretested for quality similarity, there was a significant quality difference during Time 1, such that Essay 1 was objectively rated as slightly worse ($M = 35.7$, $SD = 22.1$) than essay 2 ($M = 44.2$, $SD = 27.9$), $t(65) = -2.1$, $P < .05$. Because the essays were counterbalanced with target gender, however, I proceeded with the main analyses and this finding will be further examined in the discussion.

[4] Please note that there were some additional measures that were less relevant to the study than the ones reported here. Only the primary measures are reported in the main text body. Please contact the author if you are interested in the peripheral measures.

Principal Analyses

General quality. In order to see whether more white lies were told overall during Time 2 and whether feedback had been upwardly distorted to both targets regardless of gender, I collapsed over target gender to examine the main effect of Time (feedback phase). A multivariate repeated measures ANOVA analysis revealed that, collapsed over target gender, people gave significantly more favorable feedback to the targets during the Time 2 essay evaluation ($M = 45.6$, $SD = 25$) than they expressed privately at the Time 1 essay evaluation ($M = 39.6$, $SD = 25.7$), $F(1, 65) = 18.9$, $P < .001$.

Because Essay 1 was counterbalanced with Essay 2 across participants, and because participants did not actually find out the gender of the essay writer until after they had analyzed both essays at Time 1, instead of doing a 2 X 2 (Gender by Time) Repeated measures ANOVA, a new variable was created for the essay evaluation at Time 1 by averaging evaluations of Essay 1 and 2. A 1 X 2 Repeated Measures ANOVA was conducted to examine the difference between Time 1 essays ratings, and Time 2 essay ratings by gender.

As predicted, there was a significant interaction between 'target gender' and 'Time', where people gave more favorable general evaluations directly to Sarah ($M = 51.3$, $SD = 25.9$) at Time 2 than they privately expressed at Time 1 ($M = 39.9$, $SD = 18.8$), during objective evaluation. Whereas for Andrew, there was no difference between Time 2 ($M = 40.3$, $SD = 25.5$) and Time 1 ($M = 39.9$, $SD = 18.8$) evaluations, suggesting that participants were lying more to Sarah than to Andrew despite similar

objective evaluations, $F(1, 65) = 14.3$, $P < .001$, $\eta_p^2 = .18$. Pairwise comparisons and a paired t-test indicated that the difference between Time 1 and Time 2 was higher and significant for Sarah ($P < .001$), with a mean *positive* difference of 11.3 points on the scale, but not for Andrew, who had a mean difference of just .47 points on the scale, $t(65) = 2.9$, $P < .01$ (Fig. 1).

To see whether there were any interactions between the participant gender and the target gender on evaluations, a multivariate repeated measures ANOVA was conducted using participant gender as a between subjects variable. The data showed that there was no effect of participant gender on whether the female target was told more lies than the male target at Time 2, $F(2, 65) = .74$, $P = .47$, and both females and males were equally likely to upwardly distort their Time 2 feedback by approximately 10 points. However, pairwise comparisons revealed that female participants were significantly more generous to both targets than were male participants, $F(1,63) = 6.9$, $P < .01$. Female participants rated the general quality of both target's essays as approximately 10 percent higher ($M = 39.7$, $SD = 21.7$) in Time 1 than did male participants ($M = 30.4$, $SD = 22.1$) and at Time 2 rated the essays of both male and female targets ($M_{\text{male target}} = 47.3$, $SD = 25$; $M_{\text{female target}} = 56.2.x$, $SD = 24.6$) higher than did male participants ($M_{\text{male target}} = 30.80$, $SD = 23.4$; $M_{\text{female target}} = 44.8$, $SD = 26.7$).

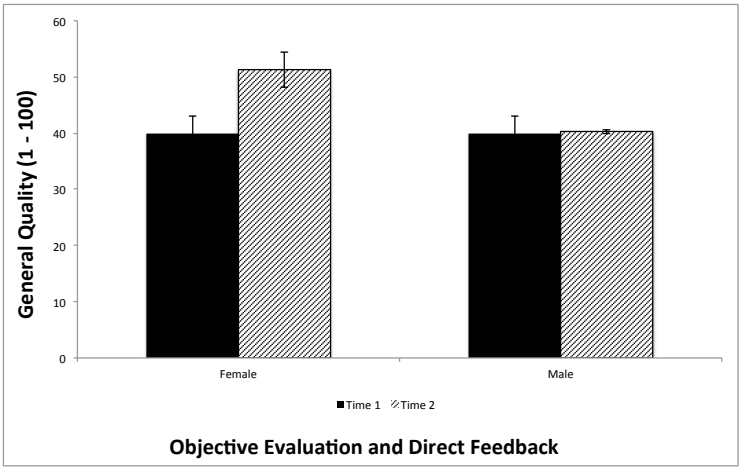


Figure 1. Bars represent participants' evaluations of General Quality of Essay at Time 1 and Time 2 for Male and Female Targets. The Time 1 series is an average of the two objective evaluations of the essays counterbalanced at Time 1, before Gender was introduced. The error bars represent the standard error of the mean.

Criteria quality. This effect replicated significantly when analyzing the five criteria: Focus, Logic, Organization, Structure, and Mechanics (FLOSM) together in a repeated measures multivariate ANOVA. A main effect of Time was significant, $F(5, 61) = 5$, $P < .001$, $\eta_p^2 = .29$ such that overall, during Time 2, higher quality ratings than during Time 1 were given as feedback directly to all participants, regardless of gender. And, as predicted, there was a significant interaction between Time and Target Gender, $F(5, 61) = 6.4$ $P<.001$, $\eta_p^2 = .35$, such that, compared to non-gendered criteria scores at Time 1 ($M = 50.9$, $SD = 17.9$), female targets received elevated scores at Time 2 ($M = 57.1$, $SD = 22.8$) while male targets at Time 2 did not ($M = 46.3$, $SD = 22.1$).

Univariate tests for each of the five criteria revealed that people lied to Sarah at Time 2 significantly more than to Andrew for all of the criteria except Focus (see Table 1 for means and significance probabilities).

Table 1

Specific writing criteria means and standard deviations for Target Gender X Time

Criteria	Time 1 (Objective Evaluation)	Time 2 (Direct Feedback)	
		Target Gender	
	Essay	Female	Male
	Mean (SD)	Mean (SD)	Mean (SD)
Focus	61.7 (19.9)	66.2 (23.3)	57.8 (25.5)
Logic***	48.1 (19.3)	58.7 (24.8)	48.9 (25.3)
Organization**	51.5 (19.8)	59.1 (24.7)	47.8 (23.5)
Support**	45.9 (20.6)	53.2 (26.6)	43.4 (24.5)
Mechanics***	37.9 (19.6)	47.9 (31)	33.5 (27.3)

Note. Repeated Measures GLM Analysis of each Criterion. †p < .10. *p < .05. ** p < .01. *** p < .001.

Likelihood to recommend. As with the previous measures, using a repeated measures multivariate ANOVA, I examined both the main effect of time overall, and the interaction between time and gender. As I had predicted, there was a significant main effect of time such that participants upwardly distorted their evaluations at Time 2 (M = 38.5, SD = 28.2) from Time 1 (M = 28.8, SD = 19.8), $F(1, 65) = 60$, $P < .001$, $\eta_p^2 = .48$.

Additionally, and mirroring the previous two measures of upward distortion, analysis revealed that the three evaluations were significantly different $F(2, 65) = 34.9$, $P < .001$, $\eta_p^2 = .52$. Pairwise comparisons indicated the difference between Phase 1 and Phase 2 was significantly higher for the female target ($P < .001$), such that, during feedback, participants indicated they were willing to recommend her work a mean *positive* difference of 13.9 points on the scale. However, this was not the case for the male target, for whom participants were willing to recommend only 1.8 points

on the scale. A paired t-test revealed that participants lied significantly more to the female target at time 2 (M = 44.6, SD = 29.5) than to the male target (M = 32.5, SD = 26.9), t (65) = 2.6, P < .01.

Trait Characteristics. As with the essays at Time 1, I averaged the traits from Essay 1 and Essay 2 so that this objective evaluation could be compared with the gendered evaluation of the targets at Time 2. I then conducted a repeated measures ANOVA using these three evaluations for each trait to examine whether there were any target differences at Time 2. There were significant differences between Time 1 and Time 2 for five of the seven traits, and one trait approached significance. However, nearly all of the traits were elevated for the female target as opposed to the male target, save the traits of 'dominance' (which was nearly significantly lower than Time 1) and 'warmth' (which was no different) (Table 2).

Table 2

Trait Characteristics of Male and Female Targets at Time 1 and Time 2

Criteria	Time 1 (Objective Evaluation)	Time 2 (Gendered Evaluation)	
		Target Gender	
	Essay	Female	Male
	Mean (SD)	Mean (SD)	Mean (SD)
Future Optimistic**	46.3 (20.7)	55.7 (25.3)	47 (23.6)
Smart**	45.9 (16.0)	55.4 (20.6)	50.7 (21.0)
Dominant†	56.4 (13.7)	52.2 (18.0)	52.5 (16.0)
Warm	58.0 (13.0)	58.9 (14.4)	59.9 (15.4)
Agreeable**	55.8 (14.9)	61.4 (14.2)	61.1 (17.9)
Competent***	43.3 (17.5)	56.4 (21.4)	49.8 (20.9)
Confident*	65.4 (13.1)	60.2 (17.9)	63.2 (15.0)

Note. †p < .10. *p < .05. ** p < .01. *** p < .001.

Bias Blind Spot. To understand whether people were truly shifting their

standards or whether they knew to some extent that they had lied during Time 2, frequencies were measured of people answering “yes,” “no,” and “maybe” to the question ‘did you give inaccurate feedback during phase two?’ Of the 66 participants, 19 answered “No,” 43 answered “Yes,” and four answered “Maybe.” A Chi Square Test revealed that a significant majority of participants (65%) reported lying overall during Phase 2, X^2 (2, 65) = 35.18, $p < .001$. Additionally, in order to evaluate whether or not participants also recognized that they were gender-biased in their evaluations at Time 2, a paired t-test analysis was used to examine the difference between the percent that participants reported lying to each of the target genders. As predicted, participants did not report significantly different percentages of truthfulness toward Sarah (M =77.9%) or Andrew (M = 75.9%) during Time 2, t(54) = .67, P = .50.

To assess whether participants who had admitted they had lied had indeed upwardly distorted the participants’ scores during the Time 2 feedback phase, I used the Time 2 – Time 1 difference scores for each participant in an ANOVA with “admitted to lying” (a “yes,” “no,” “maybe” discrete variable) as the between subjects predictor. The results indicated that those who had admitted to lying overall were not upwardly distorting significantly more (M = 4.2, SD = 30.4) than those who had not (M = 19. 1, SD = 39.5), or those who had indicated maybe (M = 12, SD = 28.7), F(2, 64) = 1.3, p = .26.

To assess whether there were any interactions between Target Gender and “admitted liars” I analyzed the Time 2 – Time 1 difference scores for each participant in a repeated measures ANOVA with “admitted to lying” (a “yes,” “no,” “maybe” discrete variable) as the between subjects predictor. The results indicated that the

interaction between Target Gender and "admitted liars" was also not significant, $F(2, 64) = 2.3$, $p = .1$, suggesting that even those who stated they did not lie, lied as much as those who said they did.

Similarly, when using participants' indicated percentage of truthfulness to the female and male target, there was no relationship between indicated truthfulness and actual upward distortion for male targets, $r = -.10$, $p = .46$, $n = 55$, nor for female targets, $r = .12$, $p = .38$, $n = 55$, suggesting that participants are unaware (or not reporting) that they are lying.

Finally, after analyzing the degree of difficulty in giving feedback to each target using a paired t-test, there was no difference in reported difficulty of feedback toward Sarah ($M = 32.4\%$, $SD = 25.4$) versus Andrew ($M = 36.5\%$, $SD = 29.7$), $t(65) = 1.1$, $P = .27$.

Discussion

Study 1 provides compelling support for the hypothesis that women are given more inaccurate feedback than their male counterparts about under-performance. In a near life-like simulation of person-to-person feedback, participants told a female target that she had done more than 11% better on an essay writing task than they had originally evaluated her, compared to a male target to whom they gave the same, truthful feedback as originally evaluated. The results also add support to the possibility that it was not "shifting standards" that was driving this effect.[5] That is, because the majority of participants admitted to lying, it appears that the reason feedback differed

[5] The alternative account of shifting standards will be discussed further and the end of the chapter and experimental data from a preliminary study analyzed to argue against this account.

from objective evaluations on the quantitative measures was not due to reassessment of the essays after gender was revealed. However, people did appear to either not be aware of or did not report that they themselves might be gender-biased or have lied more to the female rather than the male target. And even those who indicated that they had not lied, did not lie significantly less than those who had admitted to it.

However, contrary to the original hypothesis that after giving direct feedback people would elevate the more feminine traits (like warmth) of the female target while lowering the more masculine traits (like competence), the results suggest that nearly all of the traits except dominance were elevated for women and not for men. There are a few possible explanations for why this might be the case. First, people often feel cognitive dissonance (Festinger, 1957) after telling a white lie, and as Argo & Shiv (2011) have noted, the teller attempts to reduce the dissonance of lying by engaging in a behavior that favors the wrongdoer, such as prosocial act. In this case, the act of elevating their evaluations of the female target's traits may be an attempt to reduce the dissonance. Another possible explanation is self-perception theory (Bem, 1967), whereby the participants may examine, post-hoc (perhaps not consciously), their decision to upwardly distort information, and therefore assume that they must think better of the target.

Though I do not have experimental data to rule out self-perception as possible reason why people are elevating all traits of only the female targets, it is the least likely of explanations, seeing how the majority of people admitted to lying and upwardly distorting their feedback. Thus they acknowledge that there is an 'objective' truth about their original evaluations. Another possibility is that participants merely

forgot or did not realize that their trait evaluations were not being passed on to the targets, and so they carried on psychologically with biased answers. If it is the case in subsequent studies that these findings hold true then that 'error' explanation can be ruled out.

The final possible explanation may be more tied in with the general hypothesis of this project, and may reside in the motivation to tell white lies to women. That is being nice or keeping someone from harm may not be the only motivation behind telling white lies (and none of which have been examined thus far). Another motivation may be future oriented, and related to feelings of warmth and compassion, but employed as a 'motivator' for intended targets. Participants may tell white lies more to women, not just because they pity them, or because they do not think they can handle the truth (though these are potentially part of the explanation), but also because they believe that telling them white lies will motivate them to improve or give them a step up (perhaps because they believe the world is unfair and sexist). This may lead to the participant's belief that he or she has just *invested* in the target, and when asked about the person's competence, the participants are forward-thinking in their assessments. In short, the elevated trait evaluations are less of evaluations at all and more of wishful thinking. There will be a subsequent lengthier discussion of this kind of 'gendered affirmative action' in the context of benevolent sexism in the general discussion.

Finally, one of the most important findings from this study is that the results suggest that women are more often the targets of upwardly distorted feedback and demonstrates for the first time that people will upwardly distort even concrete,

quantifiable information, not just "sugarcoat" language. These quantifiable changes indicate a possibility that people are making a conscious choice to lie.[6] However, though it appears that people are indeed lying, and not just changing their evaluations of the female a target's essay once finding out she was female, there is still the possibility that the upward distortions were a case of "shifting standards" (Biernat & Kobrynowicz, 1997). Thus in Study 1b, I address this alternate explanation.

[6] This concept that people are making a conscious choice to lie, should and is being examined, though there are no results yet that speak to the answer. A study seeing if people will go out of their way to tell white lies even if there is a cost to themselves (e.g. their time and effort) is in preparation stages.

Study 1b

A Case Against the "Shifting Standards" Alternative Explanation

Study 1a was designed to test the prediction that people would lie more to women than to men about their objective performance. Though the data did suggest that people were modifying their original objective ratings in order to give nicer feedback, it is difficult to tell with this data alone whether they were actually lying or whether standards for evaluations were just shifting to accommodate this new gender information (Biernat & Manis, 2007), particularly since women are stereotypically expected to be better writers than men (Swim, Aikin, Hall, & Hunter, 1995).

However, there is a possibility that the upward distortions were due to "shifting standards" in the way that people evaluate women and men. That is, people tend to evaluate women on different types of scales than men (often relative to other women or their prior experience with women). An argument could be made that even though the Time 1 ratings were 'gender-free,' because people often assign gender and other attributes automatically based on only thin bits of information (Ambady, Hallahan, & Conner, 1999; Ambady & Rosenthal, 1992; Thompson & Murachver, 2001), perhaps the Time 1 essays were actually evaluated according to relative gender standards. That is, if participants read the two poorly written essays and had assumed the writers were likely to be male and then found out one was actually female, perhaps instead of *lying*, what they were doing was re-evaluating their standards to fit the mold of *female writer = better*.

I hypothesize that shifting standards do in fact play a part in evaluations of quality, however, not in the direction that would support an alternative explanation of the upward distortion in Study 1a. I hypothesize that because women are stereotypically expected to be good writers, the quality of female-authored essays would be evaluated compared relative to other women and poorly written essays would be evaluated as worse than they would have if they had been male-authored. If this was the case it would mean that "shifting standards" would not be an explanation for the elevated ratings for the essays during Time 2 in Study 1a since Time 2 ratings should be lowered after learning the writer was female.

To test this possibility, I analyzed the four essays that I had pretested for Study 1a. For this pre-test I had asked participants to rate four essays of varying degrees of poor quality (as tested in yet another pre-test), however I varied the gender of the authors between subjects. [7] Because people tend to believe that women are better writers, people should shift their standards of evaluation and rate the female-authored essays that are poorly written relative to other females and subsequently rate them as worse than the poorly written male-authored essays. And because in Study 1a, female's scores were elevated, not lowered during feedback, this would indicate that "shifting standards" do not explain Study 1a's findings. Thus I expected in Study 1b, that participants would rate female-authored essays that were poorly written as worse than male-authored essays that were poorly written.

[7] I asked participants to read four essays and indicate their perceptions of the authors, including assumed gender, and then to assess writing quality. Similar to previous studies (Swim et al., 1995), people indicated that the essays of worse quality were slightly more likely to be male.

Method

Participants

80 participants were recruited through Amazon's Mechanical Turk (MTurk) crowdsourcing website and compensated $0.05 per minute. Two participants were eliminated for not completing the survey. 40 (52%) of the participants were female, and 37 (48%) males responding, 89% between the ages of 20 and 40, and 11% between the ages of 40 and 60. Additionally, 51% of participants reported their highest educational level to be college, with 10% reporting high school, and 39% reporting some graduate school.

Procedure

All participants were administered a survey powered by the Qualtrics survey platform. Participants were told that they would be evaluating an essay that had been written by a fellow MTurk worker, and that I would ask them to report to us on the quality of the essay's writing. I varied the gender of the authors of the essays to examine whether gender would change quality ratings of the essays.

Materials. I used four different essays that had been written by the research assistants for this project, who had written with the intention of writing a 'poor quality' essay by a college aged-student. These essays were written so that the two most similar ones could be used for the Study 1a materials. Each essay was similar in length, content, grammar, and quantity of spelling mistakes (each essay contained the same number of mistakes or misspellings (five). The topic of the essay was "why I should go to graduate school." The research assistants had been provided with an example of a 'good' essay (written by the author of this manuscript) upon which they

could model the 'bad' essays they would write. An example of the 'good' essay and a 'bad' essay are displayed below.

'Good' essay example:

> *"I believe that I would be an excellent candidate for graduate school for several reasons. First, I have always appreciated thinking deeply about big questions, and I truly love research. When I was growing up I would spend most of my free time enjoying books and conversing with my peers about big problems in society. Second, I have excelled at my studies and have consistently performed at the top of all my classes, earning praise for my academic skills. Third, I am passionate about being able to contribute to society and I believe that a graduate education can help me broaden and sharpen my knowledge in order to do so. I know that graduate school will be challenging but as Epictetus said: "Difficulties are things that show a person what they are." I am ready for the challenge and the fulfillment of a graduate education."*

'Bad' essay example (mistakes are underlined in this manuscript only):

> *"I think I would be a really great person for a graduate program because I love to work and I'm pretty much at the top of my game all of the time. I have learned alot of skills that I can apply to things, and once I helped build a doghouse for my neighbor, and she thought I could go far in life because of the fact that I did a really great job on it and because of my pacience. So as Einstein once said, it doesn't matter what you do as long as you do it well, and I stand by that every day pretty much. I also get decent grades for what Im interested in doing with my life. Also, I am a swimmer and in fact I have won numerous competitions at my level in high school. In fact, I also really love to ski too. Someday I would like to be a pro at one of these things. Maybe if I practised every day I could go to the big leages!"*

Gender Manipulation. The design was a between subject design, and each participant was presented with one of the four 'bad' essays, randomized per participant. However, half of the participants were informed that the author of the essay was male through revealing the author's name (Michael) and the other half were informed that the author of the essay was female (Lauren) (the essays and names

presented were counterbalanced). They were also told that the person who wrote the essay was a freshman in college and wrote this sample "in a few minutes" in order to make sure that participants were under the same assumptions about the age and education level of the author.

Quality measures. Participants were then asked to read the essay and evaluate its quality. Participants were asked to grade the essays using the American grading system of F (*poor*) through A+ (*excellent*), or 13 points, for the question: (a) "what is the general quality of the essay?" They were additionally asked to (b) evaluate the essay by indicating how "the quality of the essay compares to your expectations of what a good statement should look like" on a substantive Likert scale consisting of three scale points from "*worse than expectations"*, "*same as expectations*", and "*better than expectations*"; finally they were asked to indicate (c) "how likely would you be to recommend this essay?" on a scale of 0 (*not at all*) to 100% (*definitely*) likely. Demographics were then recorded.

Results

Before examining the manipulation, each of the 'bad' essays was compared for quality collapsing across gender. A univariate ANOVA revealed that the essays were significantly different from each other, $F(3, 77) = 6.6$, $P < .001$. Essay 2 was considered the worst, with a mean of 4.9 (About a C-). Essay 1 got a score of 6.2 (C), Essay 3 got a score of 6.7 (C+) and essay 4 was considered the best with a score of 8.4 (B-).

Collapsing over all four essays, there was also a main effect of Gender on

essay quality, F (1, 77), 3.9, P = .05, such that Lauren's essays were evaluated worse overall (M = 5.8 or C) than Michael's (M = 7 or C+). For the "comparison to expectations" measure, an ANOVA revealed that Lauren's essays were not evaluated as significantly worse than expected (M = 1.3) compared to Michael's (M = 1.4). Finally, participants were not significantly more likely to recommend one gender's essay over the other (P = .2), but the data was trending in the predicted direction with Lauren's essays somewhat "less likely" (M = 24%) to be recommended than Michael's (M = 32%).

In order to examine the hypothesis that women would be punished more for writing a worse essay, I ran the following analyses using only the data (n = 43) from the very worst essay (Essay 2, average grade of C-) and the very best essay from the collection of 'bad' essays (Essay 4, average grade of B-). Then, using a univariate ANOVA, I compared the two essays crossed with author gender to see if there were significant interactions between perceived Essay Quality and Gender. The interaction was not significant, F (1, 142) = .5, P = .4. However, it appeared that for the very bad essay (Essay 2) Lauren received a worse quality rating (M = 4.4, SD = 2.1) than Michael was (M = 5.5, SD = 2.2), whereas Lauren and Michael were evaluated similarly (M = 8.4, SD= 2.3 & M = 8.4, SD = 2.4, respectively) on the 'good' 'bad' essay (Essay 4).

Similar results were found for the measure "how does the essay compare to your expectations about what a good statement should look like?" Though again the interaction between Gender and Essay was not significant, F (1, 42) = 1.5, P = .2, It appeared that for the 'bad' essay, on a scale of 1 (worse than expected) to 3 (better

than expected) Lauren's version did not do significantly worse at meeting expectations (M = 1, SD = 0), than Michael's (M = 1.3, SD = .45), nor was not the case for the 'good' 'bad' essays (M = 1.8(.8) & 1.7(.5) respectively).

Finally, for the measure, on a scale of 0 to 100, "how likely would you be to recommend this essay to be used as a good example of off the cuff writing for students applying to grad school?" a univariate ANOVA revealed that there was a nearly significant interaction between author Gender and Essay, F (1, 42) = 3.6, P = .06, such that for the 'bad' essay, participants were less willing to recommend Lauren's version (M = 8, SD = 13.9) than Michael's (M = 18.9, SD = 23.4). However, the opposite was true for the 'good' 'bad' essay, with participants more willing to recommend Lauren's essay (M = 58.8, SD =16.6) over Michael's (M = 46.7, SD = 23.9) (Fig. 2).

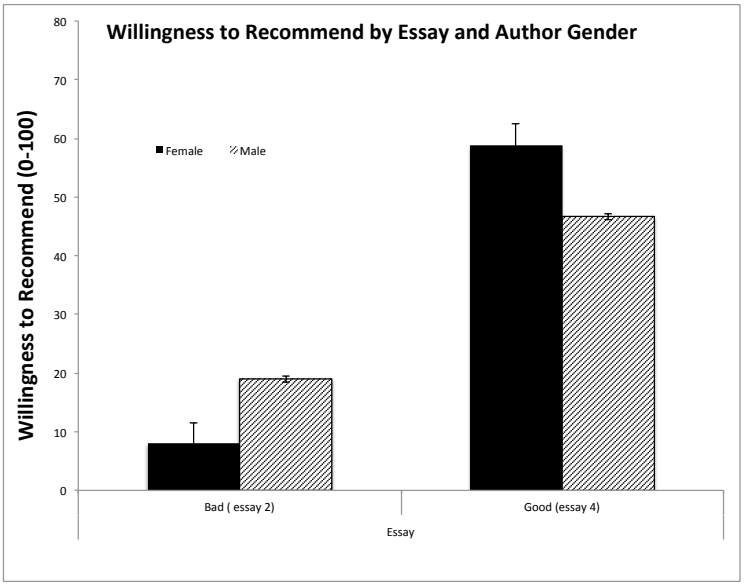


Figure 2. Chart represents the marginally significant interaction, F (1, 42) = 3.6, P = .06, between Author Gender X Essay Quality on the willingness to recommend it as an example graduate student statement. Bars represent participants' willingness to recommend the worst essay and the best essay when written by a male or female Author. The results suggest that female authors are punished more for writing very

bad essays than are male authors. The error bars represent the standard error of the mean.

Discussion

In Study 1b, participants rated poorly written essays that had been authored by females to be worse and recommended them less often than the same essays that had been authored by males. That is, it appeared the female authors were being punished for having written a bad essay, possibly because the standards for female writing are higher than for male writing.

The results were mixed in terms of significance (in particular the Quality evaluation appeared to be trending, but was not significant) however, this may be due to a underpowered sample as the study was meant to be a pre-test to make sure that the essays used for Study 1a were not extremely different. To check whether our non-significant results for the Quality evaluations were due to a lack of statistical power, I conducted a power analysis using the software G*Power (Faul & Erdfelder, 1992; for a full description see Erdfelder, Faul, & Buchner, 1996). With power (1 - β) set at 0.80, α = .05, two-tailed, and a medium expected effect size of Cohen's *f* = .25 (which was very similar to the observed effect size of *f* = .23), the a-priori analysis yielded a needed sample size of n = 128, with a critical F of 3.92 in order for group differences to reach statistical significance at the .05 level. Thus, it is possible that our negative findings could be attributed to a limited sample size.

Though not all of the measures yielded significant results, those that were significant suggest that if the author of an essay is a female, and wrote a 'bad' essay,

she would be judged even more harshly than a male who wrote the same essay. The essay that was considered the best showed no gender difference, while the essay that was considered the worst showed a more dramatic gender difference, especially in relation to participants' willingness to recommend the essays as good examples. Future studies should vary the quality of the essays in a more controlled manner, as all the essays used in this study were taken from the relatively 'bad' selection. In addition, follow-up studies should examine whether women are punished for not living up to stereotypical expectations of their writing.

In regards to the former argument about "shifting standards", the results indicate that this theory is now less likely to be an explanation of the results of Study 1a. That is, if participants found out that the writer of the 'bad' essay at Time 2 was female, a "shifting standards" account would predict that she would be punished for having written such a bad essay with even lower scores, or at least lower than her male counterpart's. That this was not the case, but that instead her Time 2 feedback was elevated despite these results suggesting she should be punished, strengthens our argument that participants are indeed distorting their answers, not just reevaluating them.

CHAPTER 3

Study 2a

Do People Believe That Others are Gender-Biased in their Feedback?

Participants in Study 1a reported that though they did upwardly distort feedback over all, they believed they were not giving biased feedback to women more often than to men, and that they had been equally truthful to both female and male targets. In reality however, the data suggest that they were giving more biased feedback to woman than to men overall. Several possibilities exist for why people reported not believing they were gender-biased, including that they were unaware that they were acting in a gender biased way, or that they were aware of their own biased actions but felt uncomfortable stating so. In order to better address this question, and to solidify the likelihood that a feedback bias does exist for women, the intention of Study 2a was to examine the possibility that gender bias during feedback is familiar and commonly understood to exist. If this is the case, then it may be more likely that people are consciously or subconsciously concealing their own biases. Additionally, if people do believe that women are told more lies than are men about their performance, this also strengthens the existence of the bias effect found in Study 1a.

The logic guiding Study 2a stemmed from past work on "bias blind spots" (Pronin, Lin, & Ross, 2002; Pronin, Gilovich, & Ross, 2004), which demonstrated that people may be entirely capable of recognizing bias in others but either do not or do not

want to recognize the same flaws or biases in their own behavior. I hypothesized that though participants in Study 1a did not recognize their own bias, they may recognize this same bias in others. Additionally, if the effect that was implied by Study 1a really is as robust as the results seemed to indicate, then there is also a high likelihood that people will recognize that in the real world women are communicated to differently than are men. That is, I expected that if people were to observe someone else giving an employee feedback that they would be able to see the difference between biased and unbiased feedback.

For Study 2a I used a business workplace scenario – a domain that is typically associated with masculine competence, and where women are judged to be less competent than men– to see whether women would still be told white lies. I additionally varied the setting so as to avoid using the same domain as Study 1a and used a task that participants would evaluate that did not stereotypically associate women with competence (as had essay writing).

In this study, instead of manipulating the gender of the target, I asked participants to judge the target's gender as a function of the accuracy of the feedback that a manager had given an underperforming employee. I hypothesized that participants who had witnessed a manager giving nicer but less accurate feedback to an employee would assume that the employee was female rather than male. Additionally, I measured individual levels of sexism to examine whether stereotypical views about women would moderate the results; I expected that people who held sexist beliefs might be less able to recognize this bias in others, as they may be more likely to deny that sexism exists at all (Swim et al., 1995).

Method

Participants

One hundred and thirty eight Cornell University undergraduate students (80 female, 55 male, 3 did not indicate), mean age 20 years old, were recruited via the internal social sciences recruitment system (SUSAN) in exchange for one course credit or $5 for a half hour. Three participants were excluded from the analysis because they did not complete the main survey, leaving 135 participants. For this study participants did not indicate their ethnicity.

Procedure

After participants were seated at a desk and had given consent, they were asked to fill out two different surveys, the experimental survey and one assessing individual differences in sexism. To make sure participants did not link the surveys and so that the answers on one would not affect the other, participants were told that the two surveys were from two different researchers and were given to them together in order to fill their 30 minute time slot. The order of the two surveys was counterbalanced to counteract any order effects.

Feedback accuracy manipulation. In the main experimental survey, participants read a scenario about a manager of a company who, as part of the manager's job, has to give monthly feedback to employees. They were also told that the goal was to give constructive feedback to the employee, for the good of the

company, and that the manager generally got along with all of the employees in the workplace (e.g., talking to them in the hallways, seeing them at lunch, etc.). I added this affability characteristic so as to represent a mostly normal, well-functioning workplace, while still making clear the need for constructive feedback. Participants then read a clear statement about an (unnamed) employee who was seriously underperforming:

> *"An employee who has been on the team for under six months is pleasant, however, upon reviewing their work progress the manager finds that it is just not up to par. In fact, the quality of the work is significantly worse than some of the other employees at the company and the manager is very disappointed with the employees work. If it continues this way, there may be repercussions including having to let that person go. Ultimately, the Manager must decide what kind of feedback will be the most constructive."*

After they had read about the manager's task they were told that the manager had six qualitative feedback options from which to choose, and were shown all six statements that had been written in varying degrees of truthfulness. The statements were numbered labeled (a) through (f) and ranged from the most truthful statement (which was also the harshest) up through the least truthful statement (which was also the nicest). All six of the statements are shown below:

> *(a) I am very disappointed in your performance so far: The quality of your work is much worse than the other employees and I expect you to do much better and work much harder. If you don't work harder and do better work you will most likely be fired.*
>
> *(b) I'm disappointed in your performance so far: The quality of your work is not quite as good as some of the other employees and I expect you to do a little better and try to work a bit harder. If you don't show a little bit of progress it might put your position in jeopardy.*
>
> *(c) I can't say that I'm extremely impressed with your performance so*

far: compared to the other employees the quality of your work could be a slightly better and I hope you'll progress a little more and you'll keep trying to work hard over time. I wouldn't want your position to suffer but I'm sure you'll keep trying hard.

(d) I'm not entirely sure what to think of your performance so far: You seem to be doing ok compared to the other employees and I'm confidant that you will keep trying hard and progressing over time. I think that keeping up your work ethic and the quality of work will result in you maintaining your position for some time.

(e) I'm pleased with your performance so far: You are doing fine compared to the other employees so just keep on trying hard and keep up the good work. I'm fairly confident that you'll keep trying and the quality of your work will get even better, resulting in good things regarding your position.

(f) I'm very pleased with your performance so far: You are doing just great compared to other employees so just keep on working hard and doing good quality work. I'm confident you'll keep working as hard as you do and performing even better, which could even result in a promotion of your position.

Participants were told that only one of these statements had been chosen by the manager to give to the employee. The feedback statement that the manager had chosen had been clearly marked on the survey, with a check mark next to the chosen statement. Feedback choices were randomized between participants so that all six options (conditions) were equally presented.

Assumption of employee gender measure. After reading about the manager and observing which feedback statement that manger had chosen, participants were asked to guess, based on the statement chosen, which gender the underperforming employee might be. After (or before, depending on survey order) guessing the gender of the employee, participants were told that the second experiment would be about romantic relationships between men and women.

Sexism measure. Participants were then given the full Ambivalent Sexism Inventory (ASI, Glick & Fiske, 1996; included in Appendix), a 26-item survey measuring stereotypical beliefs about women on a continuous scale of one (*do not at all agree*) to five (*completely agree*). When participants were finished with both surveys, they then filled out some basic demographic information, and were debriefed.

Results

Preliminary analysis

For ease of interpretation, I first categorized the ordinal feedback statements into two nominal categories, with the *one* through *three* statements categorized as "Truth" and the *four* through *six* statements categorized as "Lie."

Gender Assumptions

I then performed a Chi Square proportions analysis to see if participants were more likely to assume one gender over the other given the type of feedback the manager had selected. As I had predicted, participants were significantly more likely to guess that the employee was a woman if she had been lied to by the manager (75%), instead of told the truth (21%), Whereas participants were more likely to assume the employee was male when told the truth (79%) than when lied to (25%), X^2 (1, 137) = 26.7, P < .01 (Fig. 3). Analyses using feedback accuracy as a continuous variable in a binary logistic regression were also significant: $X^2(1, 137) = 15.2$, $p < .001$.

Sexism

Sexism (as determined by the ASI) did not interact significantly with feedback

choice ($p > .1$) however, it seemed to be trending such that participants who scored higher overall on the sexism inventory (M= 45%) were slightly less likely to think that the employee who had been lied to was a woman, compared to non-sexists (M= 55%). None of the other individual variables, such as participant gender, interacted with the effect. Also, no order effects were found for survey administration.

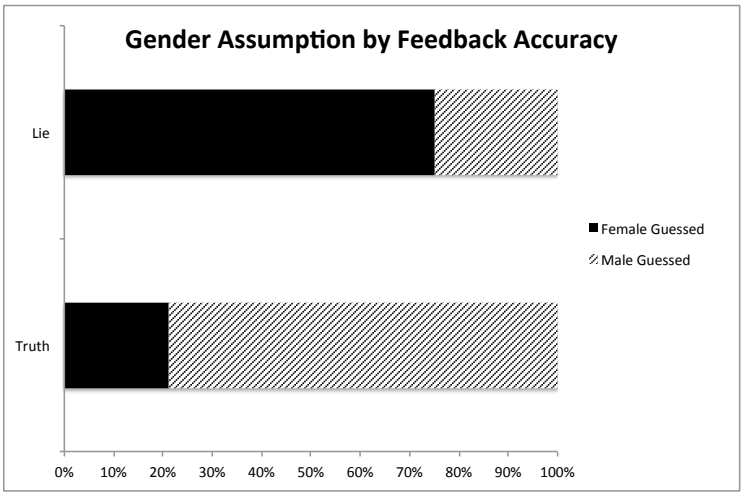


Figure 3. Represents the proportion of participants assuming an employee to be male or female depending on accuracy of feedback given. The black bars represent the proportion of people guessing the employee to be female, and the striped bars are the proportion of people guessing the employee to be male, in the categories of truthful feedback and white lies. The data reveals that when the feedback is inaccurate, the employee is more often guessed to be female and the opposite is true for accurate feedback.

Discussion

Study 2a revealed that as predicted, although people do not seem to be aware of their own bias (Study 1a) they do appear to recognize that other people are biased towards women when giving feedback in the workplace. However, though it seems that this bias is recognized as a common phenomenon, it is unclear from the data whether people really do have a bias blind spot (perhaps a motivated one), or whether

the people who are the most biased are also the least likely to see gender bias in the world at large.

That people who scored high on the sexism inventory were also less likely to assume that the target of biased feedback was female, though not significant, seems to be hinting at this possibility – that the people who are most least aware of the bias in the world are also the most likely to be biased themselves. This makes sense when considering that much of the way that sexism is perpetrated in modern times is through bias denial, or a refusal to believe that women are not on the same playing field as men, and sexism is both maintained and perpetuated through much more covert behaviors and beliefs than it used to be (Swim et al., 1995). This question needs to be more deeply examined in follow-up experiments.

Additionally, although participants in Study 2a appeared to make the *behavioral* assumption that women are lied to more than men, that is, they identified the employee as female more often when the feedback was inaccurate, it is unclear from this study whether this assumption was explicitly tied to their beliefs about other people's beliefs and actions. That is, it is not clear from this study *why* participants made these gender assumptions and whether they truly believed that other people have a gender bias during performance feedback or whether they merely recognized it. Finally, it is unclear whether participants were assuming that the manager was male or female and whether this had an effect on the assumptions about employee gender. It could be the case that participants generally believed the manager to be male, and it is also possible that people generally believe that males would be more biased than females. Thus, in Study 2b I attempted to control for manager gender by manipulating

it and I attempted to clarify participants' beliefs about why they assumed the employee to be male or female by asking them to provide qualitative justifications for their choices.

Study 2b

Why do People Believe that Less Accurate Feedback is given to Women?

Study 2a revealed a general assumption that other people tell more white lies to women than to men. However, qualitative justifications for why this is this case were not collected from the participants making it difficult to know (a) whether people only implicitly recognized bias and (b) why participants had assumed that female employees were lied to more often. Thus in Study 2b, a similar study was conducted with the purpose of gathering qualitative justifications.

The rationale draws from our overarching hypothesis about why women may be told more white lies than men. Because women are stereotypically perceived as more emotional and because benevolent sexism predicts that women need protecting from the truth, people may assume that supervisors are telling white lies to protect women's feelings. Thus I predicted that when participants discussed their rationale for assuming the employee was female because she was given inaccurate feedback, they would use more affective language composed of more emotion words than for male employees.

Additionally, in this version, two elements of the design were changed: First, I varied the gender of the manager giving the feedback, and second, I had participants listen to audio recordings of the feedback instead of reading the statements as written text. The primary objective in this study was to replicate the effect of Study 2a and to collect qualitative justifications from the participants concerning why they had guessed the target of the feedback to be male or female. The secondary objectives included examining whether the

gender of the manager giving the feedback would moderate the findings.

Research shows that not only men hold stereotypical beliefs about women. Women can exhibit stereotypical thinking as well and are often just as likely to sanction benevolent sexism, arguably because they are rewarded for conforming to a patriarchal status quo (Glick & Fiske, 2001; Moya et al., 2007). Several studies have shown that women (and other members of low-status groups) often exhibit ambivalence toward their own group, and sometimes go so far as to express out-group favoritism (e.g., Glick & Fiske, 2001; Jost & Burgess, 2000; Major, 1994; Ridgeway, 2001).

However, the general population may believe that only men are more likely to show a white lies bias toward women. This may stem in part from a belief that women are hostile and competitive with each other, especially in male-dominated domains, whereas men are paternalistic (or sexually enterprising) toward female subordinates (Ashcraft & Pacanowsky, 1996). Thus, I predicted that people would be more likely to assume an employee was female based on inaccurate feedback particularly if the manager was male.

Specifically, I hypothesized that in study 2b the gender assumption effect of Study 2a would be replicated, and furthermore that participants would use more language characterized by negative emotions in their justifications for why they had assumed the employee was female after hearing upwardly distorted feedback. I predicted that the qualitative answers about why they had selected female as the employee gender would most often include references to negative emotions in an avoidant way. That is, I expected people to bring up the emotional states of female employees more overall than for male employees, and also more often in conjunction with negations of those states, as expressed through negative emotion words (e.g. “sadness,” “upset,” “emotional”) paired with avoidance words and negations (e.g.

"didn't," "won't," "avoid"). For example, I expected a prototypical participant's rationale to sound something like "The manager didn't want to make her upset," which is made up of both an emotion word ("upset") in conjunction with an avoidance word ("didn't"). Finally, I predicted that people would be most likely to judge others as biased toward women if the manager was male.

Method

Participants

One hundred and twenty six participants were recruited through MTurk's crowdsourcing website and compensated $0.05 per minute. 69 (55%) of participants were female, and 57 (45%) males responding; 73% between the ages of 20 and 40, and 27% between the ages of 40 and 60.

Procedure

All participants were administered a seven minute survey on the Qualtrics survey platform in which I manipulated both feedback and manager gender in a 2 (Feedback: Truth or White Lie) X 2 (Voice: male or female) between subjects design. The scenario and setup of the survey was exactly identical to Study 2a, however, instead of reading the feedback statement that the manager had chosen to give the unnamed employee, participants heard an audio recording of the feedback purportedly from the manager. Participants were randomly assigned to hear either a female or male voice, and heard only one of two feedback statements.

Manager gender and feedback manipulation. Two of the original feedback

statements from Study 2a were read and recorded by two people, one female and one male.[8] Each recording was edited in Adobe Audition CC to control for length of recording, volume, and pacing. Participants only heard one voice, reading one recording. Unlike Study 2a, participants did not see or hear the other possible feedback statements and they were only presented with a truthful statement or a white lie. However, before they were told what the manager had chosen, all participants saw a statement that the manager had *intended* to give, which was the most truthful statement that corresponded with Option 1 from the original study. The intention was not described as either deliberate or just an oversight, but it was made clear that the manager had wanted to be the most truthful that he or she could be, in order to ensure that participants would construe any deviation from this feedback as a distortion. All participants saw the following description:

> *This is the feedback that the manager had **intended** to convey to the under-performing employee:*
>
> *"I am very disappointed in your performance so far: The quality of your work is much worse than the other employees and I expect you to do much better and work much harder. If you don't work harder and do better work you will most likely be fired."*
>
> *Click next to hear what the manager actually ended up telling the employee.*

The other two manipulation statements corresponded to Statement 2 (more truthful) and Statement 4 (more white lie) in the original study (see statements below).

> They either heard a truthful statement (audio):
>
> *"I'm disappointed in your performance so far: The quality of your work is not quite as good as some of the other employees and I expect you to do a little*

[8] The people who contributed their voices were: Lily Jampol and Amit Kumar.

better and try to work a bit harder. If you don't show a little bit of progress it might put your position in jeopardy."

Or they heard a white lie statement (audio):

"I'm not entirely sure what to think of your performance so far: You seem to be doing ok compared to the other employees and I'm confidant that you will keep trying hard and progressing over time. I think that keeping up your work ethic and the quality of work will result in you maintaining your position for some time."

Gender assumption measure & qualification. Identical to Study 2a, participants were first asked, based on the feedback recording they had just heard, to guess the gender (male or female) of the employee who had just received feedback. These were the instructions the participants received:

"If you had to guess based on the feedback that you just heard from the manager, what gender would you say ***the employee*** *was?"*

Critically, they were then asked to provide a qualitative justification for why they had chosen that particular gender based on the feedback they heard. Participants were allowed to respond as fully as they wanted.

Results

Replication Analysis

Before analyzing the qualitative statements and the contribution of the manager gender to the general effect, using a Chi Square proportions test, I analyzed the effect of feedback condition on employee gender assumption, and found that 58.8% of participants in the white lies condition believed the employee to be female, whereas only 44.8% percent of participants in the truth condition believed the employee to be female. However the statistical analysis did not reach significance, X^2 (1, n = 126) = 2.5, p = .11. There was no significant

main effect of manager gender ($p > .3$), nor was there a significant interaction between manager gender and feedback type on gender assumption ($p > .4$).

Qualitative analysis

To assess participants' qualitative responses, text from the two outcome groups (male assumption and female assumption) was entered into the text analysis software Linguistic Inquiry and Word Count (LIWC). LIWC is a text analysis software program designed by James W. Pennebaker, Roger J. Booth, and Martha E. Francis (2001) and calculates the degree to which people use different categories of words including positive or negative emotions, self-references, causal words, and affect. Because the primary purpose of this study was to examine how qualitative statements differed according to the hypothesis that women are shown more concern or "sugar-coating" than men, I paid special attention to the affective differences in the text. While some positive emotion language was expected, if part of the explanation behind the assumption that the employee was female was that the manager was avoiding the generation of negative reactions from women, I especially expected that negative affect language would differ between conditions (though other differences are analyzed as well).

A slight difficulty arises from the way the text statements are set up as some of the statements had the same meaning, but different structures (e.g., "he was nice to her" has the same meaning as "he was not as mean to her" but different affective words). Thus, the first step was to analyze affective words overall, and then separate the statements into negation statements and non-negation statements and analyze the statements by gender assumption, within the white lie feedback category. Only the main word categories of interest will be

reported here, and the other words of secondary interest are reported in table 3.

Affect. Before looking at gender, I analyzed the use of *affect* words over all by feedback condition (truth/lie) and found that people used significantly more affective words in the white lies condition (M = 5.35, SD = 5.19) than in the truth condition (M = 3.07, SD = 4.36), t(124) = -2.648, p = .009. However, there was no interaction between target gender and feedback condition (p > .5).

Examining each of the emotion words alone (Positive, negative, anxiety, anger, and sadness) revealed that there was a main effect of feedback condition on "negative emotions" and "sadness" such that more emotion words were expressed in the white lie conditions (see Table 3 for means and probabilities). The only marginally significant interaction between both target gender and feedback condition was in the emotion category of "sadness," F(1, 122) = 3.3, p = .07, where female employees were discussed with more sadness (M = .71, SD = 1.3) than were male employees (M = .17, SD = .94). There were no differences in emotion words for explanations of why employees were guessed to be male or female when the feedback had been truthful (see table 3 for means).

Second, I examined the correlations between *negations* (e.g., won't, don't, not, never, didn't) and negative and positive emotions in the white lie condition and found a positive correlation between *negation* words and *negative emotion* words, r = .29, n = 68, p < .01 overall. That is, *negative emotion* words were most often found next to negation words in sentences describing why a manger lied to the employee (e.g., "The manager *didn't* want her to be *sad*"). Analyzing these same word categories in a regression by employee gender, I found that this was especially the case for female employees overall, F (6, 33) = 3.2, p <.01, and especially for *affect*, *negative emotion*, and *anger*, which all significantly correlated with

negation (p's < .001) and *sadness*, which was approaching significance.

Table 3

Emotion Words by Gender Assumption and Feedback Category

Word Category		Assumed Employee Gender (White Lies)		Assumed Employee Gender (Truth)	
		Male	Male	Female	Female
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Emotion Words	*Affect*	4.55 (4.3)	5.92 (5.7)	2.72 (4.7)	3.45 (3.8)
	Sadness	0.17 (.94)	0.71 (1.3)*	0.04 (.26)	0.00 (0.0)
	Anger	0.56 (1.4)	1.50 (5.0)	0.17 (0.7)	0.58 (2.1)
	Anxiety	0.37 (1.4)	0.45 (1.1)	0.00 (0.0)	0.32 (1.1)
	Neg Emotions	1.45 (2.1)	2.80 (5.0)†	0.83 (3.0)	0.99 (2.3)
	Pos Emotions	3.04 (3.7)	2.99 (2.8)	1.89 (3.7)	2.38 (3.8)

Note. *** p < .001**p <.01, * p < .05 †p < .10. This table represents the mean number of emotion and other words represented in rationales for why participants had assumed the employee to be either male or female. The left section represents the emotions words in the white lies feedback condition and the right column for the truthful feedback condition, for male and female assumptions. Only the difference between sadness is significant between males and females in the white lie condition.

Discussion

The results only partly confirmed the primary hypotheses of Study 2b. Though the effect from Study 2a was not *significantly* replicated, the prediction that those who had assumed the employee was female after inaccurate feedback would be more likely to use affective language in their qualitative reasoning (because women are assumed to be more emotionally reactive to negative information) was supported. Overall, more *affective* words

were used when the manager has told a white lie, suggesting that people's perceived reasons for doing so are emotionally laden. Additionally, in their qualitative justifications, participants used more *negative emotion* words - especially coupled with *negations* - to describe their reasoning. That is, the responses involving female employee assumptions were most often characterized by words such as 'didn't' and 'sad' suggesting that this group was concerned with avoiding negative emotions.[9]

This may provide a clue as to why women are lied to more often in the first place: if people assume that other people try to spare female's feelings more often, then it is possible that this is really the case. However, because the construction of the sentences were so different – that is, some people indicated both why they had chosen a female, *and* why they had not chosen a male – a more complete text analysis would be beneficial to confirm these predictions and results.

The secondary hypothesis could not be confirmed by the data. That is, the gender of the manager did not seem to affect the determination of employee gender based on the different types of feedback. Before confirming a null result, however, it is worth considering some of the methods used to find the result. Because only two voices were used in this study, one female and one male, there are numerous other factors that could have (a) created much more noise in the data than anticipated and (b) created confusion as to which cues to use when determining employee gender. That is, voice intonation, tone, and pitch, perceived intentions based on character assumptions, and other factors that one can pick up from voice

[9] An example justification sentence for why someone had assumed the employee was female in the 'white lies' condition: "*Though somewhat stereotypical, it sounded like the speaker was speaking in a more "soft", passive tone making me believe she was trying to soften the blow to a female employee*."

And an example of a male employee assumption for the truth condition: "*It seemed to be more of a direct, to-the-point statement, that perhaps maybe the manager might perceive a male being able to handle a bit better than a female.*"

could all have affected the data. Participants may also have thought that the instructions were asking them to listen to the vocal cues themselves, not the content of the feedback, thus answers may have differed depending on where participants were focusing their attention and from which cues they were deducing their answers.

Finally, several participants had very complex and psychologically sophisticated qualitative justifications based on the gender dyads, which may have also reduced the effect of the main finding from Study 2a. For example, there were several participants who reasoned that there must have been romantic interest between the manager of one gender and the employee of another or even antagonistic or competitive interests between same-sex dyads. Thus, the content of the feedback alone was not the only influencing factor. This may also explain why the previously robust effect of gender assumption based on feedback of Study 2a was diluted in Study 2b (though in the predicted direction), however, based on this data, it is difficult to rule out the alternate hypothesis that presumed romantic or other motivations were driving this effect.

All in all this study adds some insight (involving some new variables) into why women are told white lies more often than men during performance feedback, and the results suggest that people who tell white lies may be assumed to be avoiding the negative emotions or reactions of others, for reasons yet to be determined. Future studies should examine more directly the motivations of those who tell white lies to more clearly understand and identify the underlying mechanism. In the next set of studies, more direct methods for answering the mechanistic question will be examined.

CHAPTER 4

Why Do People Tell White Lies to Women During Performance Feedback?

Study 1a & b and Study 2 a & b suggest that (a) people have a hard time telling the truth about performance directly to others; (b) that white lies are particularly told to female targets; (c) that though people do not admit or are not aware of this bias in themselves, they are aware that his bias exists in the world; and (d) that people use more emotional language in their reasoning for why women are told more white lies. This next section will shed more direct light on the reasons why people tell white lies to women, and who is most likely to show this bias during performance feedback.

Study 3

Do Stereotypical Beliefs about Women and Reactions to Negative Information Contribute to White Lies about Performance?

In Study 1a, people told more white lies to female targets who were underperforming than to their male counterparts, and in Study 2a, people believed that employees who had been lied to were more likely to be female, indicating that not only do people exhibit this bias themselves, they recognize it in others as well. Though the effect appears to be established, it remains unclear exactly why people tell more

white lies to women. The results from Study 2b suggest that people were accessing stereotypes about women's emotionality by using more affective language in their qualitative reasoning. However, it has not yet been directly tested whether stereotypes about how women respond to workplace feedback (and sexism more generally) contribute to the feedback bias effect.

The hypothesis that spurred these studies centered around the idea that women are perceived differently from men when it comes to receiving feedback; more specifically, that women are stereotypically perceived as more emotional, less able to take criticism, prefer biased feedback, and evoke empathic and compassionate responses (all specific correlates of the warmth and competence components of the theory of benevolent sexism). In Study 2b, there was some support for this hypothesis with the finding that people believe women are lied to more often about performance because people wish to soften the emotional blow that criticism will have. It is possible that any one of or all of these specific beliefs are driving this gender bias during feedback; however, a direct test of this hypothesis needs to be conducted.[10]

In Study 3 I tested whether stereotypical views about how well women receive feedback could be contributing to this effect. In this study I measured individual attitudes toward women and criticism and the effect of these attitudes on inaccurate feedback to women. Similar to the previous studies on MTurk, because I used a hypothetical feedback scenario, and not an anxiety-inducing direct or face-to-face paradigm, I did not expect to see a strong main effect of lying, especially considering

[10] It is also possible that there is an entirely different motivation behind the bias, including that people believe they are giving women a step up in an otherwise sexist world. This "affirmative action" motivation will be further discussed in the General Discussion section.

that most people believe that dishonesty is wrong (Kashy et al., 1996). Still, I predicted that people who held more stereotypical attitudes about women's responses to feedback would be more likely to tell white lies to women than to men.

Method

Participants.

One hundred and fifty one participants were recruited through Amazon's Mechanical Turk (MTurk) crowdsourcing website and compensated $0.05 per minute. The sample was fairly representative of the U.S. population demographics, with 82 (54%) female and 69 (46%) male respondents, 64% of which were between the ages of 20 and 40, and 26% between the ages of 40 and 60.

Procedure.

All participants were administered an eight minute survey on the Qualtrics survey platform. The hypothetical scenario setup was nearly identical to the workplace scenario used in Study 2a; however, rather than evaluating another person's feedback to an unnamed target, participants were asked to imagine themselves as a manager giving feedback to a target that was varied by gender. As in Study 2a, they were also told that their job as manager was to give constructive feedback to the employee for the good of the company, and that the manager generally got along with all of the employees in the workplace (talking with them in the hallways, seeing them at lunch etc.).

Gender manipulation. Again, exactly like Study 2a, the participants then read a hypothetical scenario about an underperforming employee who had been on the job

for about six months but whose performance was worse than expected. The gender of the employee was manipulated using a subtle name change and varied between subjects such that half the participants read about Mary as the employee and the other half read about Mark.[11] Participants read that Mark[Mary] might lose his[her] job if he[she] kept underperforming and was doing objectively worse than the other employees at the business. Participants were then told that today was monthly feedback session, and that they would be asked to give feedback to Mark[Mary] about his[her] performance.

Feedback accuracy measure. Each participant was then told that he or she could select one type of feedback statement to give to the employee out of six possible feedback choices (the same statements used in Study 2a and Study 2b). All six choices were visible to the participants, and they ranged in levels of accuracy, beginning with (1) the objective truth (and closest to the original description of Mark[Mary's] performance) and ending with (6), a blatant lie.

Individual beliefs about women. After the participants had selected a feedback option to give to Mark/Mary, they were asked to answer true or false to three different statements about women. The first two were statements chosen from the Ambivalent Sexism Inventory (ASI, Glick & Fiske, 1996) because they directly implicated workplace sentiments, and the third was created by the author to directly assess beliefs about reactions to criticism. The sexism statements included: (1) "Women exaggerate the problems they have at work;" (2) "When women lose to men

[11] I chose two different names in a similar vein as those in Study 1a. These names were selected from the top 20 baby names of the 1970's so as to correspond as closely as possible in sound, age, and popularity, thus reducing the possibility that participants would draw extraneous conclusions from the names themselves.

in a fair competition, they typically complain about being discriminated against;" and (3) "On average women are as good at receiving critical feedback as men."

Of particular interest was the third measure as it directly assessed participant's beliefs about how well women receive criticism as compared to men. After the participants answered these True/False questions, demographic information such as age, gender, and political orientation were surveyed.

Results

Gender.

The modal feedback response was option 2, which 42% of participants chose, and option 3 which 40% of participants choose, thus the responses overall were slightly skewed toward the truth. An independent-samples t-test revealed that the effect of target gender on feedback choice was not significant: t(149), = .75, p = .45.

Sexism

I next evaluated the three sexism items both independently and as a composite score ($\alpha = .67$) . For the composite sexism score, each of the participants was given a score from 0 (none) to 4 (all) based on how many of the items they had agreed with. Then, participants who had not agreed with any of the questions (0) or only 1 of the questions were labeled "not sexist" and participants who had agreed with 2 or more of the statements were labeled as "sexist" for the purposes of analysis. This split categorization was done in order to preserve approximately equal participant numbers

in each of the categories for the subsequent analysis.[12]

In order to examine whether sexism moderates the effect of target gender on feedback accuracy, a 2 (Gender) X 2 (Sexism) factorial ANOVA was conducted. As predicted, participants who scored high on the ASI and had been coded as 'sexist' were more likely to tell white lies to Mary (M = 3.2, SD = 1.2) than to Mark (M = 2.6, SD = .88), $F(1, 144) = 3.1$, $p = .051$, $\eta_p^2 = .03$. However non- sexists were equally likely to tell the truth to Mary (M = 2.6, SD = .87) as to Mark (M = 2.6, SD =.76). Further examination of the single sexism items independently, revealed that agreement with the third item (that "women are worse at receiving feedback than men") especially interacted with target gender on feedback accuracy such that participants who agreed that women are worse at receiving feedback told more white lies to Mary (M = 3.2, SD = 1) than to Mark (M = 2.6, SD = .88), $F(1, 147) = 4.3$, $p < .05$, $\eta_p^2 = .03$ (Fig. 4). The other two individual items were non-significant (all p's > .2).

[12] When the data were coded as 0 = "non-sexist" and 1, 2, and 3 = "sexist," this categorization did not significantly change the results.

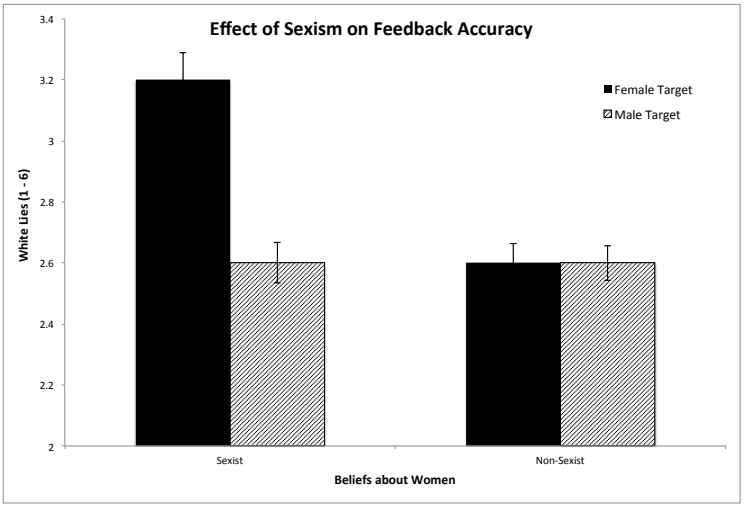


Figure 4. Chart represents the significant interaction between Target Gender and Participant Sexism on feedback accuracy (higher number indicate nicer but less accurate feedback). The results suggest that only people who indicated agreement with sexist statements about women in the workplace tell more white lies to women. The error bars represent the standard error from the mean.

Discussion

As I had predicted, and due to the hypothetical nature of paradigm, Study 3 did not significantly replicate the main effect of Study 1a, most likely because the element of social interaction that usually causes tension in person-to-person feedback was diluted due to the hypothetical nature of the experiment and the online setting.[13] However, as I had predicted, when taking into account individual degrees of sexism, the gender of the target did have an effect on the accuracy of the feedback. That is, only people who agreed with sexist statements about women in the workplace, especially that "women are worse than men at receiving feedback," tended to tell more white lies to the female employee than the male employee.

While Studies 1a, 2a, and 2b suggest that people lie more to women than to men, no matter their individual beliefs or individual differences, Study 3 found that

[13] Note that in later experiments the general effect of target gender on feedback accuracy is replicated.

only people who held sexist views about women in the workplace showed this bias. This suggests that although most people tell more white lies to women, sexists may be driving this effect. Study 3 also offers a possible explanation for why women are told more white lies and furthers the primary argument that stereotypical beliefs about women lead people to choose upwardly distorted performance feedback. These results, in combination with the findings in Study 2b that female employees are associated with more negative emotion words in the context of giving feedback, begin to paint a clearer picture of the mechanisms driving this gender bias.

CHAPTER 5

Problems and Consequences of Gender-Biased Performance Feedback

Studies 1 through 3 provide evidence that a gender bias in workplace communication exists, and is in part due to stereotypical beliefs about women. In particular, the studies suggest that the bias is associated with particular beliefs about women's excessive emotionality and their response to criticism. One might reasonably wonder if this feedback bias is in fact problematic - whether there are discernable negative consequences to the bias in terms of its impact on women. It is a fair question, for one could argue that if there are no consequences of this bias, and if it has no discernable impact on women, then I might be satisfied in knowing a discrepancy exists and go on with our lives. However, because this is a testable question, in the following two studies I considered the effects of this bias on two levels: psychological and material. First, in Study 4, I addressed the primary plausible problem: that a feedback bias is a problem for equality if women are treated differently but in reality expect and desire truthful feedback to the same extent as their male counterparts. There may be psychological and material consequences to expecting one type of feedback and being given another, including resentment, feelings of exclusion, or marginalization. In Study 5, I attempt to address the material and tangible consequences of inaccurate feedback on the productivity and performance of people who have been told white lies.

Study 4

Do Women Desire Biased Performance Feedback?

In the previous studies, participants told more white lies to females than to males. If biased feedback was indeed desired by women over truthful feedback, then one could conclude that perhaps biased feedback is not problematic. However, if women do not want to be given biased feedback, then to do so is to creating an unequal situation for women. This discrepancy between what is desired and what actually happens could have several harmful effects, including demotivation and loss of productivity, or at the very least psychological consequences such as patronizing experiences in the workplace that could result in anger and resentment (see Vescio & Gervais, 2005). Additionally, benevolent sexism has been shown to have consequences for women's cognitive performance, due to mental intrusions women experience about their sense of competence (Dardenne, Dumont, & Bollier, 2007). On the other hand, if women do expect and prefer inaccurate but nicer feedback then perhaps evaluators are merely fulfilling the wishes of women. Thus in Study 4 I tested the possibility that women do not in fact show a preference for nicer but less accurate feedback.

In order to test whether there is a discrepancy between what people think women want (or at least what their behavior suggests they do), and what kind of feedback women really want, I ran an online survey using the same measures as Study 2a and 2b, explicitly asking both women and men what kind of feedback they would *most* and *least* prefer if they were employees unwittingly performing below standards. I predicted that men's and women's responses would not be different and that both

genders would prefer feedback that was more accurate than nice. Though socially desirable responding could be a problem with this design, because people tend to believe that lying is wrong, if people did not choose the most accurate type of feedback as their most preferred option it should allay worries that participants were just choosing according to social norms. I also predicted that people's least preferred feedback would be the most patronizing kind, and that if women were in any way especially attuned to this bias in their own lives that they would be particularly likely to select a patronizing lie as their *least* preferred feedback.

Finally, I also asked participants to recall times that they had reacted badly to criticism. Previous research on emotional displays of men and women has found that women often display and are stereotypically perceived to feel more passive emotions such as sadness, whereas men display high-arousal emotions such as anger (Plant et al., 2000). Thus, I also predicted that although women would be more comfortable talking about their emotional reactions to criticism, men would recall more aggression-related (high-arousal) reactions to critical feedback and women would recall more passive (low-arousal) reactions to negative feedback. This would indicate that women and men react to criticism in different ways, negating the alternative argument that people tell more white lies to women because only women are reactive, and suggesting that if avoiding reactions is a motivating factor in telling white lies, then people should also tell white lies to men to avoid aggressive reactions, which, as previous studies have shown, they do not.

Method

Participants.

101 participants were recruited through Amazon's Mechanical Turk (MTurk) crowdsourcing website and compensated $0.05 per minute. 46 (45.5%) of the participants were female and 55 (54.5%) were male respondents, 72% of which were between the ages of 20 and 40, and 28% between the ages of 40 and 60.

Procedure.

All participants were asked to provide answers to a hypothetical workplace scenario nearly identical to that of Study 2a and Study 2b. In this scenario, however, they were asked to imagine themselves as an underperforming employee and told that that though their manager had determined they had been doing poorly, they themselves were unaware.

Feedback preference measures. In order to assess what kind of feedback each participant would be the most constructive, and what kind of feedback they would be the least constructive (the idea of "constructive" determined by the participants), all participants saw the following text followed by the six feedback statements used in studies 2 through 3:

> Most Preferred:
> "*Given that option 1 is the most truthful, and option 6 the least truthful, of the six feedback choices below that your boss could give you, choose the one that you think would be the most constructive for you. That is, which one would you want to hear so that you could improve your performance and do better at your job*?"
>
> Least Preferred:
> "*Now please pick the option that would be the LEAST constructive."*

Emotions measure. After completing the primary measures, participants were also asked to rate, on a bipolar scale of 1 (*not at all*) to 100 (*extremely*) how upset they

would have been had their boss chosen to give them the most truthful statement and the least truthful statement. Participants were also asked to select from a list of specific emotions (anger, hurt feelings, shame, annoyance, and sadness) which best characterized the 'upsetness' they had previously indicated.

Reactions to criticism. In order to assess whether there is a possibility that women really do react worse to criticism, I additionally asked all participants to answer yes or no to the following question:

> "*As an adult, have you ever gotten performance feedback from a superordinate that you have reacted to badly (showing anger, quitting, crying, fighting, flipping out etc...)*?"

I additionally asked participants to describe in a few words how they had reacted badly.

Results

Main Hypothesis Analysis

Most preferred feedback. I predicted that female participants would not respond differently than male participants in regard to what kind of feedback they believed would be the most constructive for them. Using an independent-samples t-test, I found that when selecting their most preferred option, female participants (M = 2.7, SD = .76) did not significantly differ from male participants (M= 2.71, SD = 1.1), $t(99) = .07$, $p = .95$. Though the mean response was 2.7 out of 6, the modal responses were overwhelmingly option 2 (37%) and option 3 (42%), indicating that on the scale of most truthful but harsh to least truthful but nice, all participants appeared to prefer feedback on the more truthful side, though not the harshest feedback (6%).

Least preferred feedback. When I conducted an analysis to assess gender differences in least constructive feedback again using an independent-samples t-test, I found that although there was not a significant difference between male and female preferences, t(99) = -1.8, p =.08, the data appeared to suggest a possible difference between genders. Females indicated that their least preferred feedback was nicer yet less accurate (M = 4.65, SD = 2.1) while males indicated that they also least preferred nicer feedback (M = 3.85, SD = 2.36) but to a lesser extent.

Because the variances were high, I examined the modal responses for these questions and found that the distribution of responses for male participants was bimodal, with about equal numbers indicating that their least preferred feedback was option 1 (most truthful: 36.5%) and option 6 (least truthful: 51%). Female participants, on the other hand, were more unimodal, with 71% indicating that their least preferred feedback was option 6 (least truthful), and only 23.9% indicating option 1 would be their least preferred. Thus the data suggest that female participants particularly feel that white lies were what they were least interested in (or, at least, that they were more likely to express such a preference).

Emotional Reactions. A multivariate ANOVA was conducted to determine if there were any differences in male and female self-reported emotional reactions to each of the six feedback statements. The multivariate test was not significant, $F(6, 94) = 1.57$, $p = .16$, nor were any of the Univariate tests (all p's > .1). However, because the probability was approaching significance, the means for each of the statements were compared, and suggested that female participants reported that they would have expressed higher 'upset' emotional reactions, in particular had they been given

feedback statement 1 (M = 81.1, SD = 26.8) than males (M = 72.9, SD = 26.1), though the variance was quite large.

When asked to choose which specific emotion qualified their 'upsetness' most, participants chose 'shame' (52.5%) and 'hurt feelings' (34.7%), followed by 'anger' (8.9%), 'annoyance' (3%), and lastly, 'sadness' (1%). A Chi Square test indicated that there was not a significant difference between male and female responses overall, $X^2(4, N=101)$, p = .168. However, because the cell counts were uneven, I coded and collapsed each of the emotions into two categories: 'high arousal' (anger and annoyance) and 'low arousal' (sadness, hurt feelings, and shame). When the Chi Square test was performed again on these two categories, although the majority of participants had chosen a low arousal emotion (88%), there was a significant gender difference $X^2(1, N=101)$, p = .032, such that 83% of those who identified a high arousal emotion were male (see Table 4 below for all original frequencies).

Finally, I analyzed female and male participants' yes/no answers to the question "have you ever reacted badly to criticism" using a Chi Square test of independence. Women reported reacting badly significantly more often than men, $X^2(1, N=101)$, p = .034, 43.5% percent as compared to 23.6%, though the majority of people overall reported not reacting badly. In order to calculate potential differences in the specific way that participants had reacted, qualitative answers were coded as either aggressive (screaming, talking back, flipping out, quitting) or passive (crying, hiding, depression). While there were only 27 qualitative responses, there seemed to be a significant difference in the way that males and females reacted, $X^2(1, N = 27)$, p = .05 such that 80% of the male respondents reported acting aggressively, while only 41%

of the females reported doing so. 59% of females reported acting passively while 20% of males reported doing so.

Table 4

Reported emotional reaction type to critical feedback by gender

Emotions	Male Participant Count (% within option)	Female Participant Count (% within option)
High Arousal	10 (83.3%)	2 (16.7%)
Low Arousal	45 (50.6%)	44 (49.5%)
Anger	7 (77.8%)	2 (22.2%)
Annoyance	3 (100%)	0 (0.0%)
Hurt Feelings	16 (45.7%)	19 (54.3%)
Shame	28 (52.8%)	25 (47.2%)
Sadness	1 (100%)	0 (0.0%)

Discussion

As I predicted, women did not desire different feedback then men, and all participants desired feedback that was more truthful than nice (though not the most truthful). However, because I were predicting a null hypothesis, our first aim was to determine the power of this particular study to make sure that any null results would be due to an actual lack of difference and not because I had not surveyed enough participants. In a post-hoc power analysis, by entering the expected medium effect size ($f = .25$), at the $p < .05$ level, power was indicated to be $(1- \beta) = .70$, with a critical F of 3.93. In order to get a desired power of $(1- \beta) = .80$, the sample size would only need to increase to $n = 128$, thus I could argue that that the null findings were not due to an insufficient sample size, but that, as predicted, female participants truly did not desire different feedback than their male counterparts.

Additionally, the data from the *least* desired type of feedback implies that

when it comes to feedback that people feel is not constructive, women are especially likely to feel that white lies would be the least constructive feedback, compared to men who reported that the harshest truth would be the least constructive feedback. Though there is no data yet to back up such a claim, one could argue that women may be particularly sensitive to patronizing environments due of their own experiences with such settings. Vescio and Gervais (2005) found that women who had been patronized (verbally praised but allocated less money) responded to such patronizing environments by becoming angry and less motivated to contribute. It remains to be seen whether women believe that they are more likely to be told white lies and whether this has an effect on their feedback preferences. However, the data in Study 4 suggest, at the very least, that women do not want to be told white lies any more than men do.

Finally, though there were no significant differences between how upset female and male participants reported they would get had they been given the most accurate and harshest feedback, the patterns were strong enough to discuss here: women appeared to be reporting that they would be slightly more upset than the men asked the same question. It is important to note, however, that culturally men are less likely to report emotional intensity – a component of the stereotype that women are more emotionally volatile is that they are also more likely to freely express their emotions, while men have traditionally confronted a norm to suppress their feelings (Gross & John, 2003; Gross & Levenson, 1993). Thus, these data may not be capturing anything more than cultural norms. And, when participants were asked to choose specific emotions that characterized their 'upsetness,' males chose high arousal

emotions like anger more often than women, which is consistent with the stereotypes about male anger and female sadness (Plant et al., 2000). This raises another question: If males do react to negative feedback with anger or aggression, why do people tell white lies more often to women? It seems that if an aggressive response is expected, people should take actions to avoid such a response, including upwardly distorting feedback.

I also hoped to address one final question within the scope of this project: Do negative experiences associated with giving negative feedback to women lead people to tell more white lies in future interactions? In this study, I attempted to examine one part of this question by asking men and women whether they had ever reacted badly to negative criticism. Women did report reacting badly to negative feedback significantly more than men did; however, their reactions were more characterized by passive actions (crying, hiding in the bathroom, feeling depressed, etc.). Men who had reacted badly reported that their reactions were more active or aggressive (yelling, quitting their job, talking back to their boss, etc.). Overall, however, most people reported more passive emotional reactions so this distinction is not entirely meaningful.

Thus, though it is possible that women really do react worse to feedback than men, this question was subject to similar self-report issues as previous measures. It may be that women feel more comfortable saying they had acted badly. If it is the case however, it appears that 'acting badly' for women seems to be more passive, which again brings up the question of why managers don't tell white lies to men as well, especially if their bad reactions are likely to be aggressive.

Finally, another possible limitation in this study is that by asking participants

to indicate the most and least constructive feedback only one component of the feedback process is being examined. Perhaps if participants had been asked to indicate which type of feedback would be most constructive for their self-esteem, some gender differences would have been seen, in particular women preferring softened feedback. However, because the term "constructive" was left mostly up to the participant's own understanding, it is difficult to say that the moral or self-esteem aspect was not taken into account when considering future performance. That is, participants may well have decided that accuracy is more important than self-esteem. Future studies could examine this psychological process more carefully to see how participants are determining "constructiveness."

Ultimately, Study 4 highlights one of the problems inherent in the communication gender bias: that telling white lies to women, regardless of intentions, stereotypes, or even actual experiences with negative reactions, does not match the desires of women, who, according to the data appear to want the same feedback that men do, at least in self-report. Or as one participant put it in the comments section: "Can't sugar coat some things in life. At least knowing the truth gives you a chance to work harder, or improve. But, some situations can improve, usually with direct and constructive feedback, IMO."

Study 5

Performance Consequences of Receiving Biased Feedback

Study 4 identified at least one problem with biased performance feedback: the mismatch between people's gender biased behavior and the actual needs and desires of the group that is most affected by this bias. However, there may also be material consequences of biased feedback that affect both the individuals involved and ultimately the organization to which they belong. Thus, in Study 5, I sought to investigate whether there are any tangible consequences of receiving inaccurate but nicer feedback on the recipient's performance.

The effect of different types of feedback is a widely researched topic in business and management studies, as feedback is often a key component in the effective communication between leaders and subordinates. Though feedback can have genuine and tangible effects on productivity and morale, there are many factors that determine whether feedback has a positive or negative impact. Researchers who investigate the effectiveness of feedback have not settled on the best kind of feedback in this respect. In fact, the results of many different studies contradict or qualify each other, and different settings, situations, and personalities all seem to modify the findings (For a review, see Arvey & Murphy, 1998; Kluger & DeNisi, 1996; Levy & Williams, 2004).

A meta-analysis, for instance, found that the effects of feedback are variable and that feedback impacted people negatively in about a third of the cases (Kluger &

Denisi, 1996). Though the data are mostly inconclusive, in particular there appear to be a few key variables that change the nature and effect of feedback. One is how important the task or activity being evaluated is to the person. Another is how trustworthy the feedback is (authority, respect, etc.) (Steelman & Rutkowski, 2003). Beyond those variables, researchers have found multiple moderators for the majority of findings. For example, one study found that focusing on criterial feedback (e.g., task completion) can improve task-performance, while social comparison feedback inhibits performance (Wilbert, Grosche, & Gerdes, 2010).

In terms of negative and positive feedback, the majority of the literature does establish that individuals who receive lower performance evaluations tend to improve more than those who receive higher ratings (Atwater et al., 1995; Reilly et al., 1996). For example, constructive critical feedback was found to be a better motivator than positive-only feedback, but only if said criticisms were deemed useful, which is less likely with negative feedback (Brett & Atwater, 2011). However, this is not without qualifications. For example, Brockner & Elkind (1985; 1987) found that negative feedback has negative consequences for those with chronic low self-esteem. And Cohen, Steele, and Ross (1999) found that black students reported lower ratings of motivation when receiving negative feedback but were motivated when the negative was buffered by compliments.

Though feedback type (negative or positive) has been systematically researched, the *accuracy* of information seems to have been less thoroughly examined. There have been some suggestions that accuracy of evaluations can increase performance on subsequent tasks and in turn increase job-satisfaction (Brass, 1981),

while overly positive feedback can lead to worse performance (Ellis, Mendel, & Aloni-Zohar, 2009). However, people tend to deem negative feedback inaccurate compared to positive feedback (Kluger & Denisi, 1996). Therefore, it is possible that inaccurate positive feedback may receive more attention than its negative counterpart. There is significantly less research on the connection between feedback and gender but there is evidence that men and women do interpret and receive feedback differently. Biernat and Danaher (2011) found that women interpret subjective qualitative feedback more negatively than do men, and they argue that this may lead to demotivation after critical feedback. Additionally, research suggests that women are more responsive to and more emotionally attuned to both negative and positive feedback (Roberst, 1991). Additionally, Baron (1988) found that women tend to perform worse after 'destructive' negative criticism.

Others have qualified these findings by demonstrating that ambivalence is what causes feedback to be destructive, not necessarily negative or accurate feedback; if given clear criticism, studies suggest that women interpret it no differently than men (Dweck & Gilliard, 1975). Additionally, Roberts & Nolen-Hoeksema (1989; 1994) found that women do interpret information as more self-relevant than do men, but that this does not necessarily reduce their self-esteem any more than men. They further found that women are more accurate in their self-evaluations than men and see other's evaluations as more accurate and valuable.

Taking this prior research into account, and using the finding that critical feedback tends to improve performance (even though it may be more emotionally impactful), I designed a study to test the prediction that more accurate (albeit negative)

feedback would improve performance on several quantifiable tasks. Additionally, though the research is equivocal on the role of gender in receiving and deploying feedback, I hypothesize that although women may react more negatively to critical feedback than men, that there would be no difference in their improvement on tasks. If anything, I anticipated that they would improve more because of their heightened internalization of critical information.

In this study I attempted to simulate a real feedback situation and evaluate the change in performance after receiving negative accurate feedback or positive inaccurate feedback. I attempted to ensure engagement with the task by making the topic as personally important or "ego-relevant" as possible (see Butler, 1987). Most people would like to consider themselves to be intelligent, so I used 'creativity' as a (fake) correlate of intelligence as the factor I would be measuring—a strategic choice reinforced by research suggesting that creativity is malleable and responsive to feedback and other environmental cues (Fromme, Mercadal, & Mercadal, 1976). Additionally, because I were attempting to simulate real feedback scenarios online, I planned the tasks to be difficult and obscure enough that the participant would feel the task was difficult and in turn would not be entirely sure how he or she was doing. Finally, because performance can be measured across several criteria, I chose two dimensions on which to concentrate: motivation (effort), and quality (change in performance).

Our primary hypothesis was that participants who had received nicer but less accurate feedback would (a) not be motivated to work harder to perform in the final round and (b) would not improve their performance during round 2 as much as

participants who had received negative feedback. I also predicted the women would indicate more emotional sensitivity to both negative and positive feedback, but that their levels of motivation and performance would not necessarily suffer and that overall they might be more motivated to improve their work than men after negative feedback.

Method

Participants.

Forty-six participants were recruited from Mechanical Turk, but six participants did not complete the survey, leaving 40 participants for analysis. 24 participants (60%) identified as male, while 16 (40%) identified as female. 77% of participants were between the ages of 18 and 40, and 23% were between the ages of 40 and 60.

Procedure.

All participants completed the same 20-minute survey on the Qualtrics survey platform. Before beginning the experiment, participants read that they would be completing several tasks that would test their creativity, a mental trait highly correlated with overall intelligence. This plausible-sounding yet false link was used in order to instill personal engagement with the task as I suspected that most people would want to see themselves as intelligent and thus perform as best they could:[14]

> *"While there are many different factors contributing to intelligence, creativity tends to be highly correlated with intelligence. People who*

[14] This association has not been established by research and the link between creativity and intelligence is actually unclear. Participants were debriefed and this was clarified at the end of the study.

are more creative tend to be good problem solvers, have high IQ's, and are flexible thinkers. Creativity does not necessarily mean that you are good at art or music, but that you are able to make associations between ideas and think out of the box. In the next few questions, we will ***test your creativity*** *by asking you to complete a few tasks.* ***You will be judged on how well you did*** *on the tasks and then given feedback about your performance."*

In a mixed factorial 2 (Time, within subjects) X 2 (Feedback Type, between subjects) design, participants did the same three tasks twice, once before receiving feedback, in order to get their Time 1 performance scores, and again after receiving feedback, to measure the change in their performance at Time 2.

At Time 1 participants were first asked to indicate their mood, after which they completed three different tasks designed to measure their 'creativity'. It should be noted that the tasks were designed to be quite difficult and nearly impossible to complete successfully. I achieved this by manipulating the difficulty (both real and perceived) and also by setting time limits for each task. This was done in order to make the negative feedback the most realistic type of feedback, and similarly so that the positive feedback would be the least accurate.

After Time 1, half of the sample was given negative feedback and the other half was given positive feedback (regardless of their actual performance). The participants' motivation to improve was then assessed with two different measures, a binary choice measure (yes / no) to determine their intent to complete extra tasks to possibly improve their score, and a continuous (0 – 10), subjective measure of their perceived degree of motivation. At Time 2, participants were asked to complete nearly the same three tasks they had done at Time 1 (differing only in content), and once finished were asked to indicate their mood at and demographic information such as

their gender.

Creativity measures. Three quantifiable tasks were administered to all participants in order to evaluate the change in performance from Time 1 to Time 2. Though not all the tasks were typical 'creativity' tasks, I used them because they were relatively easier to score in an objective way to get a realistic difference in performance. They were also difficult enough that participants would feel challenged but not necessarily know how they were performing. All three of these tasks were administered at Time 1 and at Time 2. Though the task was the same, the content varied slightly at the different times so that participants had to complete the same task but would not have learned the answers from the prior task. Because the content was different, the two versions of each task were counterbalanced so that half the participants saw version A at Time 1 and the other half saw version B at Time 1.

Alternate Uses Task. The first task the participants saw was Guilford's Alternate Uses task (1967) where participants are provided with one word, usually a common household item, and asked to come up with as many possible alternative uses for this object as they can. The answers are then evaluated on originality, flexibility, fluency, and elaboration. All participants were asked to read the text below and were then shown one randomly chosen object word (from a list of ten possible object words):

> *"In this **"ALTERNATE USES"** task, you will be given a word prompt and asked to **come up with as many uses** for that word as possible in 60 seconds. The page will automatically submit your answers and proceed to the next page after the 60 seconds are up. **Your goal is to come up with as many plausible answers as possible in the amount of time**.*

For example: If you are given the word "sheet," some alternate uses that you might come up with would be things like "costume" or "blanket" or "partition" or even "art canvas." Once you feel like you understand the task, please proceed to the next page when you are ready. Your word will appear immediately."

Participants were then given one minute to list as many alternate uses as possible in the text boxes provided.

***Remote Associations Test (RAT)*:** The second task that participants completed was a modified version of the Remote Associations Test (RAT: Mednick & Mednick, 1962; 1967). This test was developed to measure convergent thinking, a marker of individual creativity. In the original (1962) task participants were asked to offer a word or concept conjoining a cluster of three words, over a series of forty discrete questions. For example, the words "falling" / "actor" / "dust", would be conceptually united by the word "star" (falling star, an actor is a star, and stardust). For this study, I chose thirty RAT items (15 for each version) from the original set of RATs from Mednick & Mednick's (1962) list and some from Bowers and colleagues' (1990) list. Items were chosen based on difficulty level (moderate to difficult) as coded in a study by Bowden & Jung-Beeman (2003). Each participant was asked to read the following instructions:

"*Now you will do a REMOTE ASSOCIATION TASK. In this task you will see three different words grouped together.* ***Your task is to figure out the word that all three of these words have in common.*** *So for example, if you see the words FALLING / ACTOR / DUST, then the word and concept uniting these three would be STAR (falling star, an actor is a star, and stardust). The concept they have in common is STAR. Once you figure out the common word, type it into the box. The page will automatically proceed after the 90 seconds are up. Your task is to do as many Remote Associations in the 90 seconds allotted to this task."*

Fifteen of the remote association word groups were then presented to the participant and they were asked to fill in their answers in the boxes provided in the given time.

Cognitive Reflection Task (CRT). The final task was a critical thinking task, commonly used to assess intelligence and out of the box, system II thinking (Frederick, 2005). In this task, mathematical problems that appear simple generate an easy (and incorrect) answer that most respondents intuit rapidly. However, upon greater reflection, the correct answer can be computed. For this task I used two such CRT's, randomized to be completed at Time 1 or Time 2, purportedly to measure creativity. For both CRT's participants were asked to choose from a multiple choice list of pre-generated answers, among which the incorrect (easy) and correct (difficult) options were presented (see Szrek & Bundorf, 2011 for this methodology). Participants were shown the following CRT items:

CRT 1:
A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?
Possible answers: 3 cents, 5 cents, 10 cents, 15 cents, 20 cents

CRT 2:
If it takes 5 machines 5 minutes to make 5 widgets, how long would it take100 machines to make 100 widgets?
Possible answers: 5 minutes, 20 minutes, 50 minutes, 100 minutes, 200 minutes

Expectations measure. In order to assess both how well participants thought they did on Round 1 and 2 and if the tasks I had used were appropriately difficult, I asked participants to indicate "How well you think you did OVERALL, regardless of

whether you felt you did better on some tasks over others" on a 1 (worst) to 9 (best) point scale. Participants were also asked "What percentile would you say you fall in compared to others (higher percentile means you did better, so if you think you are in the 90th percentile, that would mean that you think you probably did better than 90% of the rest of the people taking this task)" and asked to indicate this percentile on a 0 to 100 percent scale. These questions were asked twice, once after Round 1 and once after Round 2.

Feedback Accuracy Manipulation. The feedback accuracy manipulation was included to induce either a false feeling of positivity or a more accurate feeling of negativity in the participants. Although I could not generate completely accurate feedback for each individual participant, the natural difficulty of the tasks made it more likely that the negative feedback was more accurate than the positive feedback (analysis of the actual performance data will reveal if this was in fact the case).

After completing Round 1 of the creativity tasks, participants were randomly assigned to either receive a 'positive feedback' or 'negative feedback' score. Before receiving their scores, all participants saw the following text: "*Your performance across the items will now be reviewed based on several criteria, including how you did relative to how most people do on average.*" Participants were then shown the following scoring criteria by which their score had been purportedly calculated:

> *"Based on your performance on the previous creativity tasks, we are evaluating you on a scale of 0 to 100, with 0 being the least creative, and 100 being the most.*
> *Here is how to interpret your score when you get it:*
> ***0 - 50** = You performed below average.*
> ***50 - 100** = You performed above average."*

Participants in the negative feedback condition then received a score of 48.1, along with a message that read: "*You performed: BELOW AVERAGE*." Likewise, participants in the positive feedback condition received a score of 68.1, along with a message that read: "*You performed: ABOVE AVERAGE*." These scores were chosen based on previous pilot data to accurately represent the feeling of below average and above average and the explicit messages were added to ensure that participants would not interpret the scores differently. [15]

Motivation measures (effort). This measure was intended to evaluate whether participants would be motivated enough to add more time to their survey because they desired to do better, even if there was a cost (time) to it. After participants had received their (negative or positive) feedback, they were told that they would now be given an opportunity to improve and shown the following text:

> *"Now that you have gotten your feedback and know how you did on the preliminary round,* ***you will have the opportunity to take the tests again to see if you improve, which is entirely possible****. Even though you may have gotten a different score than you predicted, preliminary tests don't always reflect your true potential. Because most people need a practice test in order to perform to their true abilities,* ***this NEXT set of tests will be the ONLY ones used for your final evaluation and recorded.***
> *You can think of the first set of tests (that you just completed) as a practice trial and this upcoming set as the actual performance measure. The instructions for each of the tasks are the same, however the content (e.g. the words, remote associations) will be*

[15] In preliminary testing, results indicated that participants who received (a) whole number scores and (b) scores that deviated too far from the mean (e.g. 20 and 80) found them to be unrealistic. Additionally, it became clear from testing that people who find their scores very close to the average even if only one or two points away (e.g. 48), still feel that they performing below average. That is, people who feel average feel nearly as bad about being average as they do about being well below average. Thus the numbers calculated and used in this experiment are meant to capture the psychological experience of performing better or worse than average, and are not numerically equivalent in terms of distance from the mean.

different."

In order to measure participant's subsequent motivation, I offered them an extra round of tasks, in order to ostensibly improve their score. In reality, there was no extra round of tasks, but participants were told they could double their time and round the scores from the purported two subsequent rounds. In essence, I was assessing whether participants would accept a time cost in order to improve themselves, as a way to measure their motivation.[16] All participants read the following text, and indicated whether they would like to accept this cost (with no extra pay):

> *"Before you begin the next round of tests,* ***you have the option of completing the round of tests twice****. Scientific research has show that* ***people can and do improve on cognitive creativity tests if the tests are practiced or taken more than once*** *(Bors & Vignau, 2003).[17] So, although this additional round of tests* ***will add approximately 2-3 minutes to the amount of time*** *it will take you to complete the tasks, you may improve your overall performance and get a higher score. In essence, there will be a Round 2 and then a Round 3 (instead of just a round 2). There is no extra payment for the time you choose to add to the experiment. Before starting the next round of tasks, you may opt for the opportunity to do an extra set of tests instead of just one, and* ***we will average the scores of the two sets for your final score****. If you choose not to do the optional round, you will only complete one round (round 2) of tests."*

Participants then indicated whether or not they would like to accept the extra time and round (yes/no) and then continued onto the next motivation measure, which asked them to indicate, on a scale of 1 (*not at all*) to 10 (*a lot*), how motivated they were to do well on Round 2 of the tests. After indicating their motivation, participants then completed the second battery of creativity tests (see measures). None of the

[16] Note that MTurk workers are paid very little for their time, and subsequently, value their time very highly. An extra 2-3 unpaid minutes is a large time sacrifice for workers on MTurk, and thus a good evaluation of their motivation and intrinsic commitment to improving their performance.
[17] This is a real citation.

participants completed a third round of tests (though the ones who had selected 'yes' believed they would).

Mood measure. Negative and positive feedback has been shown to induce different moods and also specific emotions and also motivate one to pursue goals (Fishback, Eyal, & Finkelstein, 2010). Mood was evaluated at two points during the experiment: directly before Round 1 of the creativity tests, and directly after Round 2. The mood items were intended to assess the change in participant's mood before and after receiving performance feedback. All participants completed five bipolar mood-dimension scales including: Happy-Sad, Angry-Pleased, Frustrated-Proud, Relaxed-Anxious, and Hopeless-Determined. Each bipolar dimension was measured on a scale of 0 to 100 corresponding to the polar emotions.

Manipulation check. Participants were asked whether they had seen and taken any of the creativity tasks I had given them at any point before to make sure that they did not already know the answers to the questions, which would affect the data.

Results

Data Preparation.

Each of the creativity measures was coded using the particular rubric associated with scoring that particular task.

Alternate Uses Task.

Using the original rubric for scoring this task, responses were coded for fluency (number of responses), and flexibility (the number of categories within

responses). The other two rubrics, elaboration (length of explanation), and originality were not coded as participants had only been asked to specify one or two words and because I was not interested in finding the most original participant, but rather only how they improved. Only 'flexibility' was used as a final measure for analyses as it most accurately represented the quality and quantity of the answers. On average, participants came up with 4.7 alternate uses at Time 1 and 4.6 alternate uses at Time 2.

Remote Associations (RAT).

Each participant was assigned 1 point for each RAT that they answered correctly for a total number out of 15 possible groups at Time 1 and Time 2. On average participants got 3.35 out of 15 correct at Time 1 and 3.85 at Time 2.

Cognitive Reflection (CRT).

Correct answers were coded as 'right' and all other answers as 'wrong.' On average participants got the answer right 57% of the time at Time 1 and 53% of the time at Time 2.

Manipulation Check.

Fewer than 20% of participants indicated that they had taken the Alternate Uses Test or the Remote Associations Test before. However, 65% of participants reported having taken the Cognitive Reflection Task.

Performance Quality Analyses.

Performance quality was measured by taking the difference between Time 1 and Time 2 for each of the three tasks. Each of the tasks was measured separately at Time 1 and 2, and then analyzed together first using feedback accuracy as the primary predictor in a repeated measures ANOVA.

Time. I had predicted that that there would be an overall positive difference between Time 1 and Time 2, however, the analysis did not confirm this hypothesis and the multivariate model for both continuous measures was not significant, $F(2, 38) = 1$, $p = .454$. Examining the individual tasks revealed that the Alternate Uses Task was not significantly better at Time 2, $F(1, 39) = .73$, $p = .57$ though the mean was higher at Time 2 ($M = 4.7$, $SD = 2.1$) than at Time 1 ($M = 4.57$, $SD = 2.15$). The Remote Association Task was not significant either, $F(1, 39) = 1.48$, $p = .230$, though participants appeared to do somewhat better at Time 2 ($M = 3.85$, $SD = 2.94$) than at Time 1 ($M = 3.35$, $SD = 2.58$).

I additionally used a McNemar's test for a within subjects design to compare the proportion of right answers to wrong answers for the CRT at Time 1 and 2, and found that there was no significant difference, X^2 ($N = 40$), $p = .79$. Because 65% of participants had indicated in the manipulation check that they had seen the CRT items before on several other surveys (compared to < 20% for the other measures), for all subsequent analyses this task was removed as a creativity measure.

Participant gender. Overall, there was no significant difference between male and female performance on the items (all p's > .2). Additionally, there was no significant difference between Time 1 and 2 between male and female participants (all p's > .2).

Feedback accuracy. Overall, there was no effect of feedback accuracy on performance change, (all p's > .2), however, it appeared that at Time 2, at least for the Alternate Uses Task, people in the positive feedback condition were doing slightly worse ($M = 4.14$, $SD = 1.59$) than those in the negative feedback condition ($M = 5.32$,

SD = 2.40) (see Table 5 for mean performance change in each condition).

Table 5.

Performance on creativity tasks depending on feedback accuracy and participant gender.

Task	Feedback (Male)		Feedback (Female)	
	Negative Mean (SD)	Positive Mean (SD)	Negative Mean (SD)	Positive Mean (SD)
RAT	.385 (3.66)	.091 (1.92)	1.333 (1.96)	.600 (2.06)
AUS	-.231 (2.01)	.273 (2.49)	1.166 (3.81)	-.200 (1.22)

Note: The means represent the positive and negative change in performance on each task from round 1 to round 2, for both creativity measures (all p's > .2).

Motivation. Performance effort (motivation) was measured in two ways: (1) by whether or not participants had selected to take an extra round of measures and thus extra time in order to improve their scores and (2) by the amount of motivation participants indicated they had to better on Round 2 on a one to ten scale.

Extra round choice. Only 30% of participants indicated that they would like to take the extra round of tests in order to improve their scores. A binomial generalized linear model analysis indicated that there was no significant difference between feedback accuracy conditions (p = .627) or gender (p = .109), nor was there a significant interaction between the two (p = .942). However, further Chi Square tests for gender differences indicated that a majority of the participants who had chosen to do an extra round were female (58.3%).

Motivation intensity. In order to examine the level of motivation to do better on Round 2, I first examined motivation regardless of the choice to do an extra round. There were no significant differences for amount of motivation or interaction between

gender and feedback type (all p's > .4). Next, I split the data file by choice to take the extra round to see if there were differences depending on what participants had chosen. Again, there were no significant differences in the level of motivation between gender and by feedback type if participants had chosen the extra round (all p's > .2) with the average motivation for both groups (extra rounds or not) approximately 7/10.

Emotions.

All five bipolar emotion dimension scales were entered into a repeated measures ANOVA to examine mood change from Time 1 to Time 2. The multivariate analysis was significant, $F(5, 35) = 2.89$, $p < .05$ and univariate analyses revealed that this was due in particular to the Happy – Sad dimension, $F(1, 39) = 5.8$, $p < .05$, the Angry – Pleased dimension, $F(1, 39) = 7.79$, $p < .01$, and the Relaxed – Anxious Dimension, $F(1, 39) = 14.8$, $p < .001$ (all other p's > .7). For each of the measures, participants indicated that they were feeling more negative at Time 2 than at Time 1.

While there were no significant differences between gender or between feedback type overall (p's > .2), the interaction analyses indicated that there were differences for some of the items and interactions between feedback type and time, $F(5, 32) = 3.49$, $p < .05$, and a three way interaction between participant gender, feedback and time, $F(5, 32) = 3.26$, $p < .05$. The item that was primarily significant at this level was the Relaxed – Anxious dimension. Participants reported that they were significantly more relaxed, $t(39) = -3.85$, $p < .005$ at Time 1 ($M = 31.32$, $SD = 25$) and more anxious at Time 2 ($M = 47.12$, $SD = 24.9$). Analyses using difference scores between Time 1 and 2 for this particular dimension also indicated that participants had become less anxious at Time 2 when given positive feedback (mood change: $M = 5.8$,

SD = 19) than when given negative feedback (26.84, SD = 28.45), F(1, 36) = 10.55, p < .005. And, though not significant, F(1, 36) = 2.72, p =.108, it appeared that women showed slightly greater change in anxiety when given negative feedback (M = 41.84, SD = 22.95), than did men when they were given negative feedback (M = 20, SD = 28.9). Thus there is a possibility that negative feedback could impact anxiety, especially for women, however the data are inconclusive.

Expectations.

There were two measures of expectations: (1) an absolute measure: participants rated how well they thought they did after Round 1 and 2 on a 1 – 9 point scale, and (2) a relative measure: ranked what percentile they thought they fell in.

Absolute expectations. Overall, participants were not very confident about how they did at either time, reporting scores below but close to the mean of 4.5.[18] Using a repeated measures ANOVA, results indicated that participants believed they had done significantly worse at Time 2 (M = 2.49, SD = .77) than they had at Time 1 (M = 4.11, SD = 1.99), regardless of feedback type, or gender, F(1, 33) = 19.87, p < .001. The within subject interactions with gender and feedback condition were not significant (all p's > .4).

Relative expectations. The effect of time on percentile was also significant, F(1, 36) = 5.3, p < .05, with participants reporting that they thought they fell into the 42d percentile (SD = 22.83), slightly lower than average, however, dissimilar from the from the absolute measure, participants appeared to believe that they fell into a higher, 48th percentile after round 2 (SD = 20.29). There were no feedback or gender effects

[18] This validates the design intentions of creating a mostly difficult task.

(all p's > .5). Thus it appears that though people believed they did worse after round 2, they also thought other people did worse as well.

Discussion

Although negative feedback can be emotionally difficult to receive, especially as compared to positive feedback, the long-term benefits of critical feedback can be useful for attempts at self-improvement (Freitas, Salovey, & Liberman, 2001; Mischel, Ebbesen, & Zeiss, 1973; Trope & Neter, 1994). In Study 5 I sought to demonstrate that inaccurate but positive feedback, on the other hand, can be damaging for both motivation and quality of subsequent performance, and additionally I sought to examine whether the gender of the recipient of the feedback would qualify this prediction.

The results from Study 5, however, did not support the primary hypotheses nor most of the secondary hypotheses. Participants who had been given three sets of creativity tasks to take before and after feedback (a) did not do any better on the second round, (b) did not do significantly better or worse depending on the type of feedback they were given. Furthermore, the gender of the participants did not appear to have a significant effect on performance and did not interact with feedback type either.[19]

However, it is important to note that the sample size for this experiment was somewhat low (N = 40), given the between subject variables I was interested in. That the general hypothesis was not supported within subjects is cause for concern, but it is

[19] Fromme, Mercadal & Mercadal (1976) found that people produced more RAT's after negative feedback than after positive feedback.

impossible given the sample size to draw even tentative conclusions from the between-subjects interaction and results. A repeated measures design with two independent between-subjects variables, and a minimum power of $(1-\beta) = .80$, and medium effect size $f = .25$, would necessitate $N = 212$ participants. As such the post-hoc power analysis of 40 participants, with a medium effect of $f = .25$ indicated a power of $(1-\beta) = .27$.

That said, there were some analyses that were promising and could possibly become significant with the addition of more participants assuming the pattern remained stable. In terms of primary hypotheses, it did appear as though participants were improving somewhat on the creativity measures on the second round, which makes sense considering that practice usually improves performance (Toppings, Samuels, & Paul, 2007). Additionally, though I am hesitant to draw conclusions based on this amount of data, it did appear as though performance did not increase as much on the second round after positive (and less accurate) feedback as it did after being given negative (and more accurate) feedback. Motivation to put in extra effort and do extra trials, on the other hand, did not seem to differ by gender or by feedback type, and overall was rather low (to be expected from participants for whom time is precious), although most participants reported being motivated to better on the second round.

In terms of secondary hypotheses, as predicted, there was a change in mood such that overall, participants felt more negative after the second round of tests than they had after the first round, but this negativity was particularly driven by feelings of anxiety, and also primarily by women though, surprisingly, not by feedback condition.

The gender effect is somewhat difficult to decipher in this context as it could be another case of women being more comfortable expressing their emotions, or it could be that an accompanying lack of self-confidence on creativity tasks or test-taking in general (Nguyen & Ryan, 2008; Steele, 1997) which is likely based on perceptions of having done slightly worse at Time 1 (expectations measure). However, because there was no accompanying decrement in performance for women at Time 2, it likely did not affect performance and makes it difficult to draw conclusions about both the emotional and long-term impacts of feedback on women.

While the data are not conclusive, there are reasons that the experiment, apart from being underpowered, may not have resulted in more significant findings. First, there is a possibility that the measures that were used did not accurately gauge performance. Though one's own intelligence is usually intrinsically important, three short creativity tasks may not be enough to convince participants that their actual creativity or intelligence is being measured. That is, it could be that, especially over the medium MTurk, people did not feel that these tasks were actually diagnostic of their abilities. In order to truly improve and want to improve, one has to find the task important, and the feedback diagnostic (see Sansone, 1998) and this 'importance' problem is corroborated by the relatively few people who selected to do extra rounds in order to improve their scores (< 30%).

Second, there appear to be no known gender differences in creative ability (Baer & Kaufman, 2008), however, because there are gender stereotypes concerning creativity (Kaufman, 2006) I may have opened up the possibility of stereotype threat (Steele, 1997). That is, women may have felt at a disadvantage to begin with, which

may have further demotivated them when it came to getting critical feedback. This idea is corroborated by the motivation data which suggests that men started out with more confidence in their abilities at Time 1 but less confidence at Time 2 than women, suggesting that they had farther to fall when receiving disconfirming feedback. However, the motivation data suggest that if anything, women were more likely than men to choose to do an extra round of tests in order to improve their scores, potentially highlighting a perceived threat that needed addressing.

Because the data were less than conclusive, and also because both the task and the setting created a less-than-optimal set-up for engagement and desire to improve among participants, I conclude that a similar experiment should be conducted either in the laboratory or in the field, using tasks and dimensions perceived as important to participants, motivating them to want to improve. One of the difficulties in understanding such negative consequences is that decrements in performance can often unravel slowly over time, or improve slowly over time, and as mentioned, the emotional impacts of critical feedback can painful in the immediate aftermath, but be beneficial in the long run. Thus, a short-term consequences experiment like this one would most likely tell us more about the immediate emotional impact of feedback and less about the long-term benefits (or decrements) to performance. In order to get a fuller picture of any meaningful consequences, data would have to be sampled not just at the time of feedback but also in the (as of yet undetermined amount of) time after the feedback was given.

CHAPTER 6

Study 6

An Intervention: Can Regulatory Focus Attenuate Feedback Bias?

The purpose of Study 6 was to test a possible intervention to be used in the workplace to attenuate the feedback bias. The prior studies and literature suggest that feedback is characterized by competing goals, and that people are attuned to the emotions of the recipients. This suggests that people tell white lies because they want to avoid harming those who would receive undesirable information I employed regulatory focus as the driving theory in the design of this study, testing the effects of removing subjects from an immediate, concrete, avoidance mindset and replacing it with a long-term, abstract, approach mindset. I hypothesized that this change in mindset would lead subjects to focus on long-term benefits of honesty, and in turn focus less on avoidance of more immediate pain associated with critical feedback.

In regulatory focus theory (Higgins, 1997) two different motivational foci characterize the competition between goals and the eventual dominance of one over the other. Individuals in a promotion focus tend to use forward-thinking strategies aimed at accomplishing, approaching, and responding, which other researchers have found leads to both future and global thinking (e.g. Friedman & Forster, 2001). In contrast, individuals in a preventions focus, may use a strategy of vigilance and

defensiveness and concentrate on concrete, immediate discomforts to avoid. This behavior can lead individuals to focus only on the short-term consequences of actions.

Thus, in Study 6 I tested the prediction that one's general regulatory mindset will influence both his or her behavioral approach to giving feedback and beliefs about the relative value of honesty. Focusing on the low-level, concrete experiences (and their associated rewards and punishments of critical feedback) should lead individuals to avoid emotional repercussions and choose the most comfortable and easiest action: telling a white lie. However, by focusing on the high-level, abstract goals of an action and the future outcomes for both one's own and another's well-being, individuals should choose an action providing the greatest long-term benefit: accuracy (cf., Freitas, Gollwitzer, & Trope, 2004; Liberman & Trope, 1998; Vallacher & Kaufman, 1996;).

Additionally, if women are more often the targets of white lies, this might suggest that this prevention focus is even more salient (or wins out more heavily) among those giving feedback to women. This may be the case because individuals wish to avoid hurting women who are perceived to be more emotional. Therefore, I further hypothesized that putting people in a prevention mindset (as opposed to a promotion mindset) would (a) lead them to lie more, and (b) lead them to lie more often to women than to men.

However, there may be a caveat to this hypothesis. Research has shown that gender can have an effect on confrontation and conflict management styles from a young age, with boys showing more aggression in response to conflict, and girls showing more withdrawal at a young age, and then more prosociality in their teenage

years (Lindeman et al., 1997). And, though Eagly and Johnson (1990) found in a meta-analysis that in organizational settings the leadership styles of men and women did not differ greatly, in laboratory settings, where 'managers' are not selected or self-selected into a leadership position, there are some gender-stereotypic leadership traits present. In the latter context, women tend to focus more on an interpersonal leadership style, while men tend to employ a more directive style. Thus I predicted some subtle gender differences, with women demonstrating more avoidance, and also more concern about others' feelings overall, resulting in a diminished effect of the induced approach motivation manipulation.

Finally, powerful people tend to adopt a more promotion-focused regulatory strategy, while those without power tend to be more prevention-focused (Keltner et al., 2003; Sassenberg et al., 2007), as do predominantly independent self-construed individuals (Lee et al. 2000). Because women as a social group typically have less power than men (see Ridgeway, 2001 for a review), it is possible that regulatory focus may have a differential effect on male and female participants when it comes to giving uncomfortable feedback. Thus, I additionally predicted that female participants would be particularly affected by the prevention focus and lie more overall.

Finally, I expected that the original results from Study 1a would be replicated such that, overall, people would tell more white lies to the female target than to the male target.

Method

Participants.

Two hundred and twelve participants were recruited through Amazon's Mechanical Turk crowdsourcing website and compensated $0.05 per minute. 118 (56%) of the participants were female, and 91 (43%) were male, and one person identified as 'other.' 70% of participants fell between the ages of 20 and 40, and 30% fell between the ages of 40 and 60.

Procedure.

All participants were administered a ten minute survey on the Qualtrics survey platform in which I manipulated both regulatory focus and target gender and collected participant gender in a 2 (Approach/Avoid) X 2 (Male/Female target) X 2 (Male/Female participant) between subjects design. Before beginning the experiment each participant read that they would be completing two different sections that were not related to each other. They read that the first section would ask them to complete some word and visual search tasks, and the second section would ask them to answer hypothetical questions based on a particular workplace scenario. Once the participants had been shown the instructions and given consent for their participation, they were asked to complete both sections of the survey.

Regulatory focus mindset induction. Participants were randomly assigned to either a 'prevention' focus condition or a 'promotion' focus condition. In order to manipulate regulatory focus and induce a specific mindset, I employed four different types of manipulations that represented four different but similar components of regulatory focus, (1) a directional (approach/avoid) component, (2) a vigilance

(rewards/consequences) component, (3) a temporal goal component (present/future) and (4) a processing style (local/global) component. Each of the following manipulations was designed to induce either a prevention mindset or a promotion mindset.

***Focus manipulation 1 (Directional)*:** The first manipulation was a directional manipulation modeled after Seibt & Forster's (2004, Study 1) paradigm. In this study, the authors asked participants to read twenty sentences describing a random person's implementation intentions, ten reflecting approach motivations (e.g., "because I wanted to buy something for my friend I went shopping in town") and ten representing avoidance motivations (e.g., "Because I did not want to say something stupid I did not say anything in class"). Though the authors used these sentences as a measure for memory of approach and avoidance related content, I used their method as a manipulation.

In our study, I manipulated approach or avoidance by using three new sentences of each directional motivation. However, unlike Seibt & Forster (2004), rather than selecting from various pre-written responses, participants were asked to complete the last word of each sentence. The sentences additionally differed in tense, so that the promotion sentences were future-oriented, and the avoidance sentences were present-oriented. The completion words were predictable based on the sentence content, and were designed in this way to convince participants that they were taking part in an active task that was measuring something (though it was only meant to induce their full reading of each sentences and provide the necessary regulatory manipulation). An example of each of the sentences used is below:

Promotion Focus:
"Because I want to live for a long time, I will stay healthy by going to the ____"

Prevention Focus:
"Because I do not like to seem uneducated, everyday I read the ____"

Focus manipulation 2 (Vigilance). The second manipulation was designed to further solidify the promotion or prevention manipulation by making the goal of a task either vigilant/defensive or rewarding/ approach-oriented. This task was modeled after Baas, De Dreu, and Nijstad (2011, Studies 2 & 3), who asked participants to control a mouse inside a computerized maze under two discrete circumstances: one with a promotion and another with a prevention focus In the former version, the participant's goal was to negotiate the maze to attain a piece of cheese (reward). In the latter, participants were asked to negotiate the maze in order to avoid an owl that was hovering outside (vigilance).

In our manipulation, participants saw a safari-themed find-it photo puzzle that had multiple animals in the picture and a person holding a camera in the middle (see picture below). [20]

[20] By Ravensburger Puzzle: http://www.seriouspuzzles.com/i5509.asp

In the promotion focus, participants were told that, as a photographer, their goal was to photograph all the most dangerous animals in the reserve. Their mission was to click on all the locations of the big cats (e.g. lions, leopards, cheetahs), thus inducing a promotion focus (reward):

> "*Imagine you are a photographer on safari. Your goal is to get as many pictures of the exciting big cats that you can so you can brag to your friends. On the next page you will see a picture. Please click on each of the cats that you intend to photograph and try to find them all. (The screen will highlight where you clicked). You have twenty seconds to identify and collect them all!"*

In the prevention focus, participants were told that they had to identify the big cats in order to avoid being eaten by them during their photo shoot, thus inducing a prevention focus (vigilance).

> "*Imagine you are a photographer on Safari. Your goal is to take pictures of all the exciting animals without being eaten. On the next page you will see a picture. On this picture, so that you can ensure you won't be eaten and will be safe, you need to identify where all of the dangerous big cats are by clicking on all of the locations where there are big cats. (The screen will highlight where you clicked). You have twenty seconds to identify evade them all!"*

Participants were asked to click on the screen until they had found all of the

cats, after which they could continue. If the participant did not identify all the cats, the screen auto-advanced after twenty seconds. The task was designed to be easily accomplished so that goal completion was not an issue, and only mindset was manipulated.

Focus manipulation 3 (Temporal). The third manipulation was intended to continue the promotion or prevention focus by inducing a focus on either broad future goals or concrete present goals. Several studies have found that people are more likely to adopt a promotion focus when contemplating the future, whereas they are more likely to adopt an avoidance focus when thinking about the present (e.g., Ariely & Zakay, 2001). In order to induce different temporal mindsets I asked participants to list either three of their life goals (promotion), or three tasks they had completed that day (prevention).

Focus manipulation 4 (processing style). The final manipulation intended to influence the processing style of the participants to be more global, and focused on the outcome of the group as a whole (promotion) or more local, focusing on the individual in a specific context (prevention). In order to induce these mindsets participants were given a general statement and were asked how much they agreed with that statement on a scale of 1 (*not at all*) to 10 (*completely*). The questions were designed so that they were generally agreeable in nature when presented individually and not comparatively. That is, they were designed so that people would most likely agree at least somewhat with them, and solidify the mindset I intended to manipulate.

Global Processing Style (Promotion):
"When considering a course of action, it is best to think about all

possible cases, the broader picture, and the benefit to the whole group."

Local Processing Style (Prevention):
"When considering a course of action, it is best to think about each case individually, the situation at hand, and the benefit to the person in question."

Target gender manipulation. After completing the regulatory focus measure, participants were asked to read a scenario similar to that of Study 3 in this manuscript, in which they were asked to imagine themselves as the manager of a company who had to give monthly feedback to employees. They were also told that their job as manager was to give constructive feedback to their employees for the good of the company.

The participants then were asked to read a hypothetical scenario about an underperforming employee who had been on the job for approximately six months whose performance was worse than expected. The gender of the employee was manipulated using a subtle name change and varied between subjects such that half the participants read about a female target (Sarah) and the other half read about a male target (Andrew).[21] Participants read that Sarah/Andrew might lose her/his job if he/she kept underperforming and were objectively doing worse than the other employees at the business. Participants were then told that today was monthly feedback session, and that they would be asked to give feedback to Sarah/Andrew about her/his performance. I added these details so that participants would understand that the employee was performing poorly by objective measures.

[21] These names were chosen from the top twenty baby names of the 1980's and 1990's so as to match as much as possible in sound, age, and popularity, thus reducing the possibility that participants would draw extraneous conclusions from the names themselves.

Primary measure: feedback accuracy. In order to measure the effect of regulatory focus on accuracy and to replicate the findings of Study 1 through 3, participants were shown the same six feedback statements that were used in Studies 2 through 3, ranging from the most truthful and the harshest to the most accurate and nicest feedback. Again, each of the feedback statements mirrored the content of the original scenario that had been described so as to make it more difficult for participants to stray from what they understood to be the objective truth about the employee's performance. Participants were asked to "choose which option you think you would *most likely* pick in this situation, not necessarily which option you think is the right one (though they might be the same for you)."

Accuracy/self-esteem trade-off. In order to assess the participants' explicit views about accuracy, I used a measure taken directly from Freitas, Gollwitzer, and Trope (2004, Study 2a & 2b) to assess the belief in the ideal amount of accuracy and self-esteem tradeoff during feedback. The measure was introduced as follows:

> *''Sometimes the most accurate feedback will also make a person feel quite upset, or disappointed with themselves. We can address this issue by providing feedback that is accurate only to the extent that it does not make them too upset or disappointed.''*

To answer the question, ''In your own opinion, how should I approach this issue?'' participants used a ten point scale to indicate their degree of preference with the anchors (1) ''*greatly limit accuracy in order to preserve self-esteem*'' and (10)''*maximize accuracy even if it greatly damages self-esteem*.' I did not have an explicit prediction for this measure as even explicit statements about beliefs often differ from actual behavior. That is, it is possible that though participants do tell more

white lies to female targets, this measure would not yield a significant difference in belief as explicit belief can be tied to social norms, including honesty.

Emotion measure. In order to get a sense of whether people would find the target's discomfort with receiving feedback more salient in the prevention rather than the promotion focus, and whether it would be the case more for women than for men, I additionally asked the participants to "Please list a feeling or emotion you think the employee might have felt as they received performance feedback," and subsequently to rate how strongly the participant believed that the employee experienced that emotion (Freitas, Gollwitzer, & Trope 2004, Study 2a & 2b). I expected that perceived emotional intensity would be greater for female targets for participants in the prevention focus condition, because what participants would be avoiding when delivering feedback is hurting other people's feelings (especially women's).

Strength & weakness feedback preferences. In order to evaluate whether a promotion-based mind set would make salient the more abstract benefit of receiving potentially disheartening but ultimately valuable negative feedback, I additionally measured both the beliefs of the participants about the value of providing distinct types of feedback and also their anticipation of other's desires for that kind of feedback. Using a measure also drawn from Freitas, Gollwitzer, & Trope (2004, Study 1), participants were asked to use ten-point scales from (1) '*not at all*;' to (10]) '*extremely'* to answer two questions concerning their preferences for feedback provision for others: ''How worthwhile do you think it is to provide this employee with feedback concerning their strengths [& weaknesses]?'' and two questions concerning their inferences of others' feedback preferences: ''How interested do you

think this employee would be in finding out about their strengths [& weaknesses]?''

I did not have specific predictions for this measure, as any one of the three independent variables could affect the measure in different ways. It is possible that overall, people would believe that weakness feedback would be more worthwhile to male targets than to female targets, and especially so in the prevention focus. However participant gender could also moderate the interaction effects, and beliefs about what targets want to hear could potentially differ from what participants thought was worthwhile. Again, because I was asking for their explicit beliefs, I may not be tapping into actual beliefs but instead into normative beliefs.

The order of all three of these measures was randomized across participants. As a final question, I also asked participants to indicate whether they had ever taken a survey with similar content. This was to ensure that I would be able to filter out participants who taken a very similar survey which had been running at the same time on MTurk.

Results

Preliminary Analyses.

Of the 212 participants recruited, 34 indicated that they believed they had taken a similar feedback survey recently. Thus for the following analyses, these repeat participants were removed from the sample, leaving 178 participants for analysis. No other participants were removed. [22]

Before beginning analysis on the outlined measures, the efficacy of the

[22] When all participants were analyzed, it did weaken significant effects to an extent; however, this is mostly likely due to the fact they had recently taken another survey on the same topic and with the same manipulation that had been recruiting simultaneously, thus biasing their answers.

manipulations was checked (where possible) to make sure that the manipulations were working as intended. For the first regulatory focus manipulation, sentence completion was checked to make sure that nearly all of the words used to complete the highly predictable prevention and promotion sentences were similar. There were only a few deviations from the expected word and nearly all participants had provided the same answer for the sentence completion. For example, for the sentence "*"Because I do not like to seem uneducated, everyday I read the ____,"* nearly all participants provided the word "newspaper" though some provided answers such as "scientific journals" or "twitter." These aberrations should have had no impact on the data, as the general intent was the directional focus of the sentence rather than the word itself.

The second focus manipulation, which asked participants to either avoid or pursue specific animals in a picture puzzle, had a 100% accuracy rate, and participants appeared to have no trouble clicking on the animal locations within the allotted time. The third focus manipulation asked participants to either list three broad life goals they would like to accomplish, or three tasks they had completed that day. Nearly all of the participants in the promotion condition had listed broad goals such as "financial freedom," or "happiness" or "get a PhD." And several of those in the prevention condition listed either home-related tasks they had done that day, such as "sent emails" or "brushed my teeth" while several others listed tasks that they had done in the study itself, such as "clicked on big cats" or "completed sentences."[23]

The fourth focus manipulation had asked participants to agree with either a broad collectivist statement or a specific individualized statement with the intent of

[23] If you would like to see a list of the individual qualitative answers, please contact the author.

inducing agreement and thus mindset. The mean agreement with the collectivist statement on a scale of 1-10 was M = 7.7 (1.9), and the mean agreement with the individualist statement was M = 7.9 (1.7), indicating a high level of agreement with both statements.

Primary Analyses.

Of primary interest in this study was (a) whether regulatory focus would have an effect on lying in general, (b) whether target gender and regulatory focus would interact such the women were especially told white lies in the prevention condition and (c) whether participant gender would moderate the results. I additionally had predicted that the general effect of a feedback bias towards women would be replicated.

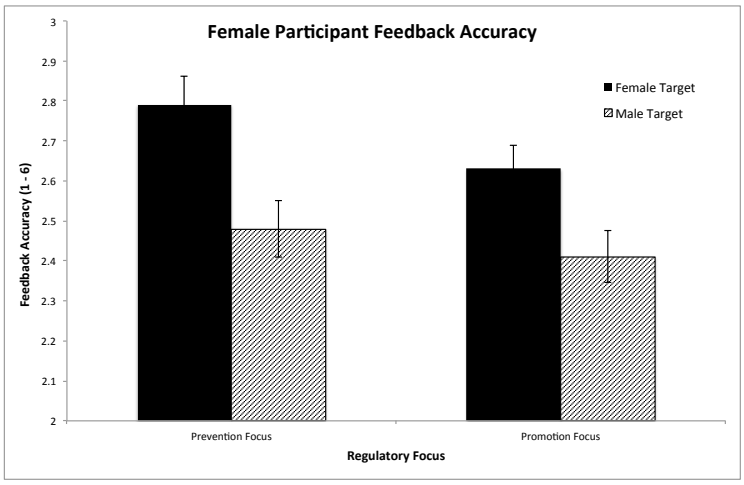
In order to analyze the primary measure of actual feedback accuracy in the hypothetical workplace scenario, a univariate factorial ANOVA was used to analyze the main effects of target gender and focus on feedback accuracy. The responses were heavily skewed towards the low end of the response choices with Options 2 and 3 as the modal feedback choices (more truthful). However, I found that there was a significant main effect of target gender on feedback accuracy, $F(1, 181) = 9$, $p < .005$, $\eta_p^2 = .05$, such that the female target ("Sarah") was given nicer feedback overall ($M = 2.88$, $SD = .97$) than was the male target ("Andrew") who was given more accurate performance feedback ($M = 2.46$, $SD = .89$). Thus the general gender bias effect from the cornerstone study was replicated.

The second hypothesis was not supported, however, and there was neither a main effect of regulatory focus, $F(1, 181) = .09$, $p = .7$ nor an interaction between

focus and target gender, $F(1, 181) = .08$, $p = .8$. However, when I entered participant gender as an independent variable into the factorial model, the main effect of target gender was still significant, as was the (marginal) main effect of participant gender on lying, $F(1, 176) = 6.5$, $p < .005$, $\eta_p^2 = .07$. That is, male participants told marginally more white lies ($M = 2.77$, $SD = .97$) than did female participants ($M = 2.57$, $SD = .89$). Examining male participants and female participants separately, I found that male participants also told significantly more lies regardless of focus, $F(1, 73) = 7.1$, $p < .01$, $\eta_p^2 = .08$, to female targets ($M = 3.0$, $SD = 1.0$) than to male targets ($M = 2.49$, $SD = .87$), whereas there was no significant difference by target gender for female participants ($p > .1$).

There was no interaction between participant gender and regulatory focus, $F(1, 176) = 2$, $p = .16$, despite male participants appearing to lie slightly more in the promotion focus ($M = 2.92$, $SD = 1$) than in the prevention focus ($M = 2.65$, $SD = .85$) condition, while female participants showed no difference between conditions. All other two-way interactions, including the Target Gender X Participant Gender interaction, were non-significant (all p's > .30.) And, the three-way interaction between focus, participant gender, and target gender was not significant $F(1, 176) = .79$, $p = .34$, though I have reported the means for possible interpretation (see Table 6 for all means by participant gender).

A



B

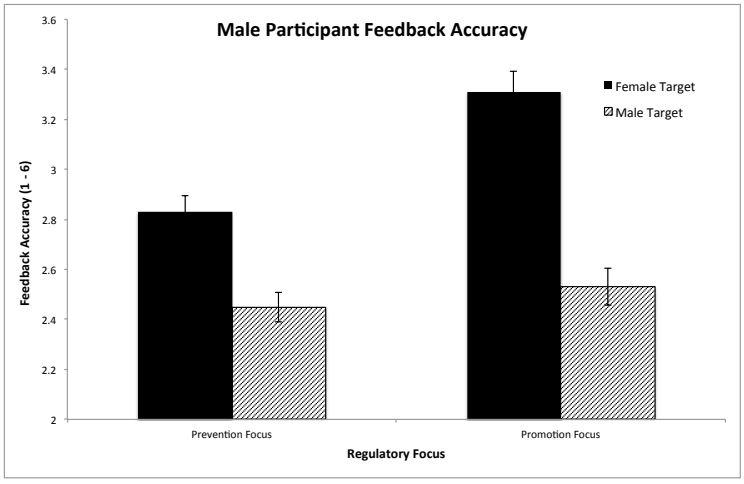


Figure 5. Graphs represent mean feedback accuracy with lower numbers signifying more truthful responses, and bars represent target gender in the different regulatory focus conditions, by (A) female participants (n =106) and (B) male participants (n = 77). Error bars indicate standard error from the mean. The interaction between focus and target gender is not significant for female or male participants. For (A) female participants target gender main effect is not significant, $p = .14$, and for (B) male participants, the target gender main effect is significant, $p < .01$.

Table 6

Feedback accuracy to male and female targets by participant gender and focus type.

	Male Participants (n = 77)		Female Participants (n = 106)	
	Prevention	Promotion	Prevention	Promotion
Female Target ("Sarah")	2.83 (.8)	3.31 (1.1)	2.79 (.9)	2.63 (.8)
Male Target ("Andrew")	2.45 (.8)	2.53 (.9)	2.48 (.9)	2.41 (.8)

Note. Means (SD) reported for three-way interaction between Participant Gender, Regulatory Focus, and Target Gender on feedback (lower number = greater accuracy). †p < .10. *p < .05. ** p < .01. *** p < .001.

Secondary Measures

Accuracy / self-esteem trade-off measure. This measure was added to assess whether participants would demonstrate an explicit belief that self-esteem was more important than accuracy in the prevention focus condition, and especially so for female targets. First, a correlation analysis was run in order to identify an existing relationship between the desire to boost self-esteem and lying behavior. The analysis revealed a significant negative correlation between altruistic lying and explicit beliefs in accuracy, $r = -.031$, $p < .001$ so that people who told more white lies were also more likely to fall on the side of self-esteem in the self esteem/accuracy tradeoff.

The factorial ANOVA results, however, indicated that people's explicit beliefs did not differ between gender conditions, $F(1, 181) = .01$, $P = .9$, or focus condition $F(1, 181) = .85$, $P = .35$, nor was there support for an interaction between the conditions $F(1, 181) = .76$, $P = .38$. However, the conditions appeared to be trending in an interesting direction, though not statistically significant. In the prevention focus, participants indicated a higher belief that accuracy was more important for the female

target (M = 5.7 SD = .9) than in the promotion condition (M = 5.4, SD = 1.4). When participant gender was entered into the factorial ANOVA, while none of the main effects or two way interactions were significant (all p's > .3), the three way interaction between target gender, focus, and participant gender was approaching significance, F(1, 176) = 2.1, P = .14, with males believing more accuracy than self-esteem was important for the female target in the prevention focus (M = 5.8, SD = .79) than the promotion focus (M = 5.2, SD = 1.3) and female participants believing more accuracy than self-esteem was important for the male target in the prevention condition (M = 5.9, SD =1.3) over the promotion condition (M = M5.5, SD = 1.2).

In order to identify which participant gender was driving the near significant interaction, I conducted univariate ANOVA's at each level of variable 'participant gender.' The split file analyses revealed that male participants were driving the effect, where the interaction between regulatory focus and target gender was verging on significance, F(1, 73) = 4.9, p =.10, while female participants showed no significant interaction between focus and target gender: F(1, 102) = .23, p =.73.

Strength & Weaknesses Feedback. In order to examine (a) how worthwhile participants thought giving strength- versus weakness-based feedback to the targets, and (b) how interested they believed the targets would be to hear such strengths- and weakness-based feedback (on a scale of 1 – 10), I first ran a multivariate factorial ANOVA to determine the main effects of target gender and regulatory focus and the interactions between the two. Though the whole model main effects and interaction were not significant (all p's > .4), there were significant differences in the univariate tests. The only *nearly* significant difference, F(1, 181) = 2.86, p =.09, was a difference

in the amount by which target genders would be interested in finding out about their own weaknesses. Participants indicated that they believed the male target would be more interested in finding out about his weaknesses (M = 6.1, SD = 2.8) than the female target (M = 5.4, SD = 2.7). However, there was a significant difference F(1, 181) = 4.1, p <.05, in participants' belief that female targets desired less feedback overall (M = 6.7) than did males (M = 7.3).

When I entered participant gender into the model, multivariate whole model tests indicated a significant effect for participant gender on all measures: F(4, 178) = 2.8, p < .05, η_p^2 = .06, with female participants believing that targets both need more feedback overall and that they desire more feedback overall than male participants. This was the case for three of the four items. However there were no interactions with target gender or focus condition (see table 7a for overall means by participant gender and table 7b for means by target gender).

Table 7a

Belief about importance and desire of strength and weakness based feedback by participant gender

	Feedback	Male Participant Mean(SD)	Female Participant Mean(SD)
How worthwhile?	Strengths†	7.35 (2.1)	7.95 (2.3)
	Weaknesses	7.25 (2.3)	7.62 (2.4)
How interested?	Strengths***	7.77 (2.1)	8.62 (1.7)
	Weaknesses†	5.33 (2.5)	6.0 (2.9)

†p < .10. *p < .05. ** p < .01. *** p < .001.

Table 7b

Belief about importance and desire of strength and weakness based feedback by target gender

	Feedback	Male Target Mean(SD)	Female Target Mean(SD)
How worthwhile?	Strengths	7.82 (2.2)	7.53 (2.2)
	Weaknesses	7.59 (2.3)	7.30 (2.5)
How interested?	Strengths	8.42 (1.9)	8.06 (1.9)
	Weaknesses†	6.10 (2.8)	5.40 (2.7)

Note. Means are for all participants. †p < .10. *p < .05. ** p < .01. *** p < .001.

Perceived emotions. Before analyzing any group differences in perceptions of target's emotional states, each of the emotions listed was coded and categorized in two ways: valence (positive or negative) and arousal/activation level ('approach' emotions or 'avoidance' emotions) (see Feldman-Barrett & Russell, 1998; Higgins, 1998; Brockner & Higgins 2000). 89.2% of the listed emotions were categorized as negative, with 8.6% as positive and 2.2% uncategorizable. 20.5% were categorized as 'approach' type emotions (e.g. "angry"), and 76.2% as avoidance type emotions (e.g. "sad") with 3.2% uncategorizable.

The percentage of negative versus positive emotions varied significantly by target gender condition, X^2 (1, N = 181) = 4.1, p < .05, such that 95.5% of male target words were negative while 87% of female target words were negative. Approach and avoid emotions were not significantly different by target gender or by regulatory focus condition, nor was there an interaction between the two (all p's > .3).

Of primary interest, however, was how participants rated the intensity of the emotions they listed by target gender and condition. While emotional intensity was not significantly different for approach or avoid words (p >.5), participants did rate the

emotional intensity of the negative words (M = 7.9, SD = 1.5) as significantly higher than the intensity of the positive words: (M = 7, SD = 1.2), $t(179) = 2.32$, $p < .05$.

Because I was primarily interested in the hypothesis that perceived negative emotional intensity would affect lying behavior, I filtered out positive emotion cases for the following analyses. Using a factorial ANCOVA, with target gender and focus as factors, and feedback accuracy as a covariate, I found that while none of the main effects were significant (all p's > .3), nor was the Focus X Target Gender interaction, $F(1, 160) = .9$, $p = .34$, the pattern of results followed our prediction that participants in the prevention focus would perceive the female target to feel more emotional intensity than would participants in the promotion focus (see Table 8).

Table 8

Perceived emotional intensity of target by participant gender and focus condition.

Focus	Male Target Mean (SD)	Female Target Mean (SD)
Prevention	7.93 (1.5)	8.11 (1.6)
Promotion	7.93 (1.3)	7.67 (1.5)

Note. Means represent the perceived emotional intensity on a scale of one to ten.

Discussion

The main purpose of this study was to design and test an effective intervention, that of putting people into a promotion focus, to attenuate the gender bias during performance feedback. In order to do so, a regulatory focus manipulation was used with the principal hypothesis being that people in a promotion focus, who are more attuned to the broader and future benefits of critical feedback, are more likely to tell

the truth and thus dampen the gender-bias. I predicted that those in an avoidance focus would lie more overall, and particularly more to women because they were more attuned to immediate emotional distress. I also expected that the overall effect whereby women would be told more white lies than men would be replicated, and that participant gender might moderate any effects because of different default regulatory foci.

As I had predicted, I was able to replicate the main effect that white lies are told more often to women. However, regulatory focus did not appear to have an effect on lying. That is, a promotion focus did not lead people to lie less to either gender than did a prevention focus. There were, however, some interesting patterns regarding participant gender that, though not statistically significant, merit discussion for future experiments. Overall, male participants lied more to the female target "Sarah" than to "Andrew" and more than female participants did overall. However, female participants appeared to lie more than men while they were in the prevention as opposed to the promotion condition. These results, though not conclusive due to the lack of significance, possible echo the predictions made about female participant's sensitivity to prevention. These patterns suggest that when target empathy is manipulated, women were more likely to tell white lies than men (Lupoli, Jampol, & Oveis, *in preparation*).

Additionally, while explicit beliefs about the importance of accuracy versus self-esteem did correlate with lying behavior – that is, people who lied were more likely to place importance on self-esteem – they did not differ by regulatory focus of target gender. Similarly, regulatory focus did not have a general effect on whether

participants believed that strength- or weakness-based feedback were important for recipients or believed recipients desired these types of feedback. However, it did appear that participants believed that male targets desired more weakness-based feedback than did females.

Finally, I predicted that the perceived emotional intensity of the recipients would be more salient in a prevention focus, and (though not significant) it appeared that the data was trending that way; participants perceived Sarah to have more intense negative emotions in the prevention focus condition. However, the data also suggested that overall, male recipients were perceived to have more negative emotions than females, thus the emotionality predictions were only partially supported.

Overall, the intervention had mixed results in terms of mitigating white lies during feedback, and it is difficult to make any conclusive statements from the data. There is a possibility that the sample is slightly underpowered for a three-way (2 x 2 x 2) interaction, and a post-hoc analysis using an effects size f = .25, at the alpha = .05 level, indicated a critical F of 2.27 such that power was only somewhat low, $(1-\beta) = 0.65$, given the parameters. However, this most likely does not explain the null effect. One of the most likely problems is that the intervention of putting people into particular focus mindsets did not have the intended effect. Unfortunately, no manipulation check was conducted so it is difficult to know whether participants were indeed in a promotion or prevention mindset, and which if the manipulation items worked or did not. Thus, because the main predictions were not supported, this study merits replication in a laboratory setting with a larger sample size and with either manipulations that have been verified to induce intended mind-sets or manipulation

checks to assure their effectiveness.

CHAPTER 7

GENERAL DISSCUSSION

Much of the discussion about the "glass ceiling" and barriers for women in the workplace has moved away from explicit sexism and external factors that might affect women's achievement and success. Instead there has been a growing interest in identifying and understanding the more covert and implicit biases and behaviors that accumulate to maintain disadvantage. One such covert bias may be the way in which women are given performance feedback. Specifically, I theorized that because giving feedback can be difficult, and because people hold stereotypes about women's reactions to critical feedback, women are told more white lies about their performance than men.

Summary of Findings

There were several hypotheses that needed to be tested to validate this theory. First, a gender bias during performance feedback needed to be identified in order to support the hypothesis that women are communicated to differently than men. The results from Study 1a found that, indeed, when evaluating both women and men, people were significantly more likely to upwardly distort quantitative evaluations when giving direct feedback to women about their performance. This effect was replicated in subsequent studies (e.g., Study 6) where people told more white lies to women through qualitative feedback and selecting nicer but less accurate feedback statements. Thus, this component of the overarching proposal was supported.

The findings that women are told more white lies than men also represent significant differences from and important additions to the literature on biased language communication to women. In study 1a, participants first evaluated two essays on objective standards and then, when asked to give feedback to the authors once they had found out their genders, upwardly distorted the original rating for the female target over 11% during direct feedback. Though many researchers have examined language as a vehicle for bias and some have shown that language about women is often "sugar-coated" (e.g., Biernat & Danaher, 2011), the results from Study 1a demonstrate for the first time that *quantitative* evaluation scores can be upwardly distorted as well.

Additionally, the data support the possibility that people may be deliberately bending an objective truth. An alternative explanation for the data in study 1a that needed addressing was that this bias effect is not due to white lies, but that people were merely re-evaluating female-authored work once finding out the gender of the authors ("shifting standards," Biernat & Kobrynowicz, 1997). That is, because women are often evaluated on different standards than men, and are evaluated relative to other women, perhaps people were not lying but instead changing their opinions after finding out the gender of the author. However, in Study 1b, I tested this hypothesis and found that people evaluate badly written essays that have been authored by women to be worse then if men had written the essays. This finding implied that, had the results in Study 1a been due to "shifting standards" for women then participants should have distorted their feedback downwards not upwards. The "shifting standards" theory has been the putative explanation for overly positive language about women in

the workplace, but this theory cannot account for our data, thus this is - as far as I know - the first work to demonstrate that people are actively distorting information to women through white lies.

However, 'actively distorting' does not imply that people are consciously distorting information. Though study 1a and b demonstrated that people were gender-biased when it came to feedback, they did not recognize or did not admit to treating female and male targets any differently. This was despite admitting that they did indeed upwardly distort their feedback overall when having to give it directly to the targets. Nevertheless, when witnessing others give inaccurate feedback to an employee, people assumed that the employee was female (Study 2a). This implies that that not only do people have a "bias blind spot" (Pronin, Lin, & Ross, 2002) but that it is likely that this gendered feedback bias is commonly recognized to exist in the workplace, further corroborating the overarching proposal. It also suggests that people are aware, on some level, that women are treated differently than men during performance feedback, and that they do not want to seem biased.

Why, then, do people tell white lies to women? There are a few theories that could explain why people impart nicer but less accurate feedback to women than to men. The first is a theory that stems from the benevolent sexism literature. That is, because women are seen as warmer but less competent, and are stereotypically perceived as being more emotional, providers of feedback will feel sorry for them and spare their feelings. This empathy will translate into avoidance of the recipient's negative reactions to undesired information. This theory appears to be supported by the data from Study 2b and Study 3 which demonstrated that people describe women

who have been told white lies in more negative emotional terms, and people who hold more sexist views about women are more likely to lie to women. These findings imply that those who believe that women cannot handle the truth are upwardly distorting feedback because they feel (at least on an implicit level) that they should protect women from accurate but negative information.

However, while much of the data lends evidence to the possibility that empathic stereotypes spur people to tell white lies, empathy may not be the only motivating factor. Though people may wish to avoid hurting others' feelings, this does not imply that the action is protective or benevolent. Indeed, people may hold hostile views about women's reactivity and their lying behavior may be more reflective of their intention to avoid negative confrontation or a "scene" than hurting someone's feelings. White lies are complex and often benevolent and selfish behavior is not mutually exclusive (e.g., Dawkins, 2006). It is unclear what the exact motivations are when people choose to tell white lies, and they may differ depending on the context (Gneezy, 2005; Erat & Gneezy, 2012).

Another theory for the motivation to tell white lies may not concern negative emotions at all. In Study 1a, the character traits of female targets were elevated after having told women more white lies. I had predicted that if white lies are told in part because of stereotypical views about women's reactions to negative information then accessing these stereotypes should raise the salience of her more stereotypical feminine traits and degrade the masculine ones (Rudman, 2004; Rudman & Kilianski, 2000; Haines & Sumner, 2006). I thus expected that only the warmth dimension traits would be elevated, and that the competency traits would be lowered. However,

participants who had told white lies to the female target not only elevated her 'warmth' dimension traits, but also her 'competence' dimension traits. This could be evidence of cognitive dissonance; Once having told a lie (normatively bad) participants wish to do something good, like elevate traits, or maintain consistency in their self-perception. However, another explanation is that people are promotion-motivated. That is, people are using white lies in order to promote self-confidence and believe that it will have an effect on actual performance, not to just to avoid hurt feelings. People may do this more for women than for men, especially if they believe that women are disadvantaged (Taylor-Carter, Doverspike, & Cook, 1996).

Though these two theories entail slightly different intentions, they stem from the same overarching stereotype about women and thus may be two parts of the same mechanism: assuming that women cannot handle critical feedback leads people to modify truthful evaluations. The alternative account can can be reconciled with our original hypothesis in that it also stems from benevolent sexism, but it highlights a promotion focused motivation rather than an avoidance one. That is, perhaps people are not avoiding women's negative reactions, but instead intend to give them extra advantage when it comes to excelling (in particular when it is a male-dominated domain, Cross, 2001). In a way, this is akin to an 'affirmative action' mentality toward women. The same people who wish to see women excel may be treating them with "kid-gloves" that can have psychological and material repercussions.

Even though the data appear to corroborate our hypothesis that women are told more white lies about performance than men, is this really problematic? The data from Study 4 suggests that it is. One of the most psychologically impactful problems with

this bias is that people may be mistaken in assuming that women want to receive biased feedback. In Study 4 I found that women in fact did *not* prefer feedback that has been upwardly distorted and instead desired feedback that was more truthful at the same rate as men. They additionally reported that the *least* constructive feedback for them would be white lies about how they were performing. Though these are self-reports and further empirical testing of whether they are right should be done, this finding suggests that women at least recognize that desiring or receiving truthful feedback is probably more constructive.

This data suggest that there is a mismatch between how people are behaving towards women (thus, how they assume women want to be given feedback) and what kinds of feedback women actually want to receive. The fact that women perceive the white lies as the least constructive feedback also implies that perhaps women are more attuned to patronizing situations in the workplace because they experience it more often (see Vescio & Gervais, 2005). However, because this was self-report data, these conclusions should remain conjectures until behavioral data on how women really do react to critical feedback can be gathered.

Unexpected Results and Limitations

Though the main predictions and overarching hypothesis were mostly supported, a few of the studies revealed unexpected or null results that warrant further explanation. One such study was an experiment that attempted to reveal material consequences for receiving inaccurate but nicer feedback. In Study 5, it was expected that feedback type would have an effect on the motivation and on the performance of the participants, and that this effect might be different for men and women.

Specifically I expected that negative feedback would have a more emotional impact on people but that it would ultimately motivate them to do better on subsequent tasks. I additionally predicted that this would be even more the case for women. However, none of these predictions were supported; The only significant differences were in the emotional experience of women, who reported more anxiety than men and the finding that people believe that feedback is both less desired and less important for women than for men.

Because the main predictions, though not significant, appeared to be following the predicted pattern, it is likely that the reasons there were no major differences was that the manipulation was not strong enough and the tasks not engaging enough. In order for people to want to improve, there has to be a certain minimal amount of engagement with the material, and importance placed on both the evaluative criteria and on the diagnosticity of the task (see Sansone, 1998). Additionally, I was trying to capture motivation by giving participants an option to do extra practice round (a time cost), but MTurk may not be the best venue for such experiments, as people may place a lot of value on their time money trade-off and little value on the experimental procedure (Goodman, Cryder, & Cheema, 2012).

Another unexpected finding was that women admitted to reacting worse to critical feedback than men in the same study where they expressed the desire to not receive white lies during feedback (i.e., Study 4). I would have predicted that women who stated their preference for truthful feedback would be similarly reactive to negative information as men. There are a few reasons why the results seem to indicate that women are more reactive than men. First, it could be the case that women do react

worse than men, and that participants are not incorrect in assuming that women will have stronger emotional reactions to their feedback. It does seem to be the case that women respond differently than men to emotional information in general (Barrett et al., 1998), and that women's self-evaluations are more responsive to the valence of the evaluative feedback they receive than are men's (Roberts, 1991).

However, the findings could also be explained by the difference in the types of emotional responses that men and women report having. This is supported by the data in Study 4 that indicated female reactions to critical feedback were more passive and male reactions more aggressive, differences that have been established in the literature and perpetuated in American culture (O'Connor, Archer, & Wu, 2004). Nevertheless, the data may be somewhat difficult to interpret considering the propensity for women to be more comfortable reporting their emotions than men (Belk & Snell, 1986; Grossman & Wood, 1993). Also, because the majority of the studies here used self-report measures of emotions, one of the limitations to these findings is that they may not accurately represent actual emotional intensity or content.

Methodology was a limitation is several of the studies. In Study 6 I had predicted that because giving feedback is characterized by a tension between short-term social goals and long-term performance goals, that changing the regulatory focus of individuals would change their propensity to lie. That is, I hypothesized that when a promotion focus was induced, people would shift their attention to the long-term benefits and away from avoiding hurting people's feelings. I predicted that this would especially be the case for female targets and that this intervention could mitigate the gender feedback bias. However, none of these predictions were supported, despite the

data trending in the predicted direction and showing interesting differences between male and female participants.

This is most likely a case of ineffective methodology or low sample size, as explained in the study discussion. However, an alternative explanation is that for some people there is no tension or competition between goals when giving feedback. This may be the case because people either do not see telling white lies to women as dissonant, or because some people are promotion-focused and some are prevention-focused. To come back to the other side of the benevolent sexism argument, as stated earlier, the same outcome can have different motivations, one a prevention motivation (avoid causing emotional pain) and the other a promotion motivation (promote the self-confidence of women). Indeed, there is some evidence that different people have different motivations stemming from the same beliefs about women's emotionality: Men in Study 6 appeared to be lying more in a promotion focus while women appeared to be lying more in a prevention focus. However, a larger samples size would be needed to adequately re-test this hypothesis.

Theoretical and Practical Implications

Despite the limitations of the studies described, there are several implications of this work. First, the fact that women do not want biased feedback, yet people still tell more white lies to women implies that women may be experiencing patronizing workplaces. This has psychological consequences for women that could translate into material consequences. The biggest problem is unequal access to information. Not only could this lead to a lack of knowledge about one's performance, but could also keep women in the dark while managers know their employees are not doing well.

This exclusion could foster deeper rifts between a manager and female employee whereas a relationship with a male employee that is founded on truth and accuracy may be stronger.

Another implication of this bias is that, if it is detected, it could foster resentment and distrust among recipients of white lies who perceive this behavior as patronizing, even if the communicator has no idea that he or she is acting in a biased way. This could not only lead to anger and demotivation as Vescio and Gervais have shown (2005), but also uncertainty about where one really stands in terms of performance, and as research has demonstrated, ambivalence can often be more crippling than negative feedback (Dweck & Gilliard, 1975).

Finally, despite the lack of abundant support for this implication in the current manuscript, there is a possibility that people's short term goals of prosociality could affect the important material long-term goals of workplace success: if white lies are indeed told to women more often than men about their performance then women could very well be left with a dearth of information needed to improve their performance, especially if, as our data shows, people believe that feedback is less desired and less important for women than for men. This could leave women at a disadvantage compared to their male counterparts and actual performance could suffer.

Future Directions

In terms of future directions, it will be useful to research more deeply the implications of unequal information reception on both women and men's future performance, and understand the actual longitudinal impacts of this discrepancy. One of the biggest issues with running an experiment to test the effects of feedback on

performance is that decrements and improvement in performance can often unfold slowly over time, and this may be particularly the case for women. Roberts and Nolen-Hoeksema (1989), have suggested that in contrast to women, men may be particularly likely to respond to the competitive nature of evaluative achievement and hence to adopt a self-confident approach that leads them to deny the informational value of others' evaluations. This could mean that women actually incorporate critical feedback better and be more likely to use constructive feedback to enhance their performance in the long-run.

Research has been somewhat equivocal on how receivers of feedback and evaluations translate and react to this feedback and the literature on evaluations and feedback is complex and sometimes contradictory. For example, some research has demonstrated that accuracy of self evaluations lead to better performance while over-positive feedback leads to worse (Ellis et al., 2009). Additionally there is also work showing that people, especially women, perform worse on subsequent tasks after destructive criticism (Baron 1988). Roberts (1991) argues that this may in part be due to the differences in the amount of negative feedback young girls are given compared to young boys and that women may be less prepared for negative feedback because they are less inured to it.

What researchers *have* agreed upon, however, is that the interpretation of feedback is somewhat subjective. Descriptive words like “intelligent” and “sincere” do not have a fixed meaning, and depend on who has uttered them, and to whom (Huttenlocher, Higgins, &Clark, 1971) and “360 degree feedback” (constructive critical feedback), for example, tends to motivate people more than does positive

feedback alone, but only if the negative is deemed useful (Brett & Atwater, 2011). However, in order to address real-world impacts of inaccurate feedback on women, both short-term and long-term studies will need to be done in a variety of different contexts utilizing a variety of performance assessments. In addition, studies using person-to-person feedback in realistic situations will provide greater ecological validity to the findings.

The ironic effects of prosocial actions should also be pursued, including investigations of the finding that being kinder to women about their performance leads to giving her inaccurate information. Future research could examine, for example, the aftereffects of prosocially lying on perceptions of the target of the inaccurate feedback at a later time. That is, managers who have continued to tell white lies to female employees may also continue to perceive them as less able to handle difficult situations and this stereotype may affect perceptions of them in other competency domains as well. On the other hand, if managers are lying out of benevolent promotion intentions in order to give women a step up, there may be repercussions if women do not meet these expectations of improvement after perceived help. Future studies would do well to research the effects that white lies have not only on the targets but also on the perceptions of the person telling them.

Studying this bias in the context of other demographics - including other minority groups and marginalized populations -would also be useful for the advancement of existing research addressing stereotyping and discrimination. According to the stereotype content model (Fiske et al., 2002), women are not the only group of people who are targets of empathic biases. Elderly people elicit many of the

same responses from people and may also be affected by prosocial lying in the workplace. It is also unclear whether or not this bias would affect other minorities that do not elicit stereotypical warmth judgments.

Finally more clarification is needed to determine and hone the exact motivation that drives people to tell white lies during feedback, and whether different types of people may have different motivations. Understanding how context and individual differences such as participant gender, personality differences like trait-compassion or empathy, and differences in beliefs contribute to this bias will provide a richer understanding of it so that interventions can be designed to attenuate this bias.

There are also practical implications of the findings detailed here. In terms of policy applications in organizations, one of the easiest ways to address a feedback bias may be to make people more aware of it. There is the possibility that awareness interventions may actually work (e.g., Greenberg, 1983), a topic that has been debated in the social sciences (Tversky & Kahneman, 1986) and future research would do well to examine the possibility of educating managers on their potential for unequal behavior. For example, a policy that encourages managers to remind themselves of the goal to promote performance before giving feedback, or having a third-party give feedback to employees might help reduce differences in the feedback given to men and women. On the other hand, some action could come from women themselves, especially if they know that this bias exists. Women could explicitly ask or remind managers that they would like candid and truthful feedback for their own professional development. Future studies in businesses and other institutions might enlist both parties in correcting the propensity to tell white lies during feedback.

Conclusion

A major obstacle to eliminating benevolent sexism and its negative effects is that people (both the person in power and the target) may be unaware of these effects especially if they seem well-meaning. This is particularly problematic, because sexism that is not thought of as sexist is the most difficult type to eradicate, especially if it is proliferated through cultural routine (Stephens & Levine, 2001). Thus it is essential that schemas of covert, benevolent sexism be identified, their nuances understood, and behaviors changed through intervention. Additionally, understanding subtle sexist communication in the form of a quotidian practice like inaccurate feedback could lead to ways of eradicating this practice, even if they are just small behavioral changes that bring us one step closer to bridging the inequality gap, one interaction at a time. Our research demonstrates one such covert bias: that people withhold negative information from women more than men because of stereotypical beliefs about women's reactions to feedback. While we may not be able to completely shatter the glass ceiling any time soon, time and some well-placed cracks will eventually lead to a deeper understanding of the explicit and implicit barriers to inequality and to ways of eradicating them.

APPENDIX

Ambivalent Sexism Inventory (ASI: Glick & Fiske, 1996).

Below is a series of statements concerning men and women and their relationships in contemporary society. Please indicate the degree to which you agree or disagree with each statement using the following scale:

0 = disagree strongly; 1 = disagree somewhat; 2 = disagree slightly; 3 = agree slightly; 4 = agree somewhat; 5 = agree strongly.

___1. No matter how accomplished be is, a man is not truly complete as a person unless he has the love of a woman.
___2. Many women are actually seeking special favors, such as hiring policies that favor them over men, under the guise of asking for "equality."
___3. In a disaster, women ought not necessarily to be rescued before men.
___4. Most women interpret innocent remarks or acts as being sexist.
___5. Women are too easily offended.
___6. People are often truly happy in life without being romantically involved with a member of the other sex.
___7. Feminists are not seeking for women to have more power than men.
___8. Many women have a quality of purity that few men possess.
___9. Women should be cherished and protected by men.
___10. Most women fail to appreciate fully all that men do for them.
___11. Women seek to gain power by getting control over men.
___12. Every man ought to have a woman whom he adores.
___13. Men are complete without women.
___14. Women exaggerate problems they have at work.
___15. Once a woman gets a man to commit to her, she usually tries to put him on a tight leash.
___16. When women lose to men in a fair competition, they typically complain about being discriminated against.
___17. A good woman should be set on a pedestal by her man.
___18. There are actually very few women who get a kick out of teasing men by seeming sexually available and then refusing male advances.
___19. Women, compared to men, tend to have a superior moral sensibility.
___20. Men should be willing to sacrifice their own well being in order to provide financially for the women in their lives.
___21. Feminists are making entirely reasonable demands of men.
___22. Women, as compared to men, tend to have a more refined sense of culture and good taste.

END OF SURVEY

SCORING: To score the inventory, these numbers should be reversed for the subsequent items: (0 = 5, 1 = 4, 2 = 3, 3 = 2, 4 = 1, 5 = 0) placed in front of items 3, 6, 7, 13, 18, and 21. An overall measure of sexism is found by adding, then taking an average of the numbers placed in front of all 22 items. The hostile sexism score is found by averaging the numbers in response to items 2, 4, 5, 7, 10, 11, 14, 15, 16, 18, and 21. The benevolent sexism score is found by averaging the numbers in response to items 1, 3, 6, 8, 9, 12, 13, 17, 19, 20, and 22.

REFERENCES

Aarts, H., Custers, R., & Veltkamp, M. (2008). Goal Priming and the Affective-Motivational Route to Nonconscious Goal Pursuit. *Social Cognition*, *26*(5), 555–577. doi:10.1521/soco.2008.26.5.555

Achtziger, A., Gollwitzer, P. M., & Sheeran, P. (2008). Implementation intentions and shielding goal striving from unwanted thoughts and feelings. *Personality & Social Psychology Bulletin*, *34*(3), 381–93. doi:10.1177/0146167207311201

Ackert, L. F., Church, B. K., & Kuang, X. J. (2007). White Lies: Why Bother? *Business Ethics Quarterly*, (August).

Allport, G. (1954). *The Nature of Prejudice*. Cambridge, MA: Perseus Books.

Ambady, N., Hallahan, M., & Conner, B. (1999). Accuracy of Judgments of Sexual Orientation from Thin Slices of Behavior. *Journal of Personality and Social Psychology*, *77*(3), 538–547.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*(2), 256–274. doi:10.1037//0033-2909.111.2.256

Approach, A. P., & Cross, E. (2001). Training the Scientists and Engineers of Tomorrow :, 296–323.

Argo, J. J., & Shiv, B. (2012). Are White Lies as Innocuous as We Think? *Journal of Consumer Research*, *38*(6), 1093–1102. doi:10.1086/661640

Ariely, D., & Zakay, D. (2001). A timely account of the role of duration in decision making. *Acta Psychologica*, *108*(2), 187–207. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11569762

Arvey, R. D., & Murphy, K. R. (1998). Performance evaluation in work settings. *Annual Review of Psychology*, *49*, 141–68. doi:10.1146/annurev.psych.49.1.141

Ashcraft, K. L., & Pacanowsky, M. E. (1996). "A Woman's Worst Enemy": Reflections on a Narrative of Organizational Life and Female Identity. *Journal of Applied Communication Research*, *24*, 217–239.

Atwater, L., Roush, P., & Fischthal, A. (1995). THE INFLUENCE OF UPWARD FEEDBACK ON SELF ‐ AND FOLLOWER RATINGS OF LEADERSHIP. *Personnel Psychology*, *48*(1), 35-59.

Baas, M., De Dreu, C. K. W., & Nijstad, B. a. (2011). When prevention promotes creativity: the role of mood, regulatory focus, and regulatory closure. *Journal of Personality and Social Psychology*, *100*(5), 794–809. doi:10.1037/a0022981

Babchuk, W. A., Hames, R. B., & Thompson, R. A. (1985). Sex differences in the recognition of infant facial expressions of emotion: The primary caretaker hypothesis. *Ethology and Sociobiology*, *6*(2), 89-101.

Babcock, L., Laschever, S., Gelfand, M., & Small, D. (2003). Nice Girls Don ' t Ask Nice Girls Don ' t Ask. *Harvard Business Review*, *October*.

Baer, J., & Kaufman, J. C. (2008). Gender differences in creativity. *The Journal of Creative Behavior*, *42*(2), 75-105.

Banks, Stitt, Curtis, H., & McQuarter, G. (1977). Perceived objectivity and the effects of evaluative. reinforcement upon compliance and self-evaluation in blacks. *Journal of Expert Mental Social Psychology*. Retrieved from http://scholar.google.com/scholar?hl=en&q=banks++stitt+curtis+mcquarter+1977&btnG=&as_sdt=1,33#0

Bannister, B. D. (1986). Performance outcome feedback and attributional feedback: Interactive effects on recipient responses. *Journal of Applied Psychology*, *71*(2), 203–210. doi:10.1037//0021-9010.71.2.203

Baron, R. a. (1988). Negative effects of destructive criticism: impact on conflict, self-efficacy, and task performance. *The Journal of Applied Psychology*, *73*(2), 199–207. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3384772

Barreto, M., & Ellemers, N. (2005). The burden of benevolent sexism: how it contributes to the maintenance of gender inequalities. *European Journal of Social Psychology*, *35*(5), 633–642. doi:10.1002/ejsp.270

Barrett, L. F., Robin, L., Pietromonaco, P. R., & Eyssell, K. M. (1998). Are Women the "More Emotional" Sex? Evidence From Emotional Experiences in Social Context. *Cognition & Emotion*, *12*(4), 555–578. doi:10.1080/026999398379565

Batson, C. D. (1998). Prosocial behavior and altruism. *The handbook of social psychology*, *4*, 282-316.

Bauer, C. C., & Baltes, B. B. (2002a). Reducing the Effects of Gender Stereotypes on Performance Evaluations 1, *47*(November).

Bauer, C. C., & Baltes, B. B. (2002b). Reducing the Effects of Gender Stereotypes on Performance Evaluations 1, *47*(November).

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529.

Bell, K. L., & DePaulo, B. M. (1996). Liking and Lying. *Basic and Applied Social Psychology*, *18*(3), 243–266. doi:10.1207/s15324834basp1803_1

Belk, S. S., & Snell, W. E. (1986). Beliefs about Women Components and Correlates. *Personality and Social Psychology Bulletin*, *12*(4), 403-413.

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*(2), 274–303. doi:10.1111/j.1464-0597.2008.00336.x

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological review*, *74*(3), 183.

Benedict, M. E., & Levine, E. L. (1988). Delay and distortion: Tacit influences on performance appraisal effectiveness. *Journal of Applied Psychology*, *73*(3), 507.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology*, *59*(5), 960–970. doi:10.1037//0022-3514.59.5.960

Beyer, S., & Bowden, E. M. (1997). Gender Differences in Seff-Perceptions: Convergent Evidence from Three Measures of Accuracy and Bias. *Personality and Social Psychology Bulletin*, *23*(2), 157–172. doi:10.1177/0146167297232005

Biernat, M., & Danaher, K. (2011). Interpreting and reacting to feedback in stereotype-relevant performance domains. *Journal of Experimental Social Psychology*, *48*(1), 271–276. doi:10.1016/j.jesp.2011.08.010

Biernat, M., & Eidelman, S. (2007). Translating subjective language in letters of recommendation: the case of the sexist professor. *European Journal of Social Psychology*, *37*(6), 1149–1175. doi:10.1002/ejsp.432

Biernat, M., Fuegen, K., & Kobrynowicz, D. (2010). Shifting standards and the inference of incompetence: effects of formal and informal evaluation tools. *Personality & Social Psychology Bulletin*, *36*(7), 855–68. doi:10.1177/0146167210369483

Biernat, M., & Kobrynowicz, D. (1997). Gender- and race-based standards of competence: lower minimum standards but higher ability standards for devalued groups. *Journal of Personality and Social Psychology*, *72*(3), 544–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9120783

Biernat, M., & Manis, M. (2007). Stereotypes and shifting standards: Assimilation and contrast in social judgment. In D. A. Stapel & J. Suls (Eds.), *Assimilation and contrast in social psychology, 75-77,* New York: Psychology Press.

Biernat, M., Tocci, M. J., & Williams, J. C. (2011). The Language of Performance Evaluations: Gender-Based Shifts in Content and Consistency of Judgment. *Social Psychological and Personality Science*, *3*(2), 186–192. doi:10.1177/1948550611415693

Biernat, M., & Vescio, T.K. (2002). She swings, she hits, she's great, she's benched: Implications of gender-based shifting standards for judgment and behavior. *Personality and Social Psychology Bulletin, 28,* 66-77.

Birnbaum, D., & Chemelski, B. (1984). Preschoolers' inferences about gender and emotion: The mediation of emotionality stereotypes. *Sex Roles*, *10*(7-8). doi:10.1007/BF00287259

Blanton, H., George, G., & Crocker, J. (2001). Contexts of System Justification and System Evaluation: Exploring the Social Comparison Strategies of the (Not Yet) Contented Female Worker. *Group Processes & Intergroup Relations*, *4*(2), 126–137. doi:10.1177/1368430201004002004

Bodenhausen, G. V. (1993). Emotions, arousal, and stereotypic judgments: A heuristic model of affect and stereotyping.

Bohnet, I. (2012). When Performance Trumps Gender Bias : Joint Versus Separate Evaluation Citation When Performance Trumps Gender Bias : Joint Versus Separate Evaluation Faculty Research Working Paper Series.

Bornmann, L., Mutz, R., & Daniel, H. (2007). Gender differences in grant peer review: A meta-analysis. *Journal of Informetrics*, *1*(3), 226–238. doi:10.1016/j.joi.2007.03.001

Bors, D. a, & Vigneau, F. (2001). The effect of practice on Raven's Advanced Progressive Matrices. *Learning and Individual Differences*, *13*(4), 291–312. doi:10.1016/S1041-6080(03)00015-3

Boshoff, A. B. (2003). POSSIBLE COGNITIVE PROCESSES UNDERLYING COMPETENCY ASSESSMENT. *Journal of Personality*, *29*(3), 45–52.

Bowers, K.S., Regehr, G., Balthazard, C.G., & parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology*, **22**, 72-110.

Bowden, E.M., & Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behavioral Research Methods, Instrumentation, and Computers*, **35**, 634-639.

Bowles, H. R., Babcock, L., & Lai, L. (2005). It Depends Who Is Asking and Who You Ask : Social Incentives for Sex Differences in the Propensity to Initiate Negotiation Social Incentives for Sex Differences in the Propensity to Initiate Negotiation. *Faculty Research Working Papers Series*, (July).

Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: sex differences in picture processing. *Emotion*, *1*(3), 300.

Brandt, M. J. (2011). Sexism and Gender Inequality Across 57 Societies. *Psychological Science*, (October). doi:10.1177/0956797611420445

Brass, D. J. (2013). Structural Relationships , Job Characteristics , and Worker Satisfaction and Performance, *26*(3), 331–348.

Brett, J. F., & Atwater, L. E. (2001). 360° feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology*, *86*(5), 930–942. doi:10.1037//0021-9010.86.5.930

Brief, A. P., & Motowidlo, S. J. (1986). Prosocial organizational behaviors.*Academy of management review*, *11*(4), 710-725.

Brockner, J., & Elkind, M. (1985). Self-esteem and reactance: Further evidence of attitudinal and motivational consequences. *Journal of Experimental Social Psychology*, *21*(4), 346–361. doi:10.1016/0022-1031(85)90035-6

Brockner, J., & Higgins, E. T. (2001). Regulatory Focus Theory: Implications for the Study of Emotions at Work. *Organizational Behavior and Human Decision Processes*, *86*(1), 35–66. doi:10.1006/obhd.2001.2972

Brummelman, E., Thomaes, S., Orobio de Castro, B., Overbeek, G., & Bushman, B. J. (2014). “That’s not just beautiful--that's incredibly beautiful!”: the adverse impact of inflated praise on children with low self-esteem. *Psychological Science*, *25*(3), 728–35. doi:10.1177/0956797613514251

Bruning, N. S., & Snyder, R. A. (1985). Sex and Position Differences Among Predictors of Perceived Work-Related Competence, *13*.

Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, *79*(4), 474–482. doi:10.1037//0022-0663.79.4.474

Bylsma, W. H., & Major, B. (1992). TWO ROUTES TO ELIMINATING GENDER DIFFERENCES IN PERSONAL ENTITLEMENT: Social Comparisons and Performance Evaluations. *Psychology of Women Quarterly*, *16*(2), 193–200. doi:10.1111/j.1471-6402.1992.tb00249.x

Cane, D. B., & Gotlib, I. H. (1985). Depression and the effects of positive and negative feedback on expectations, evaluations, and performance. *Cognitive Therapy and Research*, *9*(2), 145–160. doi:10.1007/BF01204846

Carney, D. R., Dubois, D., Nichiporuk, N., Brinke, L., & Rucker, D. D. (2001). The Deception Equilibrium : The Powerful Are Better Liars but the Powerless Are Better Lie- Detectors, *2393*.

Carragee, K. M. (1993). A critical evaluation of debates examining the media hegemony thesis. *Western Journal of Communication*, *57*(3), 330–348. doi:10.1080/10570319309374457

Catalyst. (2004). *catalyst-study-reveals-financial-performance-is-higher-for-companies-with-more-women-at-the-top*. Retrieved from http://www.catalyst.org/press-release/2/catalyst-study-reveals-financial-performance-is-higher-for-companies-with-more-women-at-the-top

Catalyst. (2012). *Statistical Overview of Women in the Workplace*. Retrieved from http://www.catalyst.org/publication/219/statistical-overview-of-women-in-the-workplace

Cikara, M., Lee, T. L., Fiske, S. T., & Glick, P. (2009). Ambivalent Sexism at Home and at Work : How Attitudes Toward Women in Relationships Foster Exclusion in the Public Sphere. In J. T. Jost, A. C. Kay, & H. Thorisdottir (Eds.), *Social and Psychological Bases of Ideology and System Justification* (pp. 444–462). New York, New York, USA: Oxford University Press.

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in psychology*, *9*, 287-299.

Clawson, J. G., & Kram, K. E. (n.d.). Managing Cross-Gender Mentoring.

Clawson, J. G., & Kram, K. E. (1984). Managing Cross-Gender Mentoring. *Business Horizons*, *May-June*, 22–32.

Cohen, G. L., Steele, C. M., & Ross, L. D. (1999a). The Mentor's Dilemma: Providing Critical Feedback Across the Racial Divide. *Personality and Social Psychology Bulletin*, *25*(10), 1302–1318. doi:10.1177/0146167299258011

Cohen, G. L., Steele, C. M., & Ross, L. D. (1999b). The Mentor's Dilemma: Providing Critical Feedback Across the Racial Divide. *Personality and Social Psychology Bulletin*, *25*(10), 1302–1318. doi:10.1177/0146167299258011

Cotter, D., Hermsen, J. M., & Vanneman, R. (2011). The end of the gender revolution? Gender role attitudes from 1977 to 2008. *AJS; American Journal of Sociology*, *117*(1), 259–89. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22003521

Cralley, E. L., & Ruscher, J. B. (2005). Lady, Girl, Female, or Woman: Sexism and Cognitive Busyness Predict Use of Gender-Biased Nouns. *Journal of Language and Social Psychology*, *24*(3), 300–314. doi:10.1177/0261927X05278391

Crocker, J., Voelkl, K., Testa, M., & Major, B. (1991). Social stigma: The affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology*, *60*(2).

Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: on the nonconscious operation of behavioral goals. *Journal of personality and social psychology*, *89*(2), 129.

Cwiden, R. L., El, M., Mc, T., & Ilson, A. V. (1984). White Lies in Interpersonal Communication : A Taxonomy and Preliminary Investigation of Social Motivations. *Most*, *48*(Fall), 309–325.

Dardenne, B., Dumont, M., & Bollier, T. (2007). Insidious dangers of benevolent sexism: consequences for women's performance. *Journal of Personality and Social Psychology*, *93*(5), 764–79. doi:10.1037/0022-3514.93.5.764

Deci, E. L., & Ryan, R. M. (2010). *Self‐Determination*. John Wiley & Sons, Inc..

DePaulo, B. M. (1996). By the Grace of Guile: The role of deception in natural history and human affairs-Rue, L.

DePaulo, B. (1985). Telling Ingratiating Lies:: Effects of Target Sex and Target Attractiveness on Verbal and Nonverbal Deceptive Success, *Journal of Personality and Social Psychology*. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0022351407604087

DePaulo, B. (2004). The many faces of lies. *The Social Psychology of Good and Evil*. Retrieved from http://books.google.com/books?hl=en&lr=&id=genMnQO4h5MC&oi=fnd&pg=PA303&dq=psychology+other+oriented+lies&ots=Fa8qBwo_Yl&sig=H6N3KNZqjj_4rJHrEy-JOo0jxtg

DePaulo, B., & Kashy, D. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*. Retrieved from http://smg.media.mit.edu/library/DePauloEtAl.LyingEverydayLife.pdf

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of personality and social psychology*, *74*(1), 63.

DePaulo, B. M., & Bell, K. L. (1996). Truth and investment: Lies are told to those who care. *Journal of Personality and Social Psychology*, *71*(4).

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*(1), 74–118. doi:10.1037/0033-2909.129.1.74

Derr, R., & Wolf, G. (1987). Self-esteem and Reactions to Negative Feedback : Toward Greater Generalizability, *333*, 318–333.

Diekman, a. B., & Eagly, a. H. (2000). Stereotypes as Dynamic Constructs: Women and Men of the Past, Present, and Future. *Personality and Social Psychology Bulletin*, *26*(10), 1171–1188. doi:10.1177/0146167200262001

Diener, E., Sandvik, E., & Larsen, R. J. (1985). Age and Sex Effects for Emotional Intensity, *2*(1981), 542–546.

Dijk, D. V. A. N., & Kluger, A. N. (2011). Task type as a moderator of positive / negative feedback effects on motivation and performance : A regulatory focus perspective, *1105*(July 2010), 1084–1105. doi:10.1002/job

Djamasbi, S., & Loiacono, E. T. (2008). Do men and women use feedback provided by their Decision Support Systems (DSS) differently? *Decision Support Systems*, *44*(4), 854–869. doi:10.1016/j.dss.2007.10.008

Drago, R., & Williams, C. (2010). *The Gender Wage Gap* (pp. 0–3).

Dumont, M. (2010). Be Too Kind to a Woman, She'll Feel Incompetent: Benevolent Sexism Shifts Self-construal and Autobiographical Memories Toward Incompetence. *Sex Roles*. Retrieved from http://www.springerlink.com/index/c240v34407625108.pdf

Dweck, C. S., & Gilliard, D. (1975). Expectancy statements as determinants of reactions to failure: Sex differences in persistence and expectancy change. *Journal of Personality and Social Psychology*, *32*(6), 1077.

Dumont, M., Sarlet, M., & Dardenne, B. (2008). Be Too Kind to a Woman, She'll Feel Incompetent: Benevolent Sexism Shifts Self-construal and Autobiographical

Memories Toward Incompetence. *Sex Roles*, *62*(7-8), 545–553. doi:10.1007/s11199-008-9582-4

Eagly, A. H., & Johnson, B. T. (1990). Gender and leadership style: A meta-analysis. *Psychological Bulletin*, *108*(2), 233–256. doi:10.1037//0033-2909.108.2.233

Ellis, S., Mendel, R., & Aloni-Zohar, M. (2009). The Effect of Accuracy of Performance Evaluation on Learning From Experience: The Moderating Role of After-Event Reviews. *Journal of Applied Social Psychology*, *39*(3), 541–563. doi:10.1111/j.1559-1816.2009.00450.x

Elster, J. (1982). *Belief, bias, and ideology*. *Rationality and relativism*. Oxford: Basil Blackwell. Retrieved from http://scholar.google.com/scholar?q=elster+1982+belief+bias&btnG=&hl=en&as_sdt=1,33#0

Erat, S., & Gneezy, U. (2011). White Lies. *Management Science*, 1–11. doi:10.1287/mnsc.1110.1449

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*,*28*(1), 1-11.

Escarti, A., & Guzman, J. F. (1999). Effects of feedback on self-efficacy, performance, and choice in an athletic task. *Journal of Applied Sport Psychology*, *11*(1), 83–96. doi:10.1080/10413209908402952

Etaugh, C., & Harlow, H. (1975). Behaviors of Male and Female Teachers as RElated to Behaviors and Attitudes of Elementary School Children. *The Journal of Genetic Psychology*, *127*, 163–170.

Fabes, R. A., & Martin, C. L. (1991). Gender and age stereotypes of emotionality. *Personality and social psychology bulletin*, *17*(5), 532-540.

Faul, F., & Erdfelder, E. (1992). GPOWER: A priori, post-hoc, and compromise power analyses for MS-DOS [Computer program]. *Bonn, FRG: Bonn University, Department of Psychology*.

Feldman Barrett, L., & Russell, J. a. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. doi:10.1037//0022-3514.74.4.967

Ferguson, A. J., Ormiston, M. E., & Moon, H. (2010). From approach to inhibition: the influence of power on responses to poor performers. *The Journal of Applied Psychology*, *95*(2), 305–20. doi:10.1037/a0018376

Ferguson, M. J., & Bargh, J. A. (2004). Liking is for doing: the effects of goal pursuit on automatic evaluation. *Journal of personality and social psychology*, *87*(5), 557.

Festinger, L. (1957). A theory of dissonance cognitive. *Evanston: Row*.

Fischer, A. H. (1993). Sex differences in emotionality: Fact or stereotype?. *Feminism & Psychology*, *3*(3), 303-318.

Fishbach, A., Eyal, T., & Finkelstein, S. R. (2010). How Positive and Negative Feedback Motivate Goal Pursuit. *Social and Personality Psychology Compass*, *4*(8), 517–530. doi:10.1111/j.1751-9004.2010.00285.x

Fisher, C. D. (1979). Transmission of positive and negative feedback to subordinates: A laboratory investigation. *Journal of Applied Psychology*, *64*(5), 533–540. doi:10.1037//0021-9010.64.5.533

Fisher, C. D. (1984). *The Transition of Positive and Negative Feedback to Subordinates* (pp. 84–50).

Fiske, S. (2002). Paternalistic and Envious Gender Stereotypes : Testing Predictions From the Stereotype Content Model, *47*(August).

Fiske, S., & Russell, A. M. (2010). Cognitive Processes. In J. F. Dovidio, Miles Hewstone, Peter Glick, & Victoria M. Esses (Eds.), *The Sage Handbook of Prejudice, Stereotyping and Discrimination* (p. 646). Sage. Retrieved from http://books.google.com/books?hl=en&lr=&id=PPkMy-CFZmIC&pgis=1

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. doi:10.1037//0022-3514.82.6.878

Fiske, S. T., Neuberg, S., Beattie, A. N. N. E., & Milberg, S. J. (1987). Category-Based and Attribute-Based Reactions to Others : Some Informational Conditions of Stereotyping and lndividuating Processes. *Journal of Experimental Social Psychology*, *23*, 399–427.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic behavior*, *6*(3), 347-369.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42. doi:10.1257/089533005775196732

Freitas, A. L., Gollwitzer, P., & Trope, Y. (2004). The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. *Journal of Experimental Social Psychology*, *40*(6), 739–752. doi:10.1016/j.jesp.2004.04.003

Freitas, A. L., Salovey, P., & Liberman, N. (2001). Abstract and concrete self-evaluative goals. *Journal of personality and social psychology*, *80*(3), 410.

Fromme, D. K., Mercadal, D. E., & Mercadal, P. L. (1976). The effects of positive and negative feedback and reward on originality. *Journal of Research in Personality*, *10*(2), 237-244.

Fuegen, K., Biernat, M., Haines, E., & Deaux, K. (2004). Mothers and Fathers in the Workplace: How Gender and Parental Status Influence Judgments of Job-Related Competence. *Journal of Social Issues, 60(4),* 737-754. doi:10.1111/j.0022-4537.2004.00383.x

Fujita, F., Diener, E., & Sandvik, E. (1991). Gender differences in negative affect and well-being: the case for emotional intensity. *Journal of personality and social psychology*, *61*(3), 427.

Galef, B. G., & Laland, K. N. (2005). Social Learning in Animals: Empirical Studies and Theoretical Models. *BioScience*, *55*(6), 489. doi:10.1641/0006-3568(2005)055[0489:SLIAES]2.0.CO;2

Gibbs, R. W. (1999). *Intentions in the Experience of Meaning*. Cambridge University Press.

Glass, J. L., Sassler, S., Levitte, Y., & Michelmore, K. M. (2013). What's So Special about STEM? A Comparison of Women's Retention in STEM and Professional Occupations. *Social Forces*, *92*(2), 723–756. doi:10.1093/sf/sot092

Glick, P., Diebold, J., Bailey-Werner, B., & Zhu, L. (1997). The Two Faces of Adam: Ambivalent Sexism and Polarized Attitudes Toward Women. *Personality and Social Psychology Bulletin*, *23*(12), 1323–1334. doi:10.1177/01461672972312009

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory : Differentiating Hostile and Benevolent Sexism. *Social Psychology*, *70*(3), 491–512.

Glick, P., & Fiske, S. T. (2001). An ambivalent alliance. *American Psychologist*, *56*(2), 109–112.

Glick, P., Fiske, S. T., Mladinic, A., Saiz, J. L., & Abrams, D. (2000). Beyond Prejudice as Simple Antipathy. *Journal of Personality and Social Psychology*, *79*(5), 763–775.

Gneezy, U. (2005). Deception : The Role of Consequences. *The American Economic Review*, *95*(1).

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*(3), 213–224. doi:10.1002/bdm.1753

Goodwin, S. a., Gubin, A., Fiske, S. T., & Yzerbyt, V. Y. (2000). Power Can Bias Impression Processes: Stereotyping Subordinates by Default and by Design. *Group Processes & Intergroup Relations*, *3*(3), 227–256. doi:10.1177/1368430200003003001

Gordon, a. K., & Miller, a. G. (2000). Perspective Differences in the Construal of Lies: Is Deception in the Eye of the Beholder? *Personality and Social Psychology Bulletin*, *26*(1), 46–55. doi:10.1177/0146167200261005

Graumann, C. F. (1989). Perspective setting and taking in verbal interaction. *Language processing in social context*, *54*, 95.

Grice, H. P. (1957). Meaning. *The philosophical review*, 377-388.

Grice, H. P. (1969). Utterer's meaning and intention. *The philosophical review*, 147-177.

Grice, H. P. (1970). *Logic and conversation* (pp. 41-58). Harvard Univ.

Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: physiology, self-report, and expressive behavior. *Journal of personality and social psychology*, *64*(6), 970.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being.*Journal of personality and social psychology*, *85*(2), 348.

Grossman, M., & Wood, W. (1993). Sex differences in intensity of emotional experience: a social role interpretation. *Journal of personality and social psychology*, *65*(5), 1010.

Guilford, J. P. (1967). The nature of human intelligence. New York: McGraw-Hill.

Guillemin, J., Holmstrom, L. L., & Garvin, M. (1979). Judging Competence: Letters of Recommendation for Men and Women Faculties. *The School Review*, *87*(2), 157–170.

Guinote, A., & Phillips, A. (2010). Power Can Increase Stereotyping. *Social Psychology*, *41*(1), 3–9. doi:10.1027/1864-9335/a000002

Guinote, A., Reese, G., & Wilkinson, D. T. (2009). A powerful vision: Power affects visual search behavior. *Psicología*, *13*(1), 127-148.

Haines, E., & Jost, J. T. (2000). *Placating the Powerless: Effects of Legitimate and Illegitimate Explanation on Affect, Memory and Stereotyping*. *Research Paper Series*.

Haines, E., & Sumner, K. (2006). Implicit Measurement of Attitudes , Stereotypes , and Self-Concepts in Organizations: Teaching Old Dogmas New Tricks. *Organizational Research Methods*, *9*(4), 536.

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological bulletin*, *85*(4), 845.

Haselton, M., & Buss, D. (2005). Sex, lies, and strategic interference: The psychology of deception between the sexes. *... and Social Psychology ...*. Retrieved from http://psp.sagepub.com/content/31/1/3.short

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. doi:10.3102/003465430298487

Heerwagen, H., Beach, L. E. E. R. O. Y., & Mitchell, T. R. (1985). Dealing with Poor Performance: Supervisor Attributions and the Cost of Responding' JUDITH H, 638–655.

Hess, U., Blairy, S., & Kleck, R. E. (2000). The influence of facial emotion displays, gender, and ethnicity on judgments of dominance and affiliation. *Journal of Nonverbal behavior*, *24*(4), 265-283.

Higgins, E. T. (1997). Beyond pleasure and pain. *American psychologist*, *52*(12), 1280.

Higgins, E. T. (1998). PROMOTION AND PREVENTION : REGULATORY FOCUS AS A MOTIVATIONAL PRINCIPLE, 1–46.

Hillman, L. W., Schwandt, D. R., & Bartz, D. E. (1990). Enhancing staff members' performance through feedback and coaching. *Journal of Management Development*, *9*(3), 20-27.

Hochschild, A. R. (1983). Attending to, codifying and managing feelings: Sex differences in love. *Rethinkjng sex, gender and society, Laurel Richardson and Vera Taylor, eds. Reading, MA: Addison-WesIey*, *134*.

Hoyt, C. L., Aguilar, L., Kaiser, C. R., Blascovich, J., & Lee, K. (2007). The self-protective and undermining effects of attributional ambiguity. *Journal of Experimental Social Psychology*, *43*(6), 884–893. doi:10.1016/j.jesp.2006.10.013

Humphreys, M. S., & Revelle, W. (1984). Psychological Review, *91*(April).

Huttenlocher, J., Higgins, E. T., & Clark, H. H. (1971). Adjectives, comparatives, and syllogisms. *Psychological Review*, *78*(6).

Hyde, J. S. (2005). The gender similarities hypothesis. *The American Psychologist*, *60*(6), 581–92. doi:10.1037/0003-066X.60.6.581

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, *64*(4), 349–371. doi:10.1037//0021-9010.64.4.349

Ilgen, D. R., & Knowlton Jr, W. A. (1980). Performance attributional effects on feedback from superiors. *Organizational Behavior and Human Performance*, *25*(3), 441-456.

Jackman, M. (1994). *The Velvet Glove: Paternalism and Conflict in Gender, Class, and Race Relations.* University of California Press.

Janice, R., & Sarah, L. (1999). Gender-emotion stereotypes are context specific.

Jelley, R. B., & Goffin, R. D. (2001). Can performance-feedback accuracy be improved? Effects of rater priming and rating-scale format on rating accuracy. *The Journal of Applied Psychology*, *86*(1), 134–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11302225

Jost, J. T., & Burgess, D. (2000). Attitudinal Ambivalence and the Conflict between Group and System Justification Motives in Low Status Groups. *Personality and Social Psychology Bulletin*, *26*(3), 293–305. doi:10.1177/0146167200265003

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*(3), 498–509. doi:10.1037/0022-3514.88.3.498

Jost, J. T., Pelham, B. W., & Carvallo, M. R. (2002). Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups.

Journal of Experimental Social Psychology, *38*(6), 586–602. doi:10.1016/S0022-1031(02)00505-X

Kaplar, M. E., & Gordon, A. K. (2004). The enigma of altruistic lying: Perspective differences in what motivates and justifies lie telling within romantic relationships. *Personal Relationships*, *11*(4), 489–507. doi:10.1111/j.1475-6811.2004.00094.x

Karoly, L. A., & Panis, C. W. . (2004). *The 21st century at work: forces shaping the future workforce and workplace in the United States (prepared for the US Department of Labor by the Rand Corporation)*.

Katkin, E. S., & Hoffman, L. S. (1976). Sex differences and self-report of fear: A psychophysiological assessment. *Journal of Abnormal Psychology*, *85*(6), 607.

Kaufman, J. C. (2006a). Gender Differences in Creativity, 1–38.

Kaufman, J. C. (2006b). Self-reported differences in creativity by ethnicity and gender. *Applied Cognitive Psychology*, *20*(8), 1065–1082. doi:10.1002/acp.1255

Kelly, M. M., Tyrka, A. R., Anderson, G. M., Price, L. H., & Carpenter, L. L. (2008). Sex differences in emotional and physiological responses to the Trier Social Stress Test. *Journal of Behavior Therapy and Experimental Psychiatry*, *39*(1), 87-98.

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review*, *110*(2), 265–284. doi:10.1037/0033-295X.110.2.265

Kilianski, S. E., & Rudman, L. A. (1998). Wanting It Both Ways : Do Women Approve of Benevolent Sexism ? ^. *Sex Roles*, *39*(February).

Kinicki, A. J., Prussia, G. E., Wu, B. J., & McKee-Ryan, F. M. (2004). A covariance structure analysis of employees' response to performance feedback. *The Journal of Applied Psychology*, *89*(6), 1057–69. doi:10.1037/0021-9010.89.6.1057

Kite, M. E., Deaux, K., & Haines, E. L. (2008). Gender Stereotypes. In F. L. Denmark & M. Paludi (Eds.), *Psychology of Women: Handbook of Issues and Theories* (2nd ed., pp. 205–236). Westport, CT: Preager Publishers.

Kite, M.E., and Whitley, B.E.; *The Psychology of Prejudice and Discrimination* .Wadsworth, Cengage Learning, 2010

Klonis, S. C., Plant, E. A., & Devine, P. G. (2005). Internal and external motivation to respond without sexism. *Personality & Social Psychology Bulletin*, *31*(9), 1237–49. doi:10.1177/0146167205275304

Kluegel, J. R., & Smith, E. R. (1986). *Beliefs About Inequality: Americans' Views of What Is and What Ought to Be* (Vol. 4, p. 332). Aldine. Retrieved from http://books.google.com/books?hl=en&lr=&id=rdzmjrv74w8C&pgis=1

Kluger, A. N., & Denisi, A. (1996a). The Effects of Feedback Interventions on Performance : A Historical Review , a Meta-Analysis , and a Preliminary Feedback Intervention Theory, *II*(2), 254–284.

Kluger, A. N., & Denisi, A. (1996b). The Effects of Feedback Interventions on Performance : A Historical Review , a Meta-Analysis , and a Preliminary Feedback Intervention Theory, *II*(2), 254–284.

Kobrynowicz, D., & Biernat, M. (1997). Decoding Subjective Evaluations: How Stereotypes Provide Shifting Standards. *Journal of Experimental Social Psychology*, *33*(6), 579–601. doi:10.1006/jesp.1997.1338

Krauss, R. M., & Chiu, C. Y. (1998). Language and social behavior. *The handbook of social psychology*, *2*, 41-88.

Krauss, R. M., Hall, S., Fussell, S. R., & Hall, M. (n.d.). S p m i c, *39762*(601).

Labouvie-Vief, G., Lumley, M. A., Jain, E., & Heinze, H. (2003). Age and gender differences in cardiac reactivity and subjective emotion responses to emotional autobiographical memories. *Emotion*, *3*(2), 115.

Lane, R. E. (1962). *Political ideology: why the American common man believes what he does.* Oxford: Free Press of Glencoe.

Larson Jr, J. R. (1986). Supervisors' performance feedback to subordinates: The impact of subordinate performance valence and outcome dependence. *Organizational Behavior and Human Decision Processes*, *37*(3), 391-408.

Lee, a Y., Aaker, J. L., & Gardner, W. L. (2000). The pleasures and pains of distinct self-construals: the role of interdependence in regulatory focus. *Journal of Personality and Social Psychology*, *78*(6), 1122–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10870913

Lerner, M. (1980). *The belief in a just world: A fundamental delusion* (p. 209). New York: Plenum Press. Retrieved from http://www.getcited.org/pub/102063124

Lerner, M. J., & Simmons, C. H. (1966). Observer's reaction to the" innocent victim": Compassion or rejection? *Journal of Personality and Social Psychology, 4*(2), 203–210. Retrieved from http://psycnet.apa.org/journals/psp/4/2/203/

Levy, B. R. (2001). Eradication of ageism requires addressing the enemy within. *The Gerontologist*, *41*(5), 578–9; discussion 580. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11574700

Levy, P. E., & Williams, J. R. (2004). The Social Context of Performance Appraisal: A Review and Framework for the Future. *Journal of Management*, *30*(6), 881–905. doi:10.1016/j.jm.2004.06.005

Lindeman, M., Harakka, T., & Keltikangas-Jarvinen, L. (1997). Age and Gender Differences in Adolescents ' Reactions to Conflict Situations: Aggression, Prosociality, and Withdrawl. *Journal of Youth and Adolescence*, *26*(2), 339–351.

Linell, P. (1998). *Approaching dialogue: Talk, interaction and contexts in dialogical perspectives* (Vol. 3). John Benjamins Publishing.

Lippa, R., & Beauvais, C. (1983). Gender jeopardy: The effects of gender, assessed femininity and masculinity, and false success/failure feedback on performance in an experimental quiz game. *Journal of Personality and Social Psychology*, *44*(2), 344–353. doi:10.1037//0022-3514.44.2.344

Lippitt, R., & Gold, M. (1959). Classroom Social Structure As a Mental Health Problem. *Journal of Social Issues*, *15*(1), 40–49. doi:10.1111/j.1540-4560.1959.tb01434.x

Locke, E. A., & Rouge, B. (1989). Effects of Feedback Sign and Credibility Task Performance.

Lucas, R. E., & Gohm, C. L. (2000). Age and sex differences in subjective well-being across cultures. *Culture and subjective well-being*, *3*, 291-317.

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: thinking categorically about others. *Annual Review of Psychology*, *51*, 93–120. doi:10.1146/annurev.psych.51.1.93

Major, B. (1994). From social inequality to personal entitlement: The role of social comparisons, legitimacy appraisals, and group membership. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 26, p. 293). San Diego: Academic Press.

Major, B., Spencer, S., Schmader, T., Wolfe, C., & Crocker, J. (1998). Coping with Negative Stereotypes about Intellectual Performance: The Role of Psychological

Disengagement. *Personality and Social Psychology Bulletin*, *24*(1), 34–50. doi:10.1177/0146167298241003

Malloch, K. (2001). White Lies of Leadership: Careing Dishonesty? *Nursing Administration*, *25*(3).

McCabe, A. E., & Brannon, L. A. (2004). An examination of racial subtypes versus subgroups. *Current Research in Social Psychology*, *9*(8), 109-23.

Mccornack, S., & Levine, T. (1990). When lies are uncovered: emotional and relational outcomes of discovered deception. *Communication Mo*, *57*.

McGraw-Hill Staff. (2000). *Writer's Choice Writing Assessment and Evaluation Rubrics Grade 12*. McGraw-Hill Education.

McRae, K., Ochsner, K. N., Mauss, I. B., Gabrieli, J. J. D., & Gross, J. J. (2008). Gender Differences in Emotion Regulation: An fMRI Study of Cognitive Reappraisal. *Group Processes & Intergroup Relations*, *11*(2), 143–162. doi:10.1177/1368430207088035

Milkman, Katherine L. and Akinola, Modupe and Chugh, Dolly, What Happens Before? A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway into Organizations (April 23, 2014). Available at SSRN: http://ssrn.com/abstract=2063742 or http://dx.doi.org/10.2139/ssrn.2063742

Mischel, W., Ebbesen, E. B., & Zeiss, A. R. (1973). Selective attention to the self: Situational and dispositional determinants. *Journal of Personality and Social Psychology*, *27*(1), 129.

Moskowitz, G. B., Gollwitzer, P. M., & Wasel, W. (2008). Preconscious Control of Stereotype Activation Through Chronic Egalitarian Goals, *77*(I), 167–184.

Moskowitz, G. B., Li, P., & Kirk, E. R. (2004). The Implicit Volition Model: On the Preconscious Regulation of Temporarily Adopted Goals.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474-16479.

Moya, M., Glick, P., Expósito, F., de Lemus, S., & Hart, J. (2007). It's for your own good: benevolent sexism and women's reactions to protectively justified restrictions. *Personality & Social Psychology Bulletin*, *33*(10), 1421–34. doi:10.1177/0146167207304790

Mullin, B.-A., & Hogg, M. A. (1999). Motivations for Group Membership: The Role of Subjective Importance and Uncertainty Reduction. *Basic and Applied Social Psychology*, *21*(2), 91–102. doi:10.1207/S15324834BA210202

Nathan, B. R., & Alexander, R. a. (1985). The role of inferential accuracy in performance rating. *Academy of Management Review. Academy of Management*, *10*(1), 109–15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10269938

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (1996). Lying Words : Predicting Deception From Linguistic Styles, (1901). doi:10.1177/0146167203251529

Nguyen, H.-H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *The Journal of Applied Psychology*, *93*(6), 1314–34. doi:10.1037/a0012702

O'Connor, D. B., Archer, J., & Wu, F. C. (2004). Effects of testosterone on mood, aggression, and sexual behavior in young men: a double-blind, placebo-controlled, cross-over study. *Journal of Clinical Endocrinology & Metabolism*,*89*(6), 2837-2845.

Pearce, J. L., & Porter, L. W. (1986). Employee responses to formal performance appraisal feedback. *Journal of Applied Psychology*, *71*(2), 211–218. doi:10.1037//0021-9010.71.2.211

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates*, *71*, 2001.

Peters, L. H., O'Connor, E. J., Weekley, J., Pooyan, A., Frank, B., & Erenkrantz, B. (1984). Sex bias and managerial evaluations: A replication and extension. *Journal of Applied Psychology*, *69*(2).

Plant, E. A., Hyde, J. S., Keltner, D., & Devine, P. G. (2000). The Gender Stereotyping of Emotions. *Psychology of Women Quarterly*, *24*, 81–92.

Pratto, F., & Shih, M. (2000). Social dominance orientation and group context in implicit group prejudice. *Psychological Science*, *11*(6), 515–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11202500

Pratto, F., & Walker, A. (2001). Dominance in disguise: Power, beneficence, and exploitation in personal relationships. In A. Y. Lee-Chai & J. A. Bargh (Eds.), *The use and abuse of power: Multiple perspectives on the causes of corruption* (pp. 93–114). New York: US Psychology Press.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: divergent perceptions of bias in self versus others. *Psychological review*, *111*(3), 781.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The Bias Blind Spot: Perceptions of Bias in Self Versus Others. *Personality and Social Psychology Bulletin*, *28*(3), 369–381. doi:10.1177/0146167202286008

Pulakos, E. D., & Wexley, K. N. (1983). Perceptual Similarity, Sex, and Performance Ratings in Manager-subordinate Dyads. *The Academy of Management Review*, *26*(1), 129–139.

Purdie-Vaughns, V., Steele, C. M., Davies, P. G., Ditlmann, R., & Crosby, J. R. (2008). Social identity contingencies: how diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology*, *94*(4), 615–30. doi:10.1037/0022-3514.94.4.615

Reilly, R. R., Smither, J. W., & Vasilopoulos, N. L. (1996). A longitudinal study of upward feedback. *Personnel Psychology*, *49*(3), 599-612.

Report, T., Been, H. A. S., For, C., Release, P., Restrictions, N. O., Imposed, A. R. E., … Unlimited, D. (n.d.). ervices Technical Information Agency.

Reyes, D. L., & John, J. (1988). Racism in Academia : The Old Wolf Revisited. *Harvard Educational Review*, *58*(3).

Ridgeway, C. L. (2001). Gender, Status, and Leadership. *Journal of Social Issues*, *57*(4), 637–655. doi:10.1111/0022-4537.00233

Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the Gender System: A Theoretical Perspective on Gender Beliefs and Social Relations. *Gender & Society*, *18*(4), 510–531. doi:10.1177/0891243204265269

Robbins, T. (1993). Moderators of sex bias in the performance appraisal process: A cognitive analysis. *Journal of Management*, *19*(1), 113–126. doi:10.1016/0149-2063(93)90048-R

Roberts, T. a. (1991). Gender and the influence of evaluations on self-assessments in achievement settings. *Psychological Bulletin*, *109*(2), 297–308. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2034753

Roberts, T., & Nolen-hoeksema, S. (1989). Sex Differences in Reactions to Evaluative Feedback. *Sex Roles*, *21*(11).

Roberts, T.-A., & Nolen-Hoeksema, S. (1994). Gender Comparisons in Responsiveness to Others' Evaluations in Achievement Settings. *Psychology of Women Quarterly*, *18*, 221–240.

Rowe, G., Hirsh, J. B., & Anderson, a K. (2007). Positive affect increases the breadth of attentional selection. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(1), 383–8. doi:10.1073/pnas.0605198104

Rudman, L. a, & Heppen, J. B. (2003). Implicit romantic fantasies and women's interest in personal power: a glass slipper effect? *Personality & Social Psychology Bulletin*, *29*(11), 1357–70. doi:10.1177/0146167203256906

Rudman, L. a. (2004). Sources of Implicit Attitudes. *Current Directions in Psychological Science*, *13*(2), 79–82. doi:10.1111/j.0963-7214.2004.00279.x

Rudman, L. a., & Kilianski, S. E. (2000). Implicit and Explicit Attitudes Toward Female Authority. *Personality and Social Psychology Bulletin*, *26*(11), 1315–1328. doi:10.1177/0146167200263001

Sansone, C., Krauss, B., Michela, J., Sandelands, L., Hill, C. A., Manderlink, G., & Rhodewalt, F. (1986). A Question of Competence : The Effects of Competence and Task Feedback on Intrinsic Interest, *31*(5), 918–931.

Sassenberg, K., Jonas, K. J., Shah, J. Y., & Brazy, P. C. (2007). Why some groups just feel better: the regulatory fit of group power. *Journal of Personality and Social Psychology*, *92*(2), 249–67. doi:10.1037/0022-3514.92.2.249

Seidlitz, L., & Diener, E. (1998). Sex differences in the recall of affective experiences. *Journal of personality and social psychology*, *74*(1), 262.

Scott, A., Davis, C., & Mark, A. (2001). Implicit attitudes and racism : Effects of word familiarity and frequency on the implicit associat ...

Scullen, S. E., Mount, M. K., & Goff, M. (2000). Understanding the latent structure of job performance ratings. *The Journal of Applied Psychology*, *85*(6), 956–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11125659

Seibt, B., & Förster, J. (2004). Stereotype threat and performance: how self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology*, *87*(1), 38–56. doi:10.1037/0022-3514.87.1.38

Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: How accessible alternatives affect goal pursuit. *Journal of Experimental Social Psychology*, *38*(4), 368–383. doi:10.1016/S0022-1031(02)00005-7

Sibley, C. G., Overall, N. C., & Duckitt, J. (2007). When Women Become More Hostilely Sexist toward their Gender: The System-Justifying Effect of Benevolent Sexism. *Sex Roles*, *57*(9-10), 743–754. doi:10.1007/s11199-007-9306-1

Sidanius, J., & Bobo, L. (1996). INTERPERSONAL RELATIONS AND GROUP Racism , Conservatism , Affirmative Action , and Intellectual Sophistication : A Matter of Principled Conservatism or Group Dominance ?, *70*(3), 476–490.

Sio, U. N., & Rudowicz, E. (2007). The Role of an Incubation Period in Creative Problem Solving. *Creativity Research Journal*, *19*(2-3), 307–318. doi:10.1080/10400410701397453

Smith, T. W., Ingram, R. E., & Brehm, S. S. (1983). Social anxiety, anxious self-preoccupation, and recall of self-relevant information. *Journal of Personality and Social Psychology*, *44*(6), 1276.

Smith, P. C., & Kendall, L. M. (1963). RETRANSLATION OF EXPECTATIONS : AN APPROACH TO THE CONSTRUCTION OF UNAMBIGUOUS, *47*(2), 149–155.

Spears, R., Ellemers, N., & Doosje, B. (2005). Let me count the ways in which I respect thee: does competence compensate or compromise lack of liking from the group? *European Journal of Social Psychology*, *35*(2), 263–279. doi:10.1002/ejsp.248

Spiegel, S., Grant-pillow, H., & Higgins, E. T. (2004). How regulatory fit enhances motivational strength during goal pursuit, *54*(November 2002), 39–54.

Staats, B. R. (2011). Driven by Social Comparisons : How Feedback about Coworkers ' Effort Influences Individual Productivity.

Statistics, B. of L. (2011). *HOUSEHOLD DATA 11 . Employed persons by detailed occupation , sex , race , and Hispanic or Latino ethnicity [In thousands]* (pp. 1–10).

Steele, C. M. (1997). A Threat in the Air. *American Psychologist*, *52*(6), 613–629.

Steelman, L. a., & Rutkowski, K. a. (2004). Moderators of employee reactions to negative feedback. *Journal of Managerial Psychology*, *19*(1), 6–18. doi:10.1108/02683940410520637

Stephens, N. M., & Levine, C. S. (2011). Opting Out or Denying Discrimination? How the Framework of Free Choice in American Society Influences Perceptions of Gender Inequality. *Psychological Science*. doi:10.1177/0956797611417260

Stevens, L. E., & Fiske, S. T. (1995). Motivation and Cognition in Social Life: A Social Survival Perspective. *Social Cognition*, *13*(3), 189–214. doi:10.1521/soco.1995.13.3.189

Stroessner, S. J., Haines, E. L., Sherman, J. W., & Kantrowitz, C. J. (2010). Stereotype Relevance Moderates Category Activation: Evidence From the Indirect Category Accessibility Task (ICAT). *Social Psychological and Personality Science*, *1*(4), 335–343. doi:10.1177/1948550610372144

Swim, J., Borgida, E., Maruyama, G., & Myers, D. G. (1989). Joan McKay versus John McKay: Do gender stereotypes bias evaluations? *Psychological Bulletin*, *105*(3).

Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. a. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology*, *68*(2), 199–214. doi:10.1037//0022-3514.68.2.199

Swim, J. K., & Cohen, L. L. (1997). Overt, Covert, and Subtle Sexism. *Psychology of Women Quarterly*, *21*, 103–118.

Szrek, H., & Bundorf, M. K. (2011). Age and the purchase of prescription drug insurance by older adults. *Psychology and aging*, *26*(2), 308.

Tajfel, H. (1981). *Human Groups and Social Categories: Studies in Social Psychology* (p. 369). CUP Archive. Retrieved from http://books.google.com/books?hl=en&lr=&id=ldA8AAAAIAAJ&pgis=1

Taylor, S. E., & Falcone, H.-T. (1982). Cognitive Bases of Stereotyping: The Relationship between Categorization and Prejudice. *Personality and Social Psychology Bulletin*, *8*(3), 426–432. doi:10.1177/0146167282083006

Taylor-Carter, M. A., Doverspike, D., & Cook, K. D. (1996). The effects of affirmative action on the female beneficiary. *Human Resource Development Quarterly*, *7*(1), 31–54. doi:10.1002/hrdq.3920070106

Tesser, A., & Rosen, S. (1975). The reluctance to transmit bad news. *Advances in experimental social psychology*, *8*, 193-232.

Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, *78*(5), 853–870. doi:10.1037//0022-3514.78.5.853

Thau, S., Tröster, C., Aquino, K., Pillutla, M., & Cremer, D. (2012). Satisfying Individual Desires or Moral Standards? Preferential Treatment and Group

Members' Self-Worth, Affect, and Behavior. *Journal of Business Ethics*, *113*(1), 133–145. doi:10.1007/s10551-012-1287-5

Thomson, R., & Murachver, T. (2001). Predicting gender from electronic discourse. *The British Journal of Social Psychology / the British Psychological Society*, *40*(Pt 2), 193–208. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11446227

Tice, D. M., Bratslavsky, E., & Baumeister, R. F. (2001). Emotional distress regulation takes precedence over impulse control: If you feel bad, do it! *Journal of Personality and Social Psychology*, *80*(1), 53–67. doi:10.1037//0022-3514.80.1.53

Timmers, M., Fischer, A. H., & Manstead, A. S. R. (2003). Ability versus vulnerability: Beliefs about men's and women's emotional behaviour. *COGNITION AND EMOTION*, *17*(1), 41–63.

Topping, K. J., Samuels, J., & Paul, T. (2007). Does practice make perfect? Independent reading quantity, quality and student achievement. *Learning and Instruction*, *17*(3), 253–264. doi:10.1016/j.learninstruc.2007.02.002

Trope, Y., & Neter, E. (1994). Reconciling competing motives in self-evaluation: the role of self-control in feedback seeking. *Journal of personality and social psychology*, *66*(4), 646.

Turner, R., Edgley, C., & Olmstead, G. (1975). Information control in conversations: Honesty is not always the best policy. *Kansas Journal of Sociology*, *11*(1). Retrieved from http://psycnet.apa.org/psycinfo/1976-26015-001

Tyler, J. M., & Feldman, R. S. (2004). Truth, Lies, and Self-Presentation: How Gender and Anticipated Future Interaction Relate to Deceptive Behavior1. *Journal of Applied Social Psychology*, *34*(12), 2602–2615. doi:10.1111/j.1559-1816.2004.tb01994.x

Tyler, T. R., & McGraw, K. M. (1986). Ideology and the Interpretation of Personal Experience: Procedural Justice and Political Quiescence. *Journal of Social Issues*, *42*(2), 115–128. doi:10.1111/j.1540-4560.1986.tb00228.x

Vescio, T., Gervais, S., Heiphetz, L., & Bloodhart, B. (2009). The stereotypic behaviors of the powerful and their effect on the relatively powerless. In T. D. Nelson (Ed.), *Handbook of Prejudice, Stereotyping, and Discrimination (Google eBook)* (p. 581). CRC Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=J1qeQT4bkX0C&pgis=1

Vescio, T. K., Butz, D. A., & Snyder, M. (2003). Power in Stereotypically Masculine Domains: A Social Influence Strategy × Stereotype Match Model. *Journal of Personality and Social Psychology*, *85*(6), 1078–1062.

Vescio, T. K., Gervais, S. J., Heidenreich, S., & Snyder, M. (2006). The effects of prejudice level and social influence strategy on powerful people's responding to racial out-group members. *European Journal of Social Psychology*, *36*(4), 435–450. doi:10.1002/ejsp.344

Vescio, T. K., Gervais, S. J., Snyder, M., & Hoover, A. (2005). Power and the creation of patronizing environments: the stereotype-based behaviors of the powerful and their effects on female performance in masculine domains. *Journal of Personality and Social Psychology*, *88*(4), 658–72. doi:10.1037/0022-3514.88.4.658

Viki, G. T., Abrams, D., & Hutchison, P. (2003). The " True " Romantic : Benevolent Sexism and Paternalistic Chivalry. *Sex Roles*, *49*(November).

Vrana, S. R., & Rollock, D. (2002). The role of ethnicity, gender, emotional content, and contextual differences in physiological, expressive, and self-reported emotional responses to imagery. *Cognition & Emotion*, *16*(1), 165-192.

Whitty, M., & Carville, S. (2008). Would I lie to you? Self-serving lies and other-oriented lies told across different media. *Computers in Human Behavior*. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0747563207000775

Wilbert, J. (2010). Effects of Evaluative Feedback on Rate of Learning and Task Motivation: An Analogue Experiment, *8*(2), 43–52.

Wilder, D. A. (1993). The role of anxiety in facilitating stereotypic judgments of outgroup behavior. *Affect, cognition, and stereotyping: Interactive processes in group perception*, 87-109. Forsythe et al., 1994

Williams, J. C., Manvell, J., & Bornstein, S. (2006). " Opt Out" or Pushed Out?: How the Press Covers Work/Family Conflict. *The Center for WorkLife Law*.

Willness, C. R., Steel, P., & Lee, K. (2007). a Meta-Analysis of the Antecedents and Consequences of Workplace Sexual Harassment. *Personnel Psychology*, *60*(1), 127–162. doi:10.1111/j.1744-6570.2007.00067.x

Wold, A., & Wennerås, C. (1997). Nepotism and sexism in peer review. *Nature*,*387*(6631), 341-343