

Balancing Professional Prototypes Increases the Valuation of Women in Male-Dominated
Professions

Felix Danbold

Leonard N. Stern School of Business

New York University

Corinne Bendersky

Anderson School of Management

University of California, Los Angeles

Keywords:

Diversity in Organizations, Group Processes and Performance, Psychological Processes

Author Contributions:

Both authors contributed equally to the study design, data collection, analyses, writing, and revision of this manuscript.

Acknowledgments:

Thanks to the department administration and local union leaders of the Los Angeles County, Culver City, and Santa Monica Fire Departments.

ABSTRACT

We tackle the persistent problem of people from specific demographic groups (e.g., women) being undervalued in professional contexts where traits associated with their group do not align with the traits perceived to be essential for success (the professional prototype). We introduce the concept of “balancing” professional prototypes, such that group membership becomes irrelevant to determining an individual’s prototypicality. Using a novel technique called “prototype inversion,” we emphasize the importance of professional traits that are typically associated with an underrepresented group, without dismissing those that are associated with the currently prototypical group. By balancing the prototype in this way, it becomes easier to recognize the professional potential of members of underrepresented groups, without incurring backlash from the currently prototypical group. We conducted a full-cycle research project to demonstrate the effectiveness of this strategy in the extreme context of women in firefighting, using qualitative and quantitative methods and participants from both the lab and the field.

Think of a firefighter. Like most people, the person that came to your mind was probably a man.¹ This may not be surprising, as the fire service today is over 95% men (Haynes & Stein, 2014; National Fire Protection Association, 2013). Now think of a business leader. Or a scientist. Or a computer programmer. Chances are you first thought about men for these professions as well (Chambers, 1983; Cheryan, Plaut, Davies, & Steele, 2009; Koenig, Eagly, Mitchell, & Risitkari, 2011). The strong association between being a man and being a firefighter, a scientist, or a business leader, makes it difficult for people to picture a woman succeeding in these professions (McClellan, Martin, Emich, & Woodruff, 2017). This is because the set of traits perceived to be essential for success (the professional prototype) is better aligned with the set of traits that are stereotypically associated with men than the set of traits that are stereotypically associated with women (Eagly & Karau, 2002; Heilman, 1983). As a result, women in these professions often face chronic skepticism and undervaluation. This skepticism inevitably leads to further underrepresentation, perpetuating income disparities (Bureau of Labor Statistics, 2015; Lips, 2013) and undermining women's potential contributions to those professions (Ely & Meyerson, 2010; Galinsky, et al., 2015, Herring, 2009).

In this paper, we build upon a large body of scholarship that examines how the fit (or lack thereof) between a category² (e.g., firefighter) and a group (e.g., women) impacts relative group success within that category. By theorizing about the content and malleability of category prototypes,³ we argue that because prototypes are generally weighted in favor of historically-

¹ Although they are meaningfully distinct, in this paper we use sex (male, female) and gender (man, woman) terminology interchangeably. We acknowledge that the experience of non-cisgender individuals for whom sex and gender identity do not align may not be fully captured in this paper.

² Consistent with the language of self-categorization theory (e.g., Mummendey & Wenzel, 1999; Turner, 1987), we use the term "group" to refer to a classification of people based on a shared characteristic like gender. We use the term "category" to refer to a superordinate and more inclusive classification (e.g., a profession), such that a category contains multiple groups nested inside.

³ We define a prototype as the constellation of traits that make up what a category is or should be like. Although the terms prototype and schema are often used interchangeably (Smith, 1998) we adhere to the usage of prototype as the more specific term. We also distinguish prototypes from stereotypes. A prototype

dominant groups (e.g., men), the disadvantage this causes non-dominant, or underrepresented groups (e.g., women) can be reduced by *balancing* these prototypes. Within a balanced prototype, the legitimate category traits that align with stereotypes about the non-dominant group are considered equally important as the category traits that align with stereotypes about the dominant group. For example, the professional prototype of business leaders has traditionally prioritized stereotypically masculine traits (e.g., assertiveness) over legitimately important, but stereotypically feminine traits (e.g., emotional intelligence), making it difficult for people to picture women in leadership roles. A balanced version of the business leader prototype would give emotional intelligence (and other stereotypically feminine traits) equal importance as assertiveness (and other stereotypically masculine traits), such that gender would no longer be predictive of success as a business leader, and both men and women could be seen as having equal potential in the profession.

Although many have tried to develop strategies for reducing group-based biases (e.g., the undervaluation of women in male-dominated professions), prototype balancing is a distinct approach to redefining category prototypes. For example, one approach has been to replace a prototype that advantages one group with one that advantages a different group (e.g., making a masculine prototype feminine instead; Helgesen, 1990; Rosener, 1995). Another approach has been to make the category prototype so unclear that no one group can claim to be prototypical (e.g., Alexandre, Waldzus, & Wenzel, 2016; Waldzus, Mummendey, Wenzel, & Weber, 2003). In contrast, prototype balancing articulates a clear, inclusive prototype for which fit assessments are based on the constellation of traits a person has, rather than the group to which they belong. This offers a new way of addressing a long-standing problem for organizations struggling with

is a set of traits associated with category membership, whereas stereotypes are individual traits that positively, negatively, or neutrally color the object, identity, or category which they're associated with (Smith & Semin, 2007).

diversity, and advances theory about the malleability of prototype content. Empirically, we focus on gender and generate balanced prototypes that subsequently improve attitudes towards women in male-dominated professions.

To test our predictions, we conducted a full-cycle research project and present multi-method evidence of the effects of prototype balancing from both the field and the laboratory. We test our theory in the context of the fire service, a profession that has an extremely masculine prototype and is overwhelmingly male-dominated. To explore if and how one could legitimately balance this masculine professional prototype, we first engaged in an inductive qualitative field study in several fire departments. From this study, we developed specific hypotheses that we test in deductive experiments with a field sample of professional firefighters and a general population sample (for replication and enhanced internal validity). From this “extreme case” (Eisenhardt, 1989) of firefighters, we suggest general strategies for balancing professional prototypes to increase their alignment with more demographic groups. In doing so, we introduce a powerful new tool for reducing the devaluation demographic groups from professions in which they are underrepresented.

The Causes and Consequences of Misalignment Between a Group and a Category Prototype

Decades of cognitive science scholarship has shown that people develop category prototypes as they learn about what characteristics both unite members of a category and distinguish them from members of other categories (Oakes, Haslam, & Turner, 1994; Rosch, 1975; Turner, 1987). Once formed, these prototypes are used to judge how well different members of the category represent the category as a whole. For example, in the category of scientist, the prototype may be defined by innate intelligence and analytical skills. These traits not only define what is essential to be considered a scientist (i.e., the traits that all scientists share) but also clearly distinguish this category from other categories. Individuals within the category of

scientist are then judged against this prototype: Those who are seen as best aligning with the prototype are rewarded and regarded as true scientists, whereas those who do not are viewed as deviating from norms and marginalized. This issue is often talked about in broad terms as person-situation fit (Cable & DeRue, 2002; Muchinsky & Monahan, 1987, Kristof 1996; Lewin, 1935).

Just as an individual can struggle to succeed when they do not fit in well with a particular context, entire demographic groups can face disadvantage when the traits they are believed to share do not align with the expectations of their environment (Hall, Galinsky, & Phillips, 2015; Kristof-Brown, Zimmerman, & Johnson, 2005; Sy, et al., 2010). For example, within the category of scientists, there are both men and women. However, when considering the stereotypes generally associated with these groups (e.g., men are stereotypically assumed to be more intelligent and analytical than women), men as a group are regarded as more prototypical than are women as a group. When one group within a category is more prototypical than the others, the most prototypical group enjoys a default sense of belonging and higher status, while non-prototypical groups face being viewed as inferior and expected to conform to the prototypical group's norms (Mummendey & Wenzel, 1999).

Scholars in both sociology (e.g., status characteristics theory, Ridgeway, 1991; Ridgeway, Li, Erickson, Backor, & Tinkler, 2009) and psychology (e.g., lack of fit model, Heilman, 1983; and role congruity theory, Eagly & Karau, 2002), have further enhanced our understanding of the importance of prototypes in group-category fit, and in particular their relevance to women within masculine-prototyped professions. These literatures have highlighted how the misalignment between the traits stereotypically associated with women (e.g., warmth, compassion, etc.) and the traits associated with masculine professional prototypes (e.g., competence, assertiveness, etc.) results in the prejudice that many women face in the workplace. Indeed, the more masculine a profession is, the more disadvantageous it is to be associated with

stereotypically feminine traits (Cejka & Eagly, 1999; Lyness & Heilman, 2006). As a result, even when their performance is identical to men, women and their contributions are perpetually devalued in professions with masculine prototypes (Heilmen & Haynes, 2005; Inesi & Cable, 2015; McClean, Martin, Emich, & Woodruff, 2017).

Problems with Changing Groups to Match Professional Prototypes

Several strategies have been proposed to correct for the inequalities created when the stereotypes about a group are misaligned with professional prototypes. One approach has been to change the way in which groups or individual group members are perceived. However, shifting existing stereotypes about and biases against specific demographic groups has proven tremendously difficult and resistant to a wide array of interventions (Lai, et al., 2014; Lenton, Bruder, & Sedikides, 2009). Confronting people with individuals who defy the stereotypes associated with their group rarely changes perceptions of that group (see Richards & Hewstone, 2001 for review). More often, individuals who defy their group stereotypes are subtyped and marked as “exceptions to the rule” and “fenced off” from the group they originally belonged to. Alternatively, subgrouping may occur, splitting a once homogenous group into smaller groups with separate sets of stereotypes and associated group status. For example, women in male-dominated professions could be split up into passive women who are likable but unlikely to succeed and more aggressive “queen bee” women who are expected to succeed but are unlikable (Derks, Ellemers, van Laar, & de Groot, 2011; Fiske, Cuddy, Glick, & Xu, 2002; Wessel, Hagiwara, Ryan, & Kermond, 2014). Those women whose success is attributed to their perceived adoption of masculine tendencies face additional penalties for defying gender stereotypes (Prentice & Carranza, 2002; Rudman & Glick, 1999) and threatening men’s privileged status (Berdahl, 2007; Vandello & Bosson, 2013). Thus, they are not rewarded with the same professional recognition and outcomes as are members of the currently prototypical group who

behave the same way (Brescoll, Dawson, & Uhlmann, 2010; Heilman, Wallen, Fuchs, & Tamkins, 2004; Ryan & Haslam, 2007).

Challenges with Changing Professional Prototypes

With so many problems arising from attempts to force groups to conform to existing prototypes, a better solution may lie in changing professional prototypes themselves. Prototypes are dynamic and sensitive to contextual influence (DuRue & Ashford, 2010, Lord, Brown, Harvey, & Hall, 2001; Smith, 1998), and so are susceptible to change. However, to date, no scholarship has presented an effective strategy for generating truly inclusive prototypes in which multiple groups can be recognized as clearly and equally prototypical.

One primary barrier to changing professional prototypes so that they no longer disadvantage non-dominant groups is that this may be seen as a threat to the dominant, currently prototypical group. Because of the many benefits to being prototypical, one would not expect currently prototypical groups to voluntarily redefine professional prototypes in ways that diminish their relative advantage (Wenzel, Mummendey, & Waldzus, 2007). In fact, dominant groups typically use their power to reinforce their privileged prototypical status by promoting representations and conceptualizations of the category that favor their group (Mummendey & Wenzel, 1999; Legault, Gutsell, & Inzlicht, 2011). Consider the case of computer programming. In the early days when computer programming was largely characterized as clerical work, women were considered ideal for the profession (Ensmenger, 2012). However, as computing became more technical, high-status, and well-compensated, the prototype of a computer programmer transformed into the masculinized image of a programmer that we have today (Ehrlinger, Plant, Hartwig, Vossen, Columb, & Brewer, 2018). With the exception of a few recent gender diversity interventions, men in computer programming have typically preferred representations of their

profession that reinforced the notion that only men could be true programmers (Cheryan, Plaut, Handron, & Hudson, 2013). Therefore, intervention is needed to address this issue.

Several researchers have proposed methods for more deliberate prototype change, but each has noteworthy limitations. For example, some scholars have argued for replacing prototypes that privilege one group with prototypes that privilege a different group, such as substituting the masculine leadership prototype with a stereotypically feminine one (Helgesen, 1990; Rosener, 1995) or defining professional goals to better align with the more communal goals of women (Diekmann, Brown, Johnston, & Clark, 2010). Some researchers have found that in the context of negotiations, replacing the masculine prototype with a feminine one has improved the performance of women (e.g., Kray, Galinsky, & Thompson, 2002). Others have found that downplaying the perceived masculinity of jobs by removing Star Trek posters from a computer science classroom (Cheryan, et al., 2009), or reducing the use of masculine words in job postings (e.g., replacing “strong communicator” with “excellent communicator”; Gaucher, Friesen, & Kay, 2011; Sczesny, Formanowicz, & Moser, 2016), results in greater interest from female applicants.

Although these strategies can increase women’s negotiation outcomes and interest in male-dominated professions, replacing or negating existing masculine prototypes produces some negative outcomes, as well. For instance, scholars have shown that feminizing professional prototypes by adding a set of stereotypically feminine traits (e.g., helpfulness and sensitivity to others’ needs) to an otherwise masculine or gender-neutral job description can benefit the perceived hireability of stereotypically “communal” female applicants (Rudman & Glick, 1999; 2001). However, by feminizing the professional prototype in this way, counter-stereotypical “agentic” women experience drops in perceived hireability, exacerbating the penalties these women already face for violating prescriptive gender stereotypes (Rudman & Fairchild, 2004). In addition, feminized prototypes may be seen as threatening to men’s prototypicality. Changing

prototypes from masculine to feminine may trigger *prototypicality threat*, the concern among members of dominant groups that they may lose their prototypicality status (Danbold & Huo, 2015; 2017). In defense, threatened men will express greater opposition to gender diversity, and may label feminized prototypes as illegitimate, undermining any meaningful change (Johnson et al., 2006; Tost, 2011).

Other scholars have tried to create a unified prototype by increasing the perceived similarities among groups within organizations (e.g., Chatman, Polzer, Barsade, & Neale, 1998). This approach, however, may be interpreted as downplaying critical between-group differences, upsetting members of the non-prototypical group (Plaut, Thomas, & Goren, 2009) and doing little to alleviate the threat felt by members of the currently prototypical group. Another related approach has been to present a prototype as obscure or complex (e.g., saying there is no consensus on the prototype; Alexandre, Waldzus, & Wenzel, 2016; Waldzus, Mummendey, Wenzel, & Weber, 2003), thus preventing any group from claiming prototypicality. Although this may reduce the unequal distribution of prototypicality among groups, having no consensually agreed upon prototype introduces *prototype ambiguity*, which impairs group cohesion and coordination in professional settings (Bartel & Wiesenfeld, 2013). In sum, in professions with strong masculine prototypes, maintaining the traditional prototype disadvantages women, shifting to a feminine prototype elicits threat and backlash, and erasing or obscuring the prototype creates dysfunctional ambiguity.

Balanced Prototypes

In order to avoid these pitfalls and create an inclusive, widely accepted, and clearly defined prototype, we introduce the concept of a *balanced prototype*. Rather than thinking about prototypes in terms of which one group is exclusively prototypical and which others are not (e.g., men are prototypical and women aren't), we conceptualize prototypes as a constellation of traits

essential for success and membership in that category. In many categories, there is not a perfect match between this constellation of essential traits and the reductive set of traits that are stereotypically associated with the most prototypical group. However, because the currently prototypical group generally has the status and power to enforce a definition of the professional prototype that works to their advantage (Wenzel, et al., 2007), traits that are associated with them tend to be prioritized when assessing fit, while traits that are not, despite being legitimately valuable, are downplayed (Heilman, 2001). Returning to the example of computer programmers, men, as the currently prototypical group, are likely to strongly emphasize the category trait of technical skills, which are stereotypically masculine, to inflate their group's prototypicality. However, programming is almost always done in teams, and social skills (which are stereotypically feminine) are incredibly important to professional success (Joshi, Trauth, Kvasny, & McPherson, 2013). Nevertheless, if technical skills are valued more than social skills in selection processes, it enhances the positive association between stereotypical masculinity and being a true and effective computer programmer, and minimizes the prevalence of necessary social skills among computer programmers.

With a balanced prototype, we are trying to weaken the association between trait masculinity and the importance of traits in the prototype. By leveraging the legitimacy of important but devalued category traits, a broader set of groups can be considered representative of the prototype. In other words, if social skills (and other stereotypically feminine traits) were recognized as equally important as technical skills (and other stereotypically masculine traits) in fit assessments of computer programmers, then gender would lose its predictive power in determining category fit. Furthermore, because both men and women would be able to see potential alignment between their group and essential traits in the professional prototype, there is

little risk that this change would elicit prototypicality threat among men. Finally, because the prototype traits are clearly defined and legitimate, there should be no prototype ambiguity.

Using Prototype Inversion to Balance Prototypes

One might expect that the simplest way to create a balanced prototype is through the clear articulation and promotion of such a prototype (e.g., stating that stereotypically feminine and masculine traits are *equally* important to the profession at hand). We argue, however, that such an approach would be insufficient to offset the default prioritization of masculine traits in an established prototype that advantages men. Instead, the best way to achieve a balanced prototype is through exposure to a carefully constructed *inverted* version of the existing prototype; providing a deliberate counterweight to the pre-existing weighting of traits. For example, an inverted prototype intervention in computer programming would make the persuasive case that social skills are the most essential skills to succeed in coding, while still recognizing the importance of technical skills, thereby emphasizing most the trait that needs to be elevated to achieve a balanced prototype. So long as it is viewed as legitimate, we predict that the counterweight offered via prototype inversion would serve as an effective means of creating a balanced prototype in which masculine and feminine traits are considered equally important.

Once a prototype inversion intervention has been applied, rather than shift the privilege of prototypicality from one group to the other (i.e., femininizing the prototype as in Rudman & Glick, 1999), a representative member of this more balanced prototype – male or female – will need to manifest a holistic mix of both stereotypically masculine and feminine traits. Thus, we predict that prototype inversion could be perceived as legitimately expanding the range of demographic groups that fit the professional prototype, making it more inclusive of both men and women who have the mix of competencies necessary for professional excellence.

Conceptualizing and testing balanced prototypes in this way represents an important advance in understanding how to make prototypes more inclusive.

RESEARCH CONTEXT: THE U.S. FIRE SERVICE

We conducted qualitative and quantitative research to determine if it is possible to create a more inclusive prototype of the extremely male-dominated context of the U.S. fire service. Of the approximately 350,000 professional (not volunteer) U.S. firefighters, only 3.7% are female (Haynes & Stein, 2014). This low representation of women is especially stark when compared to other professions with similar physical requirements and risks, including the military (15%) and police (13%). For a female firefighter, it is extremely rare to have another woman on the same shift, let alone in the same station, and for a male firefighter, it is extremely common to have no female colleagues in the same station or shift.

Numerous academic and industry studies of fire departments indicate that success in the fire service requires a range of personal characteristics related to the core functions of fighting fires (e.g., physical strength and courage), providing medical care (e.g., intelligence and compassion), and working as effective, resilient teams (e.g., companionate love, joviality, and trust) (Chetkovich, 1997; Delsohn, 1996; Desmond, 2006; Hardison, et al., 2015; O'Neill & Rothbard, 2017; Pratt, 2017). For instance, in O'Neill and Rothbard's (2017) study of the emotional cultures of fire departments, they found that the two most functionally-adaptive emotions were joviality ("markedly good humor" p. 9), which helped build team cohesion and deal with trauma, and companionate love ("warmth, connection, and affection" p. 11), which created family-like bonds among firefighters.⁴ These observations suggest that *social and*

⁴ Although there is some conceptual overlap between the emotions of joviality and companionate love and some of the traits of the firefighter prototype we identify (e.g., compassion), emotional cultures and the traits of a professional prototype are distinct constructs. We also note that the expression of these emotions (e.g., communicating joviality via humor) may be done in a stereotypically masculine way, highlighting the

emotional skills are critical to success in the fire service; thus, they should be core traits of the firefighter category prototype. Because these kinds of traits are typically congruent with the feminine stereotype dimension of warmth (Fiske, Cuddy, Glick, & Xu, 2002; Prentice & Carranza, 2002), it should be relatively easy to see women as fitting the firefighter prototype. So why is female representation in the fire service less than four percent?

Although we had theoretically-derived predictions about the role of masculine prototypes in the underrepresentation of women in the fire service, we wanted to first determine how our theory manifests in this context, and how prototype inversion might be effectively operationalized. Therefore, we conducted an inductive, qualitative study about the firefighter prototype and the consequences it has for non-prototypical women firefighters.

STUDY 1: INDUCTIVE INVESTIGATION OF THE MASCULINE FIREFIGHTER PROTOTYPE

Method

Participants. The authors conducted 60 to 90-minute one-on-one interviews with 14 female firefighters and 15 male firefighters from eight departments nationwide. In addition, we met in focus groups with 72 male and one female on-duty firefighters during 15 four- to seven-hour station visits in four different departments that included one of the largest in the country with roughly 3,000 professional firefighters, and three average-sized departments with 65 to 100 personnel each. We targeted one of the smaller departments for complete contact with all stations and shifts to make sure our interviews were not vulnerable to sampling bias. We also had 25 one-on-one interviews with fire service management and union leaders (from 30 to 90 minutes each) and observed two days of fire academy training and testing. We continued seeking data collection

fact that not all social and emotional skills are stereotyped as feminine, especially in such strongly masculine professions such as firefighting.

opportunities until we reached a point of saturation where we consistently heard similar information and little new information, and our observations were acknowledged by subject-matter experts as representative of their experiences (Strauss & Corbin, 1998; Yin, 2013).

Procedures. Semi-structured interviews were generally conducted by a female and a male interviewer together, but on some station visits, only one or the other was present. We did not note any differences in responses when only one interviewer was present based on the interviewer's gender. We told the participating firefighters that we were broadly interested in their experience as firefighters, and opened conversations with general questions about their daily life in the fire service. We asked participants to share with us how they identify successful firefighters, and what traits they deemed to be the most important. We also asked about their experiences as and/or perceptions of rookie (recruit and probationary) firefighters, and about female and ethnic minority firefighters of all tenures. We probed for their key concerns, noted inconsistencies in their statements, and asked for detailed examples and stories that illustrated their opinions.

When possible, we audio recorded the interviews and we always took detailed handwritten notes, which we reviewed immediately after each station visit or interview. Following recommended inductive data analysis techniques (Eisenhardt, 1989; Miles & Huberman, 1994; Strauss & Corbin, 1998), we iteratively identified quotes and examples that both illustrated and contradicted emerging themes, then followed up in subsequent interviews/observations to reconcile our data. After completing data collection, the authors independently reviewed the full set of interview notes through two rounds of coding. In the first round we identified the traits that comprised the firefighter prototype. In the second round of coding, we focused on the experiences of female firefighters, given the prototype traits that we identified. Following data collection and analysis, we presented our preliminary findings to seven groups of fire department chief officers

and union leaders for their feedback, who validated the conclusions we drew as accurate to the fire service.

Although we believe that we spoke to a sufficiently large sample of firefighters to reach a point of saturation, our data are not amenable to quantitative analysis because we deliberately oversampled women firefighters. Therefore, presenting quantitative descriptives of these data (e.g., the percentage of interviewees who endorse a particular trait as central to the fire service), risks being unrepresentative of the underlying population. To more faithfully describe the main themes that we heard throughout our interviews, we present illustrative quotes where similar sentiments were endorsed by the majority of our sample. Where quotes are presented, we indicate if the speaker was male (M) or female (F).

Results

Prototype traits. We found that firefighters identified a wide range of traits when describing what was important to being a successful, prototypical firefighter. These included team-orientation (e.g., “The biggest concern is to make sure other crew members get home safe.” (M)), compassion (e.g., “You innately need to have some drive to help people” (F)), physical and mental strength (e.g., “Some physical requirements to do this job...being strong and in good fitness” (F)), and a drive to continuously learn and improve (e.g., “Being self-driven, willing to overcome obstacles in yourself” (M)). We recorded endorsements of all of the traits listed in Table 1, from both male and female firefighters with relatively equal frequency, and these traits were substantiated by numerous anecdotes and our observations.

Insert Table 1 about here

An illustrative example is of a prototypical male Engineer we met during a station visit. He was undoubtedly physically fit and his job competence was evidenced by the extent to which the Captain (his immediate supervisor) consulted with him. He was clearly admired by everyone

else on his crew, and his gifted sense of humor often made him the center of attention in the station. In addition, he displayed a caring heart, as evidenced in the following anecdote. We rode along on his engine to a routine medical call at a local assisted living home to which the station regularly responded. Although the Engineer was a paramedic, he did not take the patient care lead on this particular call, so he waited in the hallway outside the patient's door. There he spotted another resident – a frail, elderly woman – sitting on a bench. The Engineer walked over to her and said, “Didn’t we help you here last week? How are you feeling, my friend?” He sat down and engaged her in a spontaneous, discretionary follow-up care conversation. He was tender, gentle, and kind. When we commented to the Captain privately about how struck we were by this behavior he replied simply, “that’s what we do.”

Physical Strength as a Litmus Test for Fit. In general, when listing traits essential to the job, most of the firefighters we spoke to did not weigh any single trait over others. In fact, many said that although physical strength is the most common metric used to evaluate prospective firefighters’ capability for the job, it is not *consistently* the most critical characteristic needed to do the job well (e.g., “Brainpower is more important than physical strength” (M); “While it is important to have strength, it’s more important to know body mechanics and knowledge on the fire ground and in day to day operations.” (M)). After all, the vast majority of calls are medical and do not require exceptional strength, and many firefighting tasks (e.g., throwing ladders) require good technique as much as brute strength. Despite this wide acknowledgment of a complex firefighter prototype, however, we observed a consistent change in the weighting of traits when the conversation shifted to the topic of female firefighters.

Both men and women firefighters tended to increase their emphasis on the importance of physical capabilities when they talked about ascertaining the ability of women firefighters. Our participants tended to reiterate that there are physical requirements to do the job and that if

someone is physically deficient they endanger themselves and everyone else on the crew.

Although this is certainly a valid concern about any firefighter, broad societal stereotypes about women's relative weakness generated heightened skepticism about the strength of most women firefighters. Thus, our interviews suggest that, in both formal training and evaluation procedures and in the informal process of developing interpersonal trust and esteem, firefighters use strength as a *litmus test* to identify and weed out weak links in the department.

A particularly pervasive assumption we heard was that increasing the representation of women in the fire service required lowering the physical standards of the job: "They had to change things to make it easier to physically pass and women started getting hired just because they were women" (M). In other words, the only way these firefighters believed that enough women could qualify for the fire service to increase their percentage representation was to lower the requirements of the job in a way that endangered everyone else. Women on crews, thus, were characterized as threats rather than as resources. The pervasiveness of the perception of women as threats to crew safety was illustrated by a male Battalion Chief who told us that when a woman was recently assigned to a new station under his command, he immediately received five requests to transfer out from the men in that crew.

Social and Emotional Skills Downplayed in Fit Assessments. Whereas physical strength aligns with masculine stereotypes and was used to rationalize women's lack-of-fit with the firefighter prototype, other important traits like warmth and compassion align more with feminine stereotypes. Like the Engineer described above, interviewees consistently reiterated and reinforced the importance of compassion and empathy on calls (e.g., "we have to keep in mind that we are entering people's homes during their worst experiences" (M)), and to support one another's processing of traumatic experiences by monitoring each other's emotional needs. One

of our interviewees eloquently summed up this theme as, “we can train your body and teach your mind, but you need to have a certain *soul* to be a firefighter” (M).

In theory, stereotypes about women’s warmth should give female firefighters a comparative fit advantage on these dimensions that could potentially offset their fit disadvantage on the masculine prototype dimensions, like strength. However, when the topic of female firefighters came up, the importance of physical strength was consistently and spontaneously invoked to justify the relative absence of women in the fire service, but the importance of compassion was rarely, if ever, brought up to argue for bringing more women into the profession.

Lack-of-Fit as Experienced by Women Firefighters. Almost all the female firefighters we talked to reported being treated as if they didn’t belong in the profession. Women firefighters routinely faced skepticism and criticism and reported needing to constantly prove themselves as exceptional under an intense spotlight (e.g., “I’ve always felt that I have to prove myself on every call, every time;” “everyone expects you to fail”). The assumption of physical inferiority became a self-fulfilling prophecy when women were excluded from important training opportunities. For example, one woman reported that her crewmates had refused to take her to practice driving the fire engine while on shift – violating a departmental training obligation – so she had to find people on other battalions that were willing to teach her on her off-duty days. A man told us of observing a recruit academy where a female rookie was struggling to extend a ladder while several of the male trainers looked on. He suggested a change that improved her leverage so she was able to complete the drill. He said, “[e]ither those guys didn’t know she was doing it wrong or they chose not to teach her.” Furthermore, when we confronted some male interviewees with examples of successful female firefighters, they deemed those women “exceptional” and “unicorns” (clear examples of subtyping; Richards & Hewstone, 2001) rather than examples of the kinds of women the department could attract on a larger scale.

Another consequence of women's lack-of-fit is social exclusion. Women reported often being left out of casual socializing, undermining their ability to both demonstrate their social skills and successfully integrate into their teams. For example, "[a]t the start of each shift when the guys are all shooting the shit, when I come over to join they all stop and ask me what I need...I hate them thinking I always need something, so now I don't even bother." Many men reported an aversion to having to "behave differently" and to "walk on eggshells" around women, and expressed concern about the impact of women's emotional volatility on crew cohesion and effectiveness (e.g., "What if she cries? Or has her period?"). They would not direct pranks and teasing at their female colleagues for fear of "offending them" and "getting written up" (i.e., the women filing a formal complaint), which meant the women were excluded from participating in the jovial emotion culture and social rituals that produced resilience, signaled a good attitude, established trustworthiness, and created a sense of belonging (O'Neill & Rothbard, 2017; Pratt, 2017).

Discussion

The inductive inferences we drew from these qualitative data helped to explain the paradox of why the representation and valuation of women is so low in the fire service despite the obvious importance of social and emotional skills to firefighting. Although skills like team orientation and compassion were widely endorsed as legitimately central traits of the firefighter prototype, strength and stoicism were invoked to legitimize the idea that women don't belong in the fire service. Furthermore, women's potential comparative advantage for contributions on the prototype traits that align more with feminine stereotypes were unrecognized and actively thwarted through routine exclusion from the adaptive emotion-cultures of joviality and companionate love and undervaluation of their trustworthiness cues. Thus, despite many

characteristics being recognized as essential to the job, the firefighter prototype is, in practice, weighted such that it ensures a fit for men and not for women.

STUDY 2: MANIPULATING THE FIREFIGHTER PROTOTYPE

Prototype Inversion Hypotheses

To counter this devaluation of genuinely important stereotypically feminine traits (and by extension, women), if people were to use *both* compassion and strength to assess one's potential as a successful firefighter, then they might be able to more readily recognize the value of women (for whom the trait of compassion aligns with the stereotype of warmth) who are also strong, and of men (for whom the trait of physical strength aligns with the stereotype of competence) who are also compassionate. Because this constellation of category traits is a balanced mix of both stereotypically feminine and masculine traits, we posit that this would increase the perceived fit of women without incurring a cost to men's perceived prototypicality.

We hypothesized that we can produce this balance by inverting the relative importance of feminine stereotyped traits, without denying the importance of masculine stereotyped traits. We predicted that this kind of inverted representation of the prototype would have two concrete effects compared to the current representation of the prototype. First, it would increase the perceived ability of female firefighters and second, it would reduce opposition to female firefighters. Furthermore, we predict that both the higher perceived ability of women and lower opposition to women in the fire service when the prototype is inverted compared to alternative representations of the prototype are due to the prototype being more balanced. That is, feminine traits are recognized as relatively more essential to professional success (i.e., more important in the category prototype) and more comparably essential to masculine traits in the prototype, such that the association between trait masculinity and trait importance would be non-significant.

In addition, we expected some important null effects. Because we are affirming the importance of masculine traits and being concrete about the included traits, prototype inversion should not induce more prototypicality threat or prototype ambiguity than alternative representations of the prototype. Because we expected our approach to function without generating threat or ambiguity, we also tested if an inverted prototype would be considered as legitimate as other versions of the prototype.

We report details and results of pilot studies we conducted of the prototype inversion manipulation and full study materials in the Appendix.

Study 2a: Firefighter Field Experiment

Method

We tested if inverting the firefighter prototype would increase the perceived ability of women firefighters and lower opposition to women in the fire service and if these effects were due to feminine stereotyped traits being seen as relatively more essential to the prototype.

Participants. Four hundred and eight active duty firefighters from two departments voluntarily completed our survey. With strong endorsement from both the administration and union officials, our 15.3% response rate exceeded our expectation of 10% that was based on participation in previous voluntary surveys in the departments. Compared to actual department demographics, women (in our sample 3.81% in our sample vs. 1.67% in the participating departments), ethnic minorities (42.23% vs. 40.28%), and personnel ranking above firefighter (63.11% vs. 56.37%) were all slightly overrepresented in our sample.

Procedure. Participants were randomly assigned to one of five conditions designed to manipulate their conception of the prototypical firefighter. All participants first watched a short video in which a researcher explained the survey as being part of a project on “modern day firefighting in the United States” (transcripts of experimental stimuli are in the Appendix). In

preparation for the subsequent video, which served as the experimental manipulation, participants were then told that the researchers had interviewed and recorded firefighters reading their responses to the question, “[w]hat are the most important characteristics for modern firefighters to have in order to succeed in the fire service?”

Participants then watched a video of a White male fire Captain from a fire department that was not participating in the research study. We selected a prototypical White man to deliver the message in our manipulation because his authority and agenda wouldn’t be questioned like a woman or minority figure’s might be. The firefighter shared what he stated were, “in [his] experience, the most important characteristics to be successful as a firefighter.” In reality, he read scripts (that he had personally endorsed) describing three key traits of the firefighter prototype that we determined through pilot testing (reported in the Appendix) varied in gender-stereotypicality: physical strength (masculine), compassion (feminine), and team orientation (gender-neutral).

We created two alternative manipulations that emphasized the relative importance of physical strength (the “Traditional Prototype”) and compassion (the “Inverted Prototype”). The traits were ranked *in order of importance* as physical strength first, team orientation second, and compassion third (in the Traditional condition), and as compassion first, team orientation second, and physical strength third (in the Inverted condition). Brief rationales were given for each trait and a recap of the traits was repeated at the end in the same order. Aside from changing the ordering and importance of the characteristics all the language and delivery was identical between these conditions. Pre-testing supported our prediction that it was necessary to list these traits “in order of importance,” rather describe them as being equally important to actually change participants’ prototype perceptions. When we presented these traits as “equally important,”

participants defaulted to the ranking of traits in the Traditional condition (reported in Pilot Study 2 in the Appendix).

We created three different control conditions. In the Neutral video control condition, only team orientation was identified as an important trait and described as in the other two video conditions. We included this condition to control for exposure to the fire captain and any description of prototype characteristics. We also created a No Video control condition in which participants did not see either the introductory or the fire captain's videos and instead went straight into the dependent variable's measures. This control condition captured a pure baseline of opinions about the fire service. Finally, the Introductory Video Only control condition tested the possibility that exposure to the occupational description in the survey introductory video, but not the video of the fire captain, might affect participants' responses.

In only the video conditions, participants were asked to write in a text box "a few sentences about why [compassion/physical strength/team orientation] could be considered the most important trait of a successful firefighter, above all other characteristics" to reinforce the manipulation.

We tested the effect of these experimental conditions on the perceived ability of rookie firefighters that varied in terms of race and gender (described below). This generated a five between-group (prototype manipulation) x two within-group (gender of evaluated firefighter) x two within-group (race of evaluated firefighter) experimental design. We imposed quota to ensure a minimum of 100 observations in each of the Traditional and Inverted focal conditions.

Perceived Ability of Rookie Firefighters. To measure the perceived ability of men and women rookie firefighters, we consulted with several of the fire service professionals to develop a realistic and face-valid, complex scenario in which a number of traits would be important for success. We created an archetypical scenario of rescuing and caring for a victim trapped inside a

burning building, a situation in which both strength and compassion are necessary. We asked participants to imagine themselves as a Battalion Chief on the scene of a large fire. We made clear that a Battalion Chief is a supervisory role with evaluative responsibility for the multiple companies that are described in the scenario. From each of the four responding companies, a team of one experienced and one rookie firefighter was responsible for entering the burning building, carrying out victims and administering first aid. Participants were asked to evaluate their confidence in the abilities of the rookie firefighters to conduct this task. Participants were told that all rookies had performed well in training to eliminate any ability differentiators.

Because of the historical presence of racial discrimination in the fire service (Hulett, Bendick, Thomas, & Moccio, 2008), to validate that gender is the primary dimension along which prototypicality is determined in the fire service, and to obscure our interest in gender from participants to reduce demand characteristics, we also examined differences in the evaluation African American and White American firefighters. Rookies that varied in terms of gender and race were presented to participants in random order with a photo and stereotypical name to match (Bertrand & Mullainathan, 2003). There was a White man (Brad Taylor), a Black man (Jamal Wilson), a White woman (Allison Smith), and a Black woman (Lakisha Johnson).

Participants were then asked to indicate the extent to which they agreed (1 = strongly disagree, 7 = strongly agree) with four evaluative statements about each rookie firefighter: (e.g., "I would feel confident sending [firefighter name] into the fire to carry out and care for the victim." White Man $\alpha = .88$, Black Man $\alpha = .90$, White Woman $\alpha = .93$, Black Woman $\alpha = .92$). We conducted within-subject comparisons by race for targets of the same gender and found no differences in any conditions (all $ps > .40$), so we pooled responses by race to measure the perceptions of female firefighters (Men $\alpha = .93$; Women $\alpha = .95$). We included perceptions of

male firefighters as a control in our analyses to account for the relative perceptions of female firefighters' ability in case the manipulations also affected perceptions of men's abilities.

Opposition to Women in the Fire Service. We developed an opposition to women in the fire service scale based on the qualitative data we collected in Study 1. We told participants, "Many fire departments have made it a priority to increase the representation of women in their ranks without lowering recruitment and training standards" and were asked to rate their agreement (from 1 = strongly disagree, 7 = strongly agree) with 5 items (e.g., "I think efforts to increase the representation of women in the fire service would make the fire department less safe." $\alpha = .94$).

Mediator: Perceived Trait Importance. To test the prediction that the effects of prototype inversion would be mediated by a relatively greater valuation of feminine traits in the prototype, we included a measure of the perceived importance of traits that we had previously identified as comprising the firefighter prototype (Pilot Study 1, in the Appendix). Given survey length constraints imposed by the fire departments, we identified a subset of 11 of the traits that were the most clearly associated with masculine, feminine, and gender-neutral stereotypes, along with the three focal traits of physical strength, compassion, and team orientation that we included as manipulation check items. We excluded these items from the tests of our primary hypotheses to ensure that people were not responding to the words highlighted in the manipulation. We asked participants to indicate the extent to which they thought each of those traits was "essential to being a successful modern firefighter" (1 = somewhat essential, 7 = absolutely essential). Based on principal components factor analysis with oblique rotation (because the traits comprised the same prototype and were correlated), we created composites of the distinctly masculine and feminine traits from our prototype trait configuration that loaded onto separate factors at greater than .50. The two masculine-stereotyped traits were mechanical aptitude and physical and mental

stamina ($r = .45, p < .001$; M femininity- masculinity rating from Pilot Study 1 = 5.30, $SD = .92$). The two feminine-stereotyped traits were patience and empathy ($r = .53, p < .001$; M femininity- masculinity rating from Pilot Study 1 = 3.28, $SD = 1.07$). To model the relative value of the feminine traits in the full prototype trait configuration, we evaluated the effects of the feminine trait scale ratings controlling for those of the masculine traits scale and one gender-neutral trait, selflessness (M femininity- masculinity rating from Pilot Study 1 = 3.84, $SD = 1.24$).

Controls. As mentioned earlier, women, ethnic minorities, and personnel ranking above firefighter are all overrepresented in our sample, indicating their importance as control variables in our models. Furthermore, including participants' gender and ethnicity (dummy-coded as 0 = Non-White, 1 = White due to small cell sizes among different non-White groups) controlled for potential in-group favoritism in evaluations of female and Black firefighters. Given the hierarchical nature of the fire service and the potential variance in exposure to female firefighters that comes with more experience and supervisory responsibility, we included participant's rank in the department. Because we collected data from two samples, we also included a fire department fixed effect variable.⁵

Reactance to Manipulation. Prior to data collection, we were warned by fire service professionals that some firefighters would reject messages they perceived as coming from "outsiders telling them what to think." Cautious of the possibility that a segment of our sample would reject our manipulation outright, we had two research assistants code the text responses to the manipulations for statements where participants challenged the importance of the trait they were prompted to write about (Cohen's kappa = .90). They identified 43 cases of reactance to our manipulation (e.g., in the Traditional condition, "I disagree that physical strength is the first

⁵ Analyses excluding these control variables produce consistent results, although some contrasts fall above the $p = .05$ threshold.

priority. Physical strength is rarely needed, certainly not on a daily basis.”). Rates of reactance to our manipulation were identical between participants in the Inverted (15.7 percent) and Traditional (15.7 percent) conditions (2.4 percent of the Neutral condition exhibited reactance), suggesting there was nothing particularly objectionable about the prototype inversion framing. While 10.5 percent overall reactance to our manipulation may seem high, their overt rejection of the manipulations invalidated our ability to test the hypothesized effects of being in a state of mind that reflected prototype inversion or one of our alternate conditions. Including participants coded for reactance to our manipulation somewhat weakened our observed effects. For example, the effect of Condition on the perceived ability of women firefighters drops to $F = 2.54$ ($p = .08$) with their inclusion, though the pattern remains consistent with what we report below. The excluded respondents did not differ demographically from the 365 who were included in the analyses (reported in Table 2).

Results

We report descriptives and correlations in Table 2. We determined that there were no significant differences among any of the control conditions (Neutral, No Video, and Introductory Video Only) on any dependent variables, so we combined them into a single Pooled Control condition variable.

Insert Table 2 about here

Manipulation Checks. ANOVA and planned contrasts of the trait importance ratings of physical strength indicate a significant effect of Condition ($F = 5.42$, $p = .005$), with physical strength rated as most important in the Traditional condition ($M = 6.56$, $SE = .10$; Inverted: $M = 6.15$, $SE = .10$, $z = 3.06$, $p = .002$; Pooled control: $M = 6.23$, $SE = .08$, $z = 2.71$, $p = .007$).

ANOVA of the compassion trait importance ratings indicate that the Inverted prototype manipulation was also successful (Condition $F = 3.45$, $p = .033$), with compassion rated as most

important in the Inverted condition ($M = 6.48, SE = .12$); Traditional: $M = 6.20, SE = .12, z = 1.70, p = .089$; Pooled control: $M = 6.09, SE = .09, z = 2.61, p = .009$), although the contrast with the Traditional condition is marginally significant. Because team orientation was mentioned as similarly important in the Traditional, Inverted, and Neutral only control conditions, we did not expect it to vary by Condition ($F = .26, p = .771$).

Perceived Ability of Women Rookie Firefighters. We first determined that, controlling for the perceived ability of male firefighters,⁶ the total effect of Condition on the perceived ability of female firefighters was significant ($F = 4.19, p = .016, \eta^2_p = .014$). Ratings of female firefighters were higher in the Inverted condition ($M = 4.86, SE = .09$) than in the Traditional condition ($M = 4.51, SE = .14, z = 2.13, p = .033$), and the Pooled Control condition ($M = 4.49, SE = .12, z = 2.53, p = .012$). Descriptively, the percent by which women were rated lower than men was smallest in the Inverted condition (Pooled Control 14.4%; Traditional 14.8%; Inverted 5.4%).

To see if the increase in the perceived ability of female firefighters came at a cost to the perceived ability of male firefighters in the Inverted condition, we conducted an equivalent, post-hoc analysis with male firefighter ability ratings as the dependent variable, controlling for those of female firefighters. In this case, the total effect of Condition was not significant ($F = 1.87, p = .155, \eta^2_p = .006$) with ability ratings of male firefighters in the Inverted condition ($M = 5.10, SE = .09$) marginally lower than in the Traditional condition ($M = 5.34, SE = .11, z = 1.81, p = .071$) and no different from in the Pooled Control condition ($M = 5.26, SE = .09, z = 1.35, p = .177$). Thus, it appears that the Inversion manipulation substantially increased perceptions of female firefighters' ability and only slightly lowered perceptions of male firefighters' ability.

⁶ We also ran a multilevel repeated model looking at the perceived ability of men and women rookie firefighters simultaneously and found significant results in the same pattern. We report the results that are easier to interpret and more closely parallel the subsequent test of indirect effects, but the others are available upon request.

Opposition to Women in the Fire Service. We next tested the prediction that the prototype inversion manipulation would decrease opposition to women in the fire service. This Condition effect was marginally significant ($F = 2.82, p = .061, \eta^2_p = .011$), with the least opposition evident in the Inverted condition ($M = 3.80, SE = .14$; v. Traditional: $M = 4.14, SE = .17, z = 1.51, p = .132$; v. Pooled Control: $M = 4.25, SE = .12, z = 2.33, p = .020$). We note that although the Inverted condition significantly reduced opposition to women in the fire service compared to the Pooled Control conditions, the planned contrast with the Traditional condition was not significant. Although the hypothesis was not fully supported, we can nevertheless conclude that prototype inversion at least did not induce more backlash against women than did the traditional representation of the prototype.

Perceived Importance of Feminine Prototype Traits. To examine our mediation hypotheses, we first tested if the feminine traits were perceived as more equivalently important to the masculine traits in the prototype in the Inverted compared to the other conditions by conducting a repeated measures ANOVA on the feminine and masculine traits ratings in each condition. Only in the Inverted condition were the feminine traits ($M = 6.16, SE = .10$) rated equivalently to the male traits ($M = 6.09, SE = .10; t = .72, p = .47$). The masculine traits were rated significantly higher than the feminine traits in both the Traditional ($M_{feminine} = 5.89, SE = .10; M_{masculine} = 6.32, SE = .10; t = 4.06, p < .001$) and Pooled Control conditions ($M_{feminine} = 6.03, SE = .08; M_{masculine} = 6.20, SE = .08; t = 1.99, p = .047$).

Indirect Effects. We next conducted a series of regression models examining the indirect effect of our manipulation on a) the perceived ability of female firefighters and b) opposition to female firefighters, through the relative perceived importance of feminine traits to the firefighter prototype, followed by bootstrapped confidence intervals of the indirect effects (Table 3). In addition to the significant total effect of our manipulation on the perceived ability of women

firefighters, we observed a significant total effect on the perceived importance of feminine traits (Condition $F = 4.22$, $p = .016$, $\eta^2_p = .024$). Planned contrasts revealed that participants in the Inverted condition ($M = 6.23$, $SE = .08$) rated the importance of feminine traits significantly higher than did participants in both the Traditional condition ($M = 5.88$, $SE = .11$, $B = -.34$, $SE = .13$, $p = .010$) and the Pooled Control condition ($M = 6.00$, $SE = .07$, $B = -.23$, $SE = .12$, $p = .029$).

Insert Table 3 about here

Feminine trait ratings, in turn, were positively associated with both the perceived ability of women firefighters ($B = .26$, $SE = .11$, $p = .021$) and opposition to women in the fire service ($B = -.35$, $SE = .12$, $p = .004$).⁷ Five thousand re-sample bootstrapped and bias-corrected 95-percent confidence intervals that did not span zero indicated a significant positive indirect effect of Condition on the perceived ability of women firefighters through the greater importance of feminine traits in the Inverted condition versus the Traditional condition [-.25 to -.01], and versus the Pooled Control condition [-.16 to -.002]. Likewise, there was a significant negative indirect effect of Condition on opposition to women in the fire service through the greater importance of feminine traits in Inverted condition versus both the Traditional condition [.04 to .26], and the Pooled Control conditions [.01 to .21].

Study 2a Discussion

In Study 2a we demonstrated that inverting the firefighter prototype by emphasizing legitimate feminine traits without denying the importance of masculine traits led active-duty firefighters to increase their perceived ability of women firefighters and to not increase their opposition to women firefighters. We also found significant indirect effects for both outcomes that were due to feminine traits being seen as relatively more important to success in the fire

⁷ Although we observed a marginally significant total effect of our manipulation on opposition to female firefighters, this does not rule out the possibility of a significant indirect effect (Rucker, Preacher, Tormala, & Petty, 2011; Zhao, Lynch, & Chen, 2010).

service. These results are consistent with our theorizing that an inverted presentation of the prototype balances the prototype traits, making it harder to see a particular group as representative of the category.

Although the within-subjects nature of the perceived ability ratings made them potentially susceptible to demand characteristics, we do not believe this undermined the legitimacy of the results. If anything, demand characteristics likely led participants to under-report their gender and race bias in all conditions, yet we still observed ample evidence of a preference for male firefighters in the Traditional and Pooled Control conditions. Furthermore, the demand characteristics were no different in the Inverted than any other conditions, so cannot account for the distinctly different pattern of effects we observed there.

This study provides support for our hypotheses in a field sample of fire service professionals, but it has some limitations to the inferences we can draw from it. Primarily, because we were limited to a short survey, we measured a small number of prototype category traits and did not directly measure prototype threat, ambiguity, or legitimacy. Additionally, we noticed low but non-negligible rates of reactance to our manipulation, and that, unsurprisingly, those who failed to internalize our manipulation were less sensitive to its effects. Finally, our sample was extremely male-dominated, and we suspect that women might respond differently because of in-group bias, which potentially limits the generalizability of the findings from this context. Thus, we conducted another experiment in a general (non-firefighter) population that enabled us to address all of these limitations.

Study 2b – General Population Replication and Extension

In this study, we sought to replicate the indirect effects of prototype inversion on perceived abilities of and opposition to women firefighters, through the increased relative valuation of feminine traits in the prototype that we observed in Study 2a. Additionally, we

directly tested that the prototype inversion manipulation effectively balanced the firefighter prototype, and addressed the limitations noted above by expanding the set of prototype traits and asking questions about prototypicality threat, prototype ambiguity and prototype legitimacy. We also collected a more gender-diversified sample (but that still reflects the male-dominated characteristics of the fire service). Supported by Pilot Study 1 (reported in Appendix), we reasoned that the prototype perceptions of a general population sample are likely susceptible to being manipulated with the stimuli we created, and without the same degree of psychological identification as in the firefighter sample, less prone to reactance to the manipulation.

Method

Participants. Nine hundred and ninety-six MTurk workers based in the United States completed our survey. We intentionally oversampled men to parallel the fire service but included enough female respondents to examine the effects of participants' gender. The sample was 66 percent men, 34.67 years old on average ($SD = 11.42$), and 77 percent was White.

Procedure. Participants were randomly assigned to one of four conditions used in the prior study (Inverted, Traditional, Neutral control, and Introductory Video Only control). We excluded the No Video control condition because we wanted to ensure that our civilian participants had at least some basic knowledge about modern firefighting. This generated a four between-group (prototype manipulation) x two within-group (gender of evaluated firefighter) x two within-group (race of evaluated firefighter) experimental design.

Dependent Variables. For consistency, participants completed the same evaluations of men and women rookie firefighters as in Pilot Study 2 and in Study 2a (White Man $\alpha = .86$, Black Man $\alpha = .87$, White Woman $\alpha = .92$, Black Woman $\alpha = .91$). We determined through within-subjects analyses that perceived ability ratings within gender targets did not differ by target's race

(all $ps > .130$) and pooled the data by race (Men $\alpha = .91$, Women $\alpha = .95$). We asked participants the same items measuring their opposition to women in the fire service as in Study 2a ($\alpha = .92$).

Mediator: Perceived Trait Importance. Less confined by space than with our fire service sample, we asked participants to rate the importance of the full set of 21 firefighter traits from Pilot Study 1, including the three traits mentioned in the manipulations that we used for manipulation checks, in the same way as in Study 2a. We conducted the same principal components analysis as in Study 2a, excluding compassion, physical strength, and team orientation from the factors. We created scales from nine traits that loaded with masculine-stereotyped traits ($\alpha = .91$), and from five traits that loaded with feminine-stereotyped traits ($\alpha = .87$) cleanly onto separate factors. The gender-neutral traits that we included did not load onto a distinct factor nor hold together as a reliable scale, so we used the same single item of selflessness to represent the gender-neutral component of the prototype as in Study 2a. We again model the hypothesized mediating mechanism of a balanced prototype as the feminine-trait scale ratings controlling for ratings of the masculine-traits scale and selflessness.

Prototypicality Threat. We examined whether or not prototype inversion would trigger prototypicality threat, especially among men participants who may identify with their perceptions of men as representative of the firefighter category. Participants indicated the extent to which they agree or disagreed with the following three statements, (1 = strongly disagree, 7 = strongly agree), adapted from past research (Danbold & Huo, 2015; 2017), “I worry that in the future, men will no longer represent what it means to be a firefighter,” “I am concerned that in the future, women will represent what it means to be a fighter more than men,” “It troubles me that in the future, when people think about what it means to be a firefighter, they may not think about men” ($\alpha = .91$).

Prototype Ambiguity. We also examined whether or not prototype inversion would trigger prototype ambiguity. Participants indicated the extent to which they agree or disagree with the following three statements, (1 = strongly disagree, 7 = strongly agree), based on prior research (Bartel & Wiesenfeld, 2013), “What it means to be a successful modern firefighter is ambiguous to me,” “What it means to be an ideal firefighter is very clear to me,” (reverse-coded) “I am not sure what it means to be a true firefighter” ($\alpha = .77$).

Perceived Legitimacy of Inverted Prototype. Our theory hinges on the prototype inversion manipulation being perceived as legitimate so that it does not induce threat-based reactance and rejection. In order to examine this assumption directly, we asked participants in all conditions to, “consider the idea that the most important characteristics of a successful modern firefighter are, in order of importance: compassion, team orientation, and physical strength.” Participants then indicated the extent to which they agreed or disagreed with the following three statements (1 = strongly disagree, 7 = strongly agree), “I think this is a legitimate description of the traits of an ideal firefighter.”, “I endorse this description of the characteristics of a true firefighter.”, and “This description contradicts what I think of as an ideal firefighter” (reverse-coded). We created a scale from these items ($\alpha = .85$). For us to be confident that our inverted prototype was seen as legitimate, we would have to see scores on these statements above the midpoint (in the realm of agreement) and no lower in our Inverted condition than in the other conditions. So that exposure to this version of the prototype did not alter responses on any other dependent measure, these questions were asked at the very end of our survey.

Controls. We included controls in all of our analyses for participants’ gender and ethnicity (dummy coded as 0 = Non-White, 1 = White due to small cell sizes among different non-White groups).

Reactance to Manipulation. We again had research assistants code the text responses for reactance to our manipulation (Cohen's kappa = .996). Only one participant was coded by both RAs as showing reactance to our manipulation ("First I don't agree with this. But I guess compassion is important because we have become a weak crybaby society where the most important thing is people's feelings..."). We excluded this observation, but note that this rate of reactance to our manipulation was markedly lower than what we observed in the fire service sample. This suggests that the high rate in Study 2a was not due to the specific contents of our manipulations.

Results

We report descriptives and correlations in Table 4. We again determined that there were no significant differences between the control conditions (Neutral and Introductory Video Only) on any dependent variables, so we combined them into a single Pooled Control condition variable.

Insert Table 4 about Here

Manipulation Checks. ANOVA and planned contrasts of the trait importance ratings of physical strength indicate a significant effect of Condition ($F = 36.91, p < .001$), with physical strength rated as most important in the Traditional condition ($M = 6.63, SE = .06$; Inverted: $M = 6.39, SE = .60, z = 2.86, p = .004$; Pooled control: $M = 6.00, SE = .05, z = 8.34, p < .001$). ANOVA of the compassion trait importance ratings indicate that the Inverted prototype manipulation was also successful (Condition $F = 58.52, p < .001$), with compassion rated as most important in the Inverted condition ($M = 6.61, SE = .07$; Traditional: $M = 6.00, SE = .07, z = 6.12, p < .001$; Pooled control: $M = 5.63, SE = .06, z = 10.82, p < .001$). We again did not expect team orientation to vary by Condition ($F = .71, p = .492$).

Balancing the Firefighter Prototype. We first took advantage of our larger data set and longer list of traits to produce a descriptive illustration of how prototype inversion generated a balanced prototype. To do this, we conducted post-hoc analyses examining the relationship between the trait masculinity ratings that we collected in Pilot Study 1 (1 = extremely feminine, 7 = extremely masculine; see Appendix) and the trait importance ratings from this study. We ran a regression analysis predicting trait importance with trait masculinity, a dummy-coded Condition variable (Inverted condition as the baseline), and the interaction between the two as predictors. We observed significant interactions between trait masculinity and both the Traditional condition ($B = .33$, $SE = .16$, $p = .041$) and the Pooled control condition ($B = .44$, $SE = .16$, $p = .007$) dummies. The simple slopes indicate a positive, significant relationship between trait masculinity and trait importance in the Traditional ($B = .49$, $SE = .11$, $p < .001$) and Pooled Control ($B = .61$, $SE = .11$, $p < .001$) conditions, but not in the Inverted condition ($B = .16$, $SE = .11$, $p = .149$). Importantly, we did not find a significant *negative* association between trait masculinity and importance in the Inverted condition, demonstrating that we avoided feminizing the prototype (e.g., Rudman & Glick, 2001). When plotted (Figure 1), these data clearly show that the Inverted condition produced a substantially more balanced association between trait masculinity and trait importance in the firefighter prototype than did the other conditions. Although these analyses looking at the full constellation of individual firefighter traits provide a clear visual representation of how our prototype inversion technique can be used to generate a balanced prototype, we use the scale measures of the perceived importance of feminine and masculine traits in order to test our full predicted model.

Insert Figure 1 about Here

Perceived Importance of Feminine Prototype Traits. Mirroring Study 2a, we tested if the feminine traits were perceived as more equivalently important to the prototype in the Inverted

than in the other conditions by conducting a repeated measures ANOVA on the feminine and masculine traits ratings in each condition. Although in the Inverted condition, the feminine traits ($M = 6.07, SD = .06$) were rated as significantly lower than masculine traits ($M = 6.30, SD = .06; t = 4.44, p < .001$), this difference was much smaller than in both the Traditional ($M_{feminine} = 5.48, SD = .05; M_{masculine} = 6.14, SD = .05; t = 12.86, p < .001$) and Pooled Control conditions ($M_{feminine} = 5.42, SD = .04; M_{masculine} = 6.22, SD = .04; t = 19.58, p < .001$). This indicates a similar trend as in Study 2a.

Indirect Effects. We next tested our full model, examining the indirect effect of Condition on a) the perceived abilities of and b) opposition to women firefighters through the increased valuation of feminine traits with the same analytical approach as in Study 2a.

Replicating those findings, we observed a significant total effect of Condition on perceived importance of feminine traits ($F = 41.09, p < .001, \eta^2_p = .077$), the perceived ability of women firefighters ($F = 4.12, p = .017, \eta^2_p = .008$), and on opposition to women in the fire service ($F = 3.35, p = .036, \eta^2_p = .007$) (we report planned contrasts in the Appendix). We then tested our mediation hypotheses with the same series of regression models and bootstrapped confidence intervals as in Study 2a (see Table 5). Feminine trait ratings were positively associated with the perceived ability of women firefighters ($B = .16, SE = .05, p = .001$) and were negatively associated with opposition to women in the fire service ($B = -.18, SE = .06, p = .002$).

Bootstrapped confidence intervals indicated a significant positive indirect effect of Condition on the higher perceived ability of women firefighters through the greater importance of feminine traits in the Inverted condition versus the Traditional condition [-.14 to -.02], and versus the Pooled Control condition [-.16 to -.02]. We also found a significant negative indirect effect of Condition on lower opposition to women firefighters through the greater importance of feminine traits in the Inverted condition versus both the Traditional [.02 to .16], and the Pooled Control

[.02 to .18] conditions. These results, therefore, replicate those from Study 2a with a broader set of trait importance items as the mediating mechanism.

Insert Table 5 About Here

Prototypicality Threat. We next tested whether or not prototype inversion would induce prototypicality threat. The Condition effect was not significant ($F = .56, p = .572, \eta^2_p = .001$; Inverted condition: $M = 2.26, SE = .08$; v. Traditional: $M = 2.29, SE = .08, z = .28, p = .782$; v. Pooled Control: $M = 2.36, SE = .06, z = 1.00, p = .319$). Participant gender is significant ($F = 6.39, p = .012, \eta^2_p = .006$). Consistent with our expectations, men reported higher threat ($M = 2.39, SE = .05$) than did women ($M = 2.17, SE = .07$). In post-hoc analyses we confirmed that gender did not interact with the Condition variable ($F = .04, p = .957, \eta^2_p = .00$), meaning that men did not perceive more threat specifically in the Inverted prototype condition than did women. This suggests that the prototype inversion manipulation did not stimulate prototypicality threat among any participants.

Prototype Ambiguity. Next, we tested if the prototype inversion manipulation increased prototype ambiguity. Here, the Condition effect was significant ($F = 6.91, p < .001, \eta^2_p = .014$), with prototype ambiguity lower in the Inverted and Traditional conditions than in the Pooled Control condition (Inverted: $M = 2.74, SE = .07$; v. Traditional: $M = 2.86, SE = .07, z = 1.22, p = .221$; v. Pooled Control: $M = 3.06, SE = .06, z = 3.59, p < .001$). Not only did the prototype inversion manipulation not create prototype ambiguity, but it also reduced it compared to the baseline. We speculate that ambiguity was lower in both the Traditional and Inverted conditions than in the Pooled Control condition because more concrete information about the firefighter prototype was provided.

Legitimacy of Inverted Prototype. Finally, we examined the effect of our manipulations on the perceived legitimacy of the inverted prototype. The Condition effect was significant ($F =$

43.31, $p < .001$, $\eta^2_p = .080$), with the greatest perceived legitimacy in the Inverted condition ($M = 5.81$, $SE = .09$; v. Traditional: $M = 4.69$, $SE = .08$, $z = 9.24$, $p < .001$; v. Pooled Control: $M = 5.13$, $SE = .07$, $z = 6.15$, $p < .001$). Thus, as anticipated, people in the Inverted condition saw the prototype as legitimate, and no less so than in either of the other conditions.

Study 2b Discussion

Study 2b replicated and extended the findings from Study 2a. The Inverted condition was associated with the greatest valuation of feminine traits in the firefighter prototype. This increase in the perceived relative essentiality of feminine traits was, in turn, positively associated with greater perceived ability of women firefighters and lower opposition to women in the fire service. Furthermore, prototype inversion proved legitimate and did not trigger prototypicality threat (Danbold & Huo, 2015) or prototype ambiguity (Bartel & Wiesenfeld, 2013). Finally, we directly documented that the Inverted condition produced a more balanced prototype than the other conditions by significantly weakening the association between trait masculinity and trait importance. Thus, prototype inversion seems to have effectively improved perceptions of female firefighters without incurring the potential risks of redefining the firefighter prototype.

GENERAL DISCUSSION

At first glance, the extremely low representation of women in the fire service may seem like a paradox. Firefighters universally agree with the importance of traits like compassion and warmth, which, as stereotypically feminine characteristics, should make the fire service a natural fit for many women. However, as we learned through an inductive qualitative study of the firefighter prototype, fit in the fire service is evaluated first and foremost against masculine traits, like physical strength. Not only are women perceived as lacking in fit on those masculine traits, but also their potential contributions in terms of social and emotional skills are devalued and actively thwarted.

Nonetheless, through both field and laboratory experiments in the context of firefighting, we determined that it is possible to recalibrate the professional prototype such that legitimately important feminine traits are highlighted, while still affirming masculine traits. By using a prototype inversion technique, we were able to generate a balanced articulation of the firefighter prototype in which masculine and feminine traits were recognized as equivalently essential to professional success. This balanced prototype was perceived as legitimate, and both increased the perceived ability of women firefighters and decreased opposition to women in the fire service. Additionally, these positive results emerged without increasing levels of costly prototype ambiguity or prototypicality threat.

Implications for Theory

This research calls for a shift in the way prototypes are conceptualized. Rather than thinking reductively about how a prototype is associated with a group or not (e.g., saying the prototypical business leader is a man), it is important to recognize that the full set of traits that comprise a prototype rarely overlap completely with the traits that are stereotypically associated with any single group. By emphasizing the importance of the full constellation of traits that are essential to a category, prototypicality cannot be conflated with stereotypes of one demographic group more easily than another. This allows us to imagine prototypes as a balanced constellation of traits, eliminating the relative advantage and disadvantage that generally exists when prototypicality is weighted in favor of one group over another.

This work also reveals important insights into how prototypes are created and updated given exposure to new information. We found that if one wants to achieve a balanced prototype, it is not enough to simply articulate that prototype (e.g., to say that feminine and masculine traits are equally important). Rather, one must recognize that some traits are heavily devalued in the existing prototype (e.g., feminine traits in a masculine prototype) and counterweight that by

placing increased emphasis on these traits. Applying this prototype inversion technique, therefore, may create more inclusive prototypes in which it is easier to see more kinds of groups as fitting and succeeding.

This research also speaks to the broader issue of “person-situation fit” (Lewin, 1935) and how to improve instances of poor fit. Social stereotypes are generally so deeply entrenched that strategies focusing on the “person” (e.g., asking individuals to act counter-stereotypically) can generate more problems than they solve. Our research focuses on the “situation,” and suggests that people are receptive to concrete redefinitions of the prototype that retain its legitimately important characteristics. Doing so causes individuals to reevaluate and acknowledge women’s fit with those professions. This differs from prior approaches to reframing category prototypes. Unlike previous efforts to prevent any one group from claiming prototypicality by creating an obscure and indefinable prototype (Alexandre, Waldzus, & Wenzel, 2016; Waldzus, Mummendey, Wenzel, & Weber, 2003), including a concrete definition of the inverted category prototype seems to avoid inducing prototype ambiguity (Bartel & Wiesenfeld, 2013). Furthermore, because prototype inversion accomplishes this while still affirming the importance of masculine characteristics, it preempts men from experiencing prototypicality threat, which has been shown to be a driver of resistance to efforts to feminize previously masculine professions (Danbold & Huo, 2017).

Limitations

Although our results clearly demonstrate the effectiveness of prototype inversion in this context of firefighters, it is important to acknowledge some limitations of our findings. The first is that the specific prototype inversion manipulation we developed here worked best when participants fully internalized the manipulation. Because rates of reactance to our manipulation from firefighters were comparable in the Inverted and Traditional conditions, and virtually

nonexistent in the non-firefighter sample, this pushback likely owes more to the perception that the information provided came from outsiders than reactions to its actual content, as anticipated by the firefighters we worked with. Because this research aimed to test the predictions of internalizing an inverted prototype framing, rather than simply testing the efficacy of our specific manipulation, we encourage researchers to identify ways to optimize the delivery of prototype inversion interventions that reduce reactance to manipulations in future work. We speculate that interventions based on prototype inversion are likely to be most effective when they are delivered by in-group members and when they encourage the validity of an inverted prototype through discussion rather than lecture. To this point, it is worth acknowledging that we only tested a single video-based prototype inversion manipulation, and that amplifying the perceived importance of previously devalued traits, while affirming those already valued, can take many forms (e.g., in-person trainings, organizational cultural values, job recruitment materials, etc.). Clearly, this is an area ripe for future research.

A second limitation of this work is our focus on attitudinal outcome variables measured shortly after our manipulation, and thus may be susceptible to demand characteristics or desirability bias. Despite this possibility, we observed significant gender bias in favor of male firefighters across our samples, notably in control conditions when no manipulation was present, suggesting we were able to measure people's genuine attitudes. Additionally, our measure of the perceived ability of women firefighters used a hypothetical situation and was, therefore, consequence-free for participants. Our measures of opposition to women in the fire service, however, were based on the articulated rationale for exclusionary behaviors towards women in the fire departments that participated in Study 1. Responses on those items, therefore, are more likely associated with downstream consequential behaviors, but we were unable to test that directly in these studies. Although these variables were appropriate first steps for testing our

theory, it should be a priority of subsequent work to examine the effects of prototype inversion on more behavioral consequences and longer-term outcomes for hiring and promotion of women.

One potential criticism of prototype inversion is that emphasizing commonly held *positive* stereotypes about groups (e.g., women are compassionate, men are strong, etc.) could also reinforce *negative* stereotypes associated with those groups (Siy & Cheryan, 2016).

Prototype inversion's utilization of positively stereotyped traits also raises the question of how this technique would apply to situations in which the underrepresented group has few positive characteristics associated with it, or when their associated characteristics aren't valued within the relevant profession. Furthermore, it could especially penalize members of underrepresented groups who do not demonstrate the newly-emphasized trait that is stereotypically associated with their groups (e.g., women who are low in compassion). First, it is important to stress that we do not endorse gender stereotyping inside or outside the context of prototype inversion. Additionally, we argue that the accomplishments of prototype inversion may not be limited to seeing one individual group (e.g., women) in a more positive light. Rather, by decoupling the exclusive association between a category prototype and a single group, prototype inversion may increase the perceived prototypicality of members of the dominant group who are high in counter-stereotypical characteristics (e.g., compassionate male firefighters). It may also have spillover effects, giving members of other groups typically not associated with the firefighter identity (e.g., gay men) greater opportunity to represent an ideal firefighter.

We also focused on the demographic dimension of gender, even though ethnic minorities are also underrepresented in the fire service (National Fire Protection Association, 2013), and the history of the fire service has been marked with many instances of racial bias (Hulett, Bendick, Thomas, & Moccio, 2008). Our interviews did not indicate that this bias was particularly due to a perceived incompatibility between ethnic minorities and the firefighter prototype, nor did we

observe racial bias towards Black Americans in our measures of perceived ability. However, more work is needed to determine if any kind of racial bias in the fire service is sensitive to some kind of prototype manipulation. An interesting extension may be to look at the consequences of prototype inversion for groups like Asian men or lesbian women, who are viewed as having stereotypically incompatible identities (i.e., Asian Americans are stereotypically viewed as feminine and lesbians are stereotypically viewed as masculine; Blashill & Powlishta, 2009; Hall, Galinsky, & Phillips, 2015; Johnson, Freeman, & Pauker, 2012). These groups, which are often uniquely disadvantaged by this perceived incompatibility between their identities (Purdie-Vaughns & Eibach, 2008), may, in fact, find these disadvantages diminished when the prototypes they are being evaluated against become more balanced.

Implications for Managerial Diversity and Inclusion

Of course, women's lack of fit in masculine professions is only one of the barriers in the way of fully equal representation and inclusion. Even if a balanced prototype is successfully created and men aren't concerned about losing their prototypicality, they may still hold other concerns regarding competition over jobs and resources (Stephan, Ybarra, & Rios Morrison, 2009). As such, prototype balancing should be viewed as part of a wider set of techniques to address the issues of gender and diversity in the workplace. However, given that explicit diversity training and diversification initiatives are often seen as superficial or forced and rarely achieve all their stated objectives (Dobbin & Kalev, 2013; Kaiser, Major, Jurcevic, Dover, Brady, & Shapiro, 2013; Kalev, Dobbin, & Kelly, 2006), prototype inversion may have the advantage of achieving these goals without being explicitly linked to diversity. Framed as an expression of what traits the profession values, a prototype inversion strategy, delivered by trusted organization members, can potentially avoid the reactance that challenges so many other approaches to gender and diversity.

Applied to the fire service, we hope this intervention will help firefighters create more inclusive departments in which everyone would be proud to have their daughters serve.

By demonstrating the effectiveness of prototype inversion for increasing the perceived fit of women in the extreme masculine context of the U.S. fire service, we infer that it could similarly balance the prototypes of many other masculine professional contexts. Companies that are able to effectively communicate a balanced professional prototype to the general public would potentially enhance women's perceived sense of fit between themselves and positions within that profession, thereby increasing the recruitment of women. Within such companies, trainings and norm-setting that promotes a balanced prototype could increase the degree to which women are recognized as capable and contributing members of the profession, supporting their advancement and retention. Additionally, the theory and methods of prototype inversion may inspire new strategies for increasing representations of ethnic minorities and other traditionally underrepresented groups and addressing the challenges surrounding gender, diversity, and person-situation fit more broadly.

REFERENCES

- Alexandre, JD, Waldzus, S, Wenzel, M (2016) Complex inclusive categories of positive and negative valence and prototypicality claims in asymmetric intergroup relations. *British Journal of Social Psychology*, 55(3):457-483.
- Bartel, CA, Wiesenfeld, BM (2013) The social negotiation of group prototype ambiguity in dynamic organizational contexts. *Academy of Management Review* 38(4):503-524.
- Berdahl, JL (2007). Harassment based on sex: Protecting social status in the context of gender hierarchy. *Academy of Management Review* 32(2):641-658.
- Bertrand, M, Mullainathan, S (2003) Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *National Bureau of Economic Research*.
- Blashill, AJ, Powlishta, KK (2009) Gay stereotypes: The use of sexual orientation as a cue for gender-related attributes. *Sex Roles*, 61(11-12):783-793.
- Brescoll, VL, Dawson, E, Uhlmann, EL (2010) Hard won and easily lost the fragile status of leaders in gender-stereotype-incongruent occupations. *Psychological Science* 21(11):1640-1642.
- Bureau of Labor Statistics (2015) The Economics Daily, Women's earnings compared to men's earnings in 2014. Retrieved from <http://www.bls.gov/opub/ed/2015/womens-earnings-compared-to-mens-earnings-in-2014.htm>.
- Cable, DM, DeRue, DS (2002) The convergent and discriminant validity of subjective fit perceptions. *Journal of Applied Psychology* 87(5):875.
- Cejka, MA, Eagly, AH (1999) Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin* 25(4):413-423.
- Chambers, DW (1983) Stereotypic images of the scientist: The Draw-a-Scientist Test. *Science Education* 67(2):255-265.
- Chatman, JA, Polzer, JT, Barsade, SG, Neale, MA (1998) Being different yet feeling similar: The influence of demographic composition and organizational culture on work processes and outcomes. *Administrative Science Quarterly* 41:423.
- Cheryan, S, Plaut, VC, Davies, PG, Steele, CM (2009) Ambient belonging: how stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology* 97(6):1045-1060.
- Cheryan, S, Plaut, VC, Handron, C, Hudson, L (2013) The stereotypical computer scientist: Gendered media representations as a barrier to inclusion for women. *Sex Roles* 69: 58-71.

- Chetkovich, CA (1997) *Real heat: Gender and race in the urban fire service*. (Rutgers University Press.)
- Danbold, F, Huo, YJ (2015) No Longer “All-American”? Whites’ Defensive Reactions to Their Numerical Decline. *Social Psychological and Personality Science* 6(2):210-218.
- Danbold, F, Huo, YJ (2017) Men’s Defense of Their Prototypicality Undermines the Success of Women in STEM Initiatives. *Journal of Experimental Social Psychology*.
- Delsohn, S (1996) *The Fire Inside: Firefighters Talk about their Lives*. (New York: Harper Collins).
- Derks, B, Ellemers, N, van Laar, C, De Groot, K (2011) Do sexist organizational cultures create the Queen Bee? *British Journal of Social Psychology* 50(3):519-535.
- Desmond, M (2006) Becoming a firefighter. *Ethnography* 7:387-421.
- Diekmann, AB, Brown, ER, Johnston, AM, & Clark, EK (2010) Seeking congruity between goals and roles a new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science* 21(8):1051-1057.
- Dobbin, F, Kalev, A (2013) The Origins and Effects of Corporate Diversity Programs, 253-281, in *Oxford Handbook of Diversity and Work*, QM Roberson ed. (New York: Oxford University Press)
- DeRue, DS, & Ashford, SJ (2010) Who will lead and who will follow? A social process of leadership identity construction in organizations. *Academy of Management Review*, 35(4):627-647.
- Eagly, AH, Karau, SJ (2002) Role congruity theory of prejudice toward female leaders. *Psychological Review* 109(3):573.
- Ehrlinger, J, Plant, EA, Hartwig, MK, Vossen, JJ,, Columb, CJ, Brewer, LE (2018) Do Gender Differences in Perceived Prototypical Computer Scientists and Engineers Contribute to Gender Gaps in Computer Science and Engineering?. *Sex Roles*, 78(1-2):40-51.
- Eisenhardt, KM (1989) Agency theory: An assessment and review. *Academy of Management Review*, 14(1):57-74.
- Ely, RJ, Meyerson, DE (2010) An organizational approach to undoing gender: The unlikely case of offshore oil platforms. *Research in Organizational Behavior* 30:3-34.
- Ensmenger, NL (2012). *The computer boys take over: Computers, programmers, and the politics of technical expertise* (MIT Press)
- Fiske, ST, Cuddy, AJ, 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.

- Galinsky, AD, Todd, AR, Homan, AC, Phillips, KW, Apfelbaum, EP, Sasaki, SJ, Richeson, JA, Olayon, JB, Maddux, WW (2015) Maximizing the Gains and Minimizing the Pains of Diversity A Policy Perspective. *Perspectives on Psychological Science* 10(6):742-748.
- Gaucher, D, Friesen, J, Kay, AC, (2011) Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology*, 101(1):109.
- Hall, EV, Galinsky, AD, Phillips, KW (2015) Gender Profiling A Gendered Race Perspective on Person–Position Fit. *Personality and Social Psychology Bulletin* 41(6):853.
- Hardison, CM, Lim, N, Keller, KM, Marquis, JP, Payne, LA, Bozick, R, .Mariano, LT, Mauro, JA, Miyashiro, L, Oak, G, Saum-Manning, L (2015) Recommendations for Improving the Recruiting and Hiring of Los Angeles Firefighters. *Rand Corporation*.
- Haynes, HJG, Stein, GP (2014) US Fire Department Profile 2013. *National Fire Protection Association Fire Analysis and Research Division*.
- Heilman, ME (1983) Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior* 5:269-298.
- Heilman, ME (2001) Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues* 57(4):657-674.
- Heilman, ME, Haynes, MC (2005) No credit where credit is due: attributional rationalization of women's success in male-female teams. *Journal of Applied Psychology* 90(5):905.
- Heilman, ME, Wallen, AS, Fuchs, D, Tamkins, MM (2004) Penalties for success: reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology* 89(3):416.
- Helgesen, S (1990) *The female advantage: Women's ways of leadership*. (New York: Doubleday)
- Herring, C (2009) Does diversity pay?: Race, gender, and the business case for diversity. *American Sociological Review* 74(2):208-224.
- Hulett, DM, Bendick, M, Thomas, SY, Moccio, F (2008) A national report card on women in firefighting. *Madison, WI: International Association of Women in Fire & Emergency Services*.
- Inesi, ME, Cable, DM (2015) When accomplishments come back to haunt you: The negative effect of competence signals on women's performance evaluations. *Personnel Psychology* 68(3):615-657.
- Johnson, C, Dowd, TJ, Ridgeway, CL (2006) Legitimacy as a social process. *Annual Review of Sociology* 32:53-78.

- Johnson, KL, Freeman, JB, Pauker, K (2012) Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology*, 102(1):116.
- Joshi, KD, Trauth, E, Kvasny, L McPherson, S (2013) Exploring the Differences Among IT Majors and Non-majors: Modeling the Effects of Gender Role Congruity, Individual Identity, and IT Self efficacy on IT Career Choices. *Proceedings of the Thirty-Fourth International Conference on Information Systems* (Milan, Italy, December).
- Kaiser, CR, Major, B, Jurcevic, I, Dover, TL, Brady, LM, Shapiro, JR (2013) Presumed fair: Ironic effects of organizational diversity structures. *Journal of Personality and Social Psychology* 104(3):504.
- Kalev, A, Dobbin, F, Kelly, E (2006) Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, 71(4):589-617.
- Kray, LJ, Galinsky, AD, Thompson, L (2002) Reversing the gender gap in negotiations: An exploration of stereotype regeneration. *Organizational Behavior and Human Decision Processes*, 87(2):386-409.
- Koenig, AM, Eagly, AH, Mitchell, AA, Ristikari, T (2011) Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin* 137(4):616.
- Kristof, AL (1996) Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology* 49(1):1-49.
- Kristof-Brown, AL, Zimmerman, RD, Johnson, EC (2005) Consequences of individual's fit at work: A meta-analysis of person-job, person-organization, person-group, and person-supervisor fit. *Personnel Psychology* 58(2):281-342.
- Lai CK, Marini M, Lehr SA, Cerruti C, Shin JE, Joy-Gaba JA, . . . Nosek BA (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General* 143:1765-1785
- Legault, L, Gutsell, J, Inzlicht, M (2011) Ironic effects of antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science* 22:1472-1477
- Lenton, AP, Bruder, M, Sedikides, C (2009) A meta-analysis on the malleability of automatic gender stereotypes. *Psychology of Women Quarterly*, 33(2):183-196.
- Lewin, K (1935) *A Dynamic Theory of Personality* (New York, NY, US: McGraw-Hill)
- Lips, HM (2013) The gender pay gap: Challenging the rationalizations. Perceived equity, discrimination, and the limits of human capital models. *Sex Roles* 68(3-4):169-185.

- Lord, RG, Brown, DJ, Harvey, JL, & Hall, RJ (2001) Contextual constraints on prototype generation and their multilevel consequences for leadership perceptions. *The Leadership Quarterly*, 12(3):311-338.
- Lyness, KS, & Heilman, ME (2006). When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. *Journal of Applied Psychology*, 91(4):777.
- McClellan E, Martin SR, Emich KJ, Woodruff T (2017) The social consequences of voice: An examination of voice type and gender on status and subsequent leader emergence. *Academy of Management Journal*.
- Miles, MB Huberman, AM, (1994) *Qualitative data analysis: A sourcebook*. (Beverly Hills: Sage Publications.)
- Muchinsky, PM Monahan, CJ (1987) What is person-environment congruence? Supplementary versus complementary models of fit. *Journal of Vocational Behavior* 31(3):268-277.
- Mummendey, A, Wenzel, M (1999) Social discrimination and tolerance in intergroup relations: Reactions to intergroup difference. *Personality and Social Psychology Review* 3(2):158-174.
- National Fire Protection Association (2013) Firefighting Occupations by Women and Race. Retrieved from <http://www.nfpa.org/research/reports-and-statistics/the-fire-service/administration/firefighting-occupations-by-women-and-race>.
- Oakes, PJ, Haslam, SA, Turner, JC (1994) *Stereotyping and social reality*. (Blackwell Publishing)
- O'Neill, OA, Rothbard, N (2017) Is Love All You Need? The Effects of Emotional Culture, Suppression, and Work-Family Conflict on Firefighter Risk Taking and Health. *Academy of Management Journal*.
- Plaut, VC, Thomas, KM, Goren, MJ (2009) Is multiculturalism or color blindness better for minorities? *Psychological Science*, 20(4):444-446.
- Pratt, MG (2017) *Trust Under Fire: Trustworthiness as Temporal-Spatial Sensemaking in Early Work Relationships Among Firefighters* (Boston College.)
- Prentice, DA, & Carranza, E (2002) What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly* 26(4):269-281.
- Purdie-Vaughns, V, Eibach, RP (2008) Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles*, 59(5-6):377-391.
- Richards, Z, Hewstone, M (2001). Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review* 5(1),52-73.

- Ridgeway, CL (1991) The social construction of status value: Gender and other nominal characteristics. *Social Forces* 70(2):367-386.
- Ridgeway, CL, Li, YE, Erickson, KG, Backor, K, Tinkler, JE (2009) How Easily Does a Social Difference Become a Status Distinction? Gender Matters. *American Sociological Review*, 74(1):44-62.
- Rosch, E (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104(3):192.
- Rosener, JB (1995) *America's competitive secret: Utilizing women as a management strategy*. (New York: Oxford University Press.)
- Rucker, DD, Preacher, KJ, Tormala, ZL, & Petty, RE (2011) Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass* 5(6):359-371.
- Rudman, LA, Fairchild, K (2004) Reactions to counterstereotypic behavior: the role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, 87(2):157-176.
- Rudman, LA, Glick, P (1999) Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle-managers. *Journal of Personality and Social Psychology*, 77:1004–1010.
- Rudman, LA, Glick, P (2001) Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues* 57(4):743-762.
- Ryan, MK, Haslam, SA (2007) “The glass cliff: Exploring the dynamics surrounding the appointment of women to precarious leadership positions.” *Academy of Management Review* 32(2):549-572.
- Sczesny, S, Formanowicz, M, Moser, F (2016) Can gender-fair language reduce gender stereotyping and discrimination? *Frontiers in Psychology* 7.
- Siy, JO, Cheryan, S (2016) Prejudice Masquerading as Praise The Negative Echo of Positive Stereotypes. *Personality and Social Psychology Bulletin*, 42(7):941-954.
- Smith, ER (1998) Mental representation and memory. In DT Gilbert, ST Fiske, & G Lindzey eds. *Handbook of Social Psychology, Vol. 1* (New York: Oxford University Press).
- Smith, ER, Semin, GR (2007) Situated Social Cognition. *Current Directions in Psychological Science*, 16(3):132-135.
- Stephan, WG, Ybarra, O, RRios Morrison, K (2009) Intergroup Threat Theory. In TD Nelson ed. *Handbook of Prejudice, Stereotyping, and Discrimination* (New York: Psychology Press)

- Strauss A, Corbin JM (1990) *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. (Sage, Newbury Park, California.)
- Sy, T, Shore, LM, Strauss, J, Shore, TH, Tram, S, Whiteley, P, Ikeda-Muromachi, K (2010) Leadership perceptions as a function of race–occupation fit: The case of Asian Americans. *Journal of Applied Psychology* 95(5):902.
- Tost, LP (2011) An Integrative Model of Legitimacy Judgments. *Academy of Management Review* 36(4):686-710.
- Turner, JC (1987) A self-categorisation theory. In JC Turner, MA Hogg, PJ Oakes, SD Reicher, & MS Wetherell eds. *Rediscovering the social group: A self-categorisation theory*. (Oxford, UK: Blackwell). 42–67.
- Waldzus, S, Mummendey, A, Wenzel, M, Weber, U (2003) Towards tolerance: Representations of superordinate categories and perceived ingroup prototypicality. *Journal of Experimental Social Psychology*, 39(1):31-47.
- Wenzel, M, Mummendey, A, Waldzus, S., (2007) Superordinate identities and intergroup conflict: The ingroup projection model. *European Review of Social Psychology* 18(1):331-372.
- Wessel, JL, Hagiwara, N, Ryan, AM, Kermond, CM (2014) Should Women Applicants “Man Up” for Traditionally Masculine Fields? Effectiveness of Two Verbal Identity Management Strategies. *Psychology of Women Quarterly*.
- Vandello, JA, & Bosson, JK (2013) Hard won and easily lost: A review and synthesis of theory and research on precarious manhood. *Psychology of Men & Masculinity*, 14(2):101.
- Yin, RK (2013) *Case study research: Design and methods*. (Sage Publications.)
- Zhao, X, Lynch Jr, JG, Chen, Q (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research* 37(2):197-206.

Table 1: Traits of the firefighter prototype identified by active-duty fire service personnel, with representative quotations. Order of trait listing does not indicate their frequency. Respondent gender is indicated in parentheses (F = female, M = male).

| Trait | Representative Quote |
|--|---|
| Physical and mental strength | “You have to be physically ready to do this job.” (F) |
| Compassion | “You have to have the certain soul where you want to help people.” (M) |
| Initiative | “Always striving to get better, being non-complacent.” (F) |
| Problem-solving skills/mechanical aptitude | “Being a fast learner, being able to retain information about all sorts of different things.” (F) |
| Decisive decision-making | “Being able to analyze the problem and figure out a course of action.” (F) |
| Team orientation | “You need to be able to get along because we’re living together.” (M) |
| Selflessness | “You need to be humble.” (M) |
| Positive public demeanor | “You need to be aware that you are going into people’s homes.” (M) |
| Self-discipline | “You have to take the job seriously.” (M) |
| Warmth | “Firefighters show up to take care of you.” (M) |

Table 2: Study 2a descriptive statistics and inter-item correlations.

| | Fem. Traits | Masc. Traits | Neutral Trait | Women FF Ability | Men FF Ability | Opp. to Women FF | Department | Gender | Ethnicity | Rank |
|------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------|-------------------|---------------------|---------------------------|
| | <i>M = 6.03, SD = 1.08</i> | <i>M = 6.20, SD = 0.87</i> | <i>M = 6.14, SD = 1.28</i> | <i>M = 4.59, SD = 1.53</i> | <i>M = 5.22, SD = 1.23</i> | <i>M = 4.12, SD = 1.71</i> | <i>90.74% LACoFD</i> | <i>96.19% Men</i> | <i>57.77% White</i> | <i>36.89% Firefighter</i> |
| Fem. Traits | - | | | | | | | | | |
| Masc. Traits | .43** | - | | | | | | | | |
| Neutral Trait | .50** | .37** | - | | | | | | | |
| Women FF Ability | .14** | -.02 | .07 | - | | | | | | |
| Men FF Ability | -.00 | -.01 | .03 | .56** | - | | | | | |
| Opp. to Women FF | -.13* | .08 | -.01 | -.39** | -.14** | - | | | | |
| Department | .02 | -.01 | -.02 | .07 | .12* | .04 | - | | | |
| Gender | .04 | .03 | -.04 | .16** | .11* | -.20** | .06 | - | | |
| Ethnicity | -.11* | -.10 | -.07 | .07 | .08 | .11* | -.06 | .09 | - | |
| Rank | -.04 | .07 | -.05 | -.03 | -.02 | .15** | .03 | .11* | .05 | - |

Note: Fem. = Feminine; Masc. = Masculine; FF = Firefighter; Opp. = Opposition, Gender is coded 0 = Woman, 1 = Man; Ethnicity is coded 0 = Non-White, 1 = White.

† $p < .10$; * $p < .05$; ** $p < .01$

Table 3: Study 2a test of the indirect effect of inversion manipulation on the perceived ability of women firefighters and opposition to women firefighters, through feminine trait ratings. Regression results with unstandardized coefficients. Key path coefficients are bolded.

| | DV = Feminine Trait Ratings | DV = Perceived Ability of Women Firefighters | DV = Opposition to Women in the Fire Service |
|---|-----------------------------|--|--|
| Traditional Condition | -.34* | -.25 | .17 |
| Pooled Control Condition | -.23* | -.33* | .35 [†] |
| Feminine Trait Ratings | | .26* | -.35** |
| Masculine Trait Ratings | .34** | -.14 | .31 |
| Neutral Trait Rating | .33** | .01 | .05 |
| Perceived Ability of Men Firefighters | | .69** | |
| Department | .13 | -.04 | .23 |
| Gender | .33 | .75** | -1.45** |
| Ethnicity | -.13 | .12 | .43** |
| Rank | -.03 | .00 | .13* |
| <i>Constant</i> | <i>1.82</i> | <i>-.31</i> | <i>4.25</i> |
| <i>R</i> ² | .33** | .36** | .14** |
| Bootstrapped Bias-Corrected 95% Confidence Intervals of Indirect Effect of Prototype Inversion Manipulation on Perceived Ability of Women Firefighters through Feminine Trait Ratings Inverted vs. Traditional: [-.25, -.01] Inverted vs. Control: [-.16, -.00] | | | |
| Bootstrapped Bias-Corrected 95% Confidence Intervals of Indirect Effect of Prototype Inversion Manipulation on Opposition to Women in the Fire Service through Feminine Trait Ratings Inverted vs. Traditional: [.04, .26] Inverted vs. Control: [.01, .21] | | | |

Note: Inverted condition is the excluded baseline; Gender is coded 0 = Woman, 1 = Man; Ethnicity is coded 0 = Non-White, 1 = White; [†] $p < .10$; * $p < .05$; ** $p < .01$

Table 4: Study 2b descriptive statistics and inter-item correlations.

| | Fem. Traits | Masc. Traits | Neutral Trait | Women FF Ability | Men FF Ability | Opp. to Women FF | Inv. Prototype Legit. | Proto. Threat | Prototype Ambiguity | Gender | Ethnicity |
|-----------------------|--------------------------------------|-------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|---------------|-----------------|
| | <i>M</i> = 5.62, <i>SD</i> = 1.08 | <i>M</i> = 6.22, <i>SD</i> = .80 | <i>M</i> = 6.04, <i>SD</i> = 1.23 | <i>M</i> = 4.37, <i>SD</i> = 1.38 | <i>M</i> = 5.35, <i>SD</i> = 1.05 | <i>M</i> = 3.20, <i>SD</i> = 1.49 | <i>M</i> = 5.19, <i>SD</i> = 1.49 | <i>M</i> = 2.31, <i>SD</i> = 1.31 | <i>M</i> = 2.92, <i>SD</i> = 1.18 | 67.04% Men | 76.68% White |
| Fem Traits | - | | | | | | | | | | |
| Masc Traits | .59** | - | | | | | | | | | |
| Neutral Trait | .54** | .64** | - | | | | | | | | |
| Women FF Ability | .15** | .09** | .05 | - | | | | | | | |
| Men FF Ability | .25** | .35** | .27** | .39** | - | | | | | | |
| Opp. to Women FF | -.23** | -.21** | -.18** | -.53** | -.14** | - | | | | | |
| Inv. Prototype Legit. | .35** | .28** | .24** | .11** | .15** | -.24** | - | | | | |
| Proto. Threat | -.15** | -.34** | -.24** | -.21** | -.24** | .42** | .26** | - | | | |
| Prototype Ambiguity | -.22** | -.33** | -.24** | -.10** | -.33** | .10** | -.24** | .42** | - | | |
| Gender | -.19** | -.17** | -.18** | -.13** | -.07* | .18** | -.13** | .08** | .08** | - | |
| Ethnicity | -.08** | -.05 | -.07* | .05 | .04 | .03 | -.04 | -.07 | -.04 | -.02 | - |

Note: Fem. = Feminine; Masc. = Masculine; FF = Firefighter; Opp. = Opposition, Inv. = Inverted, Legit. = Legitimacy, Proto. = Prototypicality, Gender is coded 0 = Woman, 1 = Man; Ethnicity is coded 0 = Non-White, 1 = White.

† $p < .10$; * $p < .05$; ** $p < .01$

Table 5: Study 2b test of the indirect effect of inversion manipulation on the perceived ability of women firefighters through feminine trait ratings, regression results with unstandardized coefficients. Key path coefficients are bolded.

| | DV = Feminine Trait Ratings | DV = Perceived Ability of Women Firefighters | DV = Opposition to Women in the Fire Service |
|---|-----------------------------|--|--|
| Traditional Condition | -.46** | -.26* | .19 |
| Pooled Control Condition | -.55** | -.03 | .05 |
| Feminine Trait Ratings | | .16** | -.18** |
| Masculine Trait Ratings | .55** | .16** | -.17* |
| Neutral Trait Rating | .22** | -.10 | -.03 |
| Perceived Ability of Men Firefighters | | .54** | |
| Gender | -.17 | -.33** | .45** |
| Ethnicity | -.14 | .12 | .04 |
| <i>Constant</i> | <i>1.47**</i> | <i>2.44**</i> | <i>5.04**</i> |
| <i>R</i> ² | .45** | .19** | .08** |
| Bootstrapped Bias-Corrected 95% Confidence Intervals of Indirect Effect of Prototype Inversion Manipulation on Perceived Ability of Women Firefighters through Feminine Trait Ratings | | | |
| Inverted vs. Traditional: [-.14, -.02] | | Inverted vs. Control: [-.16, -.02] | |
| Bootstrapped Bias-Corrected 95% Confidence Intervals of Indirect Effect of Prototype Inversion Manipulation on Opposition to Women in the Fire Service through Feminine Trait Ratings | | | |
| Inverted vs. Traditional: [.02, .16] | | Inverted vs. Control: [.02, .18] | |

Note: Inverted condition is the excluded baseline; Gender is coded 0 = Woman, 1 = Man; Ethnicity is coded 0 = Non-White, 1 = White; † $p < .10$; * $p < .05$; ** $p < .01$

Figure 1: Study 2b post-hoc analysis of the balance of prototype traits.

