The (Re)Production of Inequality in Evaluations: A Unifying Framework Outlining the Drivers of Gender and Racial Differences in Evaluative Outcomes[*]

Mabel Abraham

Tristan L. Botelho

Gabrielle Lamont-Dobbin

Abstract

Evaluations play a critical role in the allocation of resources and opportunities. Although evaluation systems are a cornerstone of organizational and market processes, they often reinforce social and economic inequalities. The body of organizational research on inequality and evaluations is extensive, but it is also fragmented, siloed within specific contexts and types of evaluations (e.g., hiring, performance). As a result, we currently lack a systemic understanding of the conditions under which inequalities emerge. This paper provides a unifying framework to identify how gender and racial inequality is produced and reproduced in evaluations across professional contexts (e.g., digital platforms, entrepreneurship, traditional employment). Our framework categorizes the drivers of inequality into three main areas: prevailing beliefs in evaluative contexts, the design and structure of evaluation processes, and the characteristics of evaluators. Our approach not only sheds light on the common processes that exacerbate inequality but also underscores why an integrative framework is critical for both theoretical advancement and enacting effective reforms.

Keywords: evaluations, gender, inequality, organizations, professionals, race

Introduction

Evaluations—that is, the processes determining the value or quality of a candidate's contributions, performance, or work—are fundamental to professional contexts.[1] They govern the disbursal of resources such as employment opportunities, compensation, and recognition (Botelho, 2024; Botelho & Abraham, 2017; Bowers & Prato, 2018; Rosette et al., 2008). Whether it is in a firm, where performance evaluations dictate promotions and compensation or in academia and entrepreneurship, where review systems determine who receives funding, the impact of evaluative systems is far-reaching (Abraham, 2017; Beckman & Phillips, 2005; Bohnet et al., 2016; Lee & Huang, 2018; Rivera & Tilcsik, 2019). Even in the digital sphere, evaluations via platforms are consequential, determining the success of customer-facing businesses, entrepreneurial ventures, and gig workers (Brooks et al., 2014; Leung & Koppman, 2018).

Yet, an accumulating body of research warns that evaluations are often not equitable, fortifying entrenched social and economic disparities (Valentino & Vaisey, 2022). At first blush, evaluation processes commonly appear to judge a candidate's objective performance such that final assessments reflect their capabilities and quality (Alon & Tienda, 2007; Castilla, 2008). However, factors unrelated to a candidate's capabilities can affect evaluative outcomes (Bowers & Prato, 2018; Correll et al., 2007; Rivera, 2020), especially disadvantaging historically marginalized groups (Abraham, 2020; Botelho & Gertsberg, 2022; Gorbatai et al., 2023). These inequities emerge even in purportedly merit-based and impartial evaluation systems where evaluators are guided and even incentivized to prioritize objective performance factors over demographic ones (Botelho & Abraham, 2017; Dobbin et al., 2015). Consequently, while there

---

[1] For simplicity, we will use the term "evaluator" to refer to the individual (or group) that is responsible for assessing a "candidate" (or set of candidates), which we define as the individual(s) or target(s) of the evaluation.

have been considerable advancements in educational attainment and professional opportunities for White women and racially marginalized groups over the past few decades, members of these groups continue to significantly trail behind White men in their professional achievements (England et al., 2020; Skrentny, 2014; Stainback & Tomaskovic-Devey, 2012).

Although scholars have devoted ample attention to identifying the wide-spread nature of these inequities, research at the intersection of evaluations and inequality is disconnected. Existing scholarship has documented evaluative disparities in distinct contexts and specific types of evaluations (e.g., customer ratings, investment in startups, personnel decisions). For instance, it is well established that women and racially marginalized groups experience disadvantages in receiving interviews for jobs (Kang et al., 2016; Pager & Pedulla, 2015), being promoted (DiPrete & Soule, 1988), earning top performance ratings (Castilla, 2008; Rivera & Tilcsik, 2019), and accessing funding to support their entrepreneurial ventures (see Botelho et al., 2024 for a brief review; Brooks et al., 2014; Greenberg & Mollick, 2017). These examples are concrete and informative, but the prevailing approach is atomistic, inadvertently erecting research silos which hinders our ability to identify conditions *common across* evaluative contexts that reproduce inequality. In other words, the fragmented nature of existing research leaves unclear whether and how the drivers reproducing inequality transcend contexts.

The current analysis provides a unifying framework that cuts across contexts and types of evaluations to uncover the common drivers that exacerbate gender and racial inequality across contexts and types of evaluations (see Figure 1 for an overview of the drivers and Table 1 for further detail). Given the volume of research on evaluations and inequality across disciplines, it is important to begin by outlining several scope conditions. First, the framework primarily builds on research examining inequality in evaluations and evaluative outcomes within professional

contexts, including those common both within organizations (e.g., hiring, performance reviews, salary determinations) and in the economy more generally (e.g., grant allocation, online ratings, crowdfunding, gig work). Professional contexts are defined as environments—including but not limited to formal organizational and employment contexts—where individuals engage in tasks or provide services requiring specialized knowledge or skills. Evaluations in these contexts tend to follow largely agreed-upon performance metrics and are typically conducted by individuals with a baseline level of expertise (Botelho, 2024). Second, this framework centers on the demand-driven determinants of evaluative outcomes rather than the precursors that may also contribute to unequal outcomes. Notably, it does not consider the factors affecting the formation and composition of candidate pools. Although extensive research reveals supply-side processes also contribute to inequality—often skewing the gender and racial composition of candidate pools (Abraham & Burbano, 2022; Fernandez et al., 2000; Fernandez-Mateo & Kaplan, 2018; Wald et al., 2024)—our focus is on inequality in the direct assessment of candidates included. Third, the review builds primarily on research stemming from American and, to a lesser extent, European professional contexts, and thus speaks most directly to evaluation processes in Western settings.

Fourth, the framework focuses on gender- and race-based inequality. Granted, these are far from the only bases of inequality in professional contexts; there is considerable evidence that discrimination emerges with respect to age, sexual orientation, disability, and other characteristics (Martin & North, 2022; Mishel, 2020; Rivera & Tilcsik, 2023). But gender and race have received the greatest amount of attention from organizational scholars. We also note that organizational research on racial inequality has been primarily focused on differences between White and Black individuals. We therefore have limited opportunities to discuss

whether the drivers we have identified also shape inequality among members of other racial minority groups (e.g. Asian, Latin American).

Before presenting our framework, we provide an overview of the common explanations for why gender and race generally shape evaluative outcomes. This literature is reviewed in detail elsewhere (Eagly & Koenig, 2008; Ridgeway & Correll, 2006), and thus our aim is to provide a basic overview. From there, we introduce our framework (see Figure 1), which groups drivers of inequality across types of evaluations into the following three categories: (i) How Prevailing Beliefs Affect Inequality, (ii) How the Design and Structure of Evaluation Processes Shape Inequality, and (iii) How Evaluator Characteristics Influence Inequality. For each, we also highlight high-impact research areas to further deepen our theoretical understanding of when and how evaluations perpetuate inequality. Beyond advancing research, our framework provides a strategic blueprint for how policymakers, organizational architects, and social stakeholders can reform evaluation systems to be more equitable.

Common Explanations for Why Gender and Race Shape Evaluative Outcomes

Evaluations do not occur in a vacuum; they happen within broader contexts that are defined by social norms and the hierarchical positioning of different social groups. Individuals almost immediately categorize others based on ascriptive (or demographic) characteristics, like gender and race (Zarate & Smith, 1990). Not only are these characteristics readily apparent, but they are also linked to a range of prevalent beliefs and expectations (Correll & Ridgeway, 2003). Specifically, gender and race are associated with widely held beliefs that there are meaningful differences between women and men and between various racially marginalized groups and White people. Because evaluators strive to ascertain qualities of candidates they often cannot perfectly observe—such as expected on-the-job performance and underlying competence—they

4

draw inferences about valued attributes, using ascriptive characteristics as "culturally convenient coordinating device(s)" to assign value (Ridgeway, 2011, p. 93). Thus, across social settings, ascriptive processes commonly facilitate the application of gendered and racialized beliefs (e.g., status and stereotypes).

Though the nature of these beliefs is complex, in professional contexts gender and racial beliefs about competence or ability are particularly consequential. The primary goal of most evaluations—and thus of most evaluators—is to identify the best or most qualified candidate(s). However, assessing candidate competence and ability can be rather challenging because evaluators rarely have access to the requisite information to evaluate quality and predict performance. In hiring, for instance, managers and human resource professionals who are sifting through applicant profiles have access to information about candidates, including past work experience and educational credentials. However, these are imperfect proxies for how successful any given candidate will be in a specific role. Challenges with discerning candidate competence and ability are ubiquitous, emerging across many types of evaluations, from the allocation of prestigious awards and grants (Belz et al., 2022) to admittance decisions by universities (Castilla & Rissing, 2019).

By contrast, information about candidate race and gender is easily accessible: It can often be determined simply by reading a candidate's name (Pager et al., 2009). Ascriptive characteristics are associated with expectations about competence and ability (de Vaan & Stuart, 2022; Rudman & Phelan, 2008; Wayne et al., 2023). Although gender and race are typically *not correlated* with valued underlying attributes, such as competence (Brezina & Winder, 2003; Darity et al., 2022), evaluators nonetheless rely on them as proxies, particularly when they face uncertainty about a candidate's capabilities. Moreover, status-based theories describe how men

and White individuals are accorded greater social worth, creating expectations that they are stronger performers and should be evaluated favorably (Correll & Benard, 2006). Research on statistical discrimination describes making gender- and race-based inferences as rational: These inferences are thought to be based on evaluators' perceptions of the distribution of valued skills and attributes within a demographic group (Phelps, 1972; also see Rivera, 2020). Therefore, ascriptive characteristics are commonly perceived to be reliable indicators of actual performance, even though they are often unrelated to quality.

A broad set of social stereotypes connected to competence and ability shape the assumptions made about marginalized groups. For example, the feminine stereotype attributes communal characteristics to women, such as domesticity and passivity, while the African American stereotype characterizes Black people as intellectually inferior and unreliable (Moss & Tilly, 2001). In contrast, the masculine stereotype confers (predominately White) men with the agentic characteristics most valued in professional contexts, including perseverance and assertiveness (Fiske & Stevens, 1993). As a result of these stereotypes, evaluators tend to favor White men in their assessments in professional contexts.

Ascriptive characteristics inform a range of expectations about workers beyond those centered on competence (Ellemers, 2018). In many work domains, for instance, evaluators are concerned with a candidate's expected commitment to work (Rudman et al., 2012), which is arguably even more difficult and costly to discern than ability or competence. Men are more readily seen as fulfilling notions of the ideal worker who can and will prioritize work over all else (Acker, 1990). In contrast, assumptions that family responsibilities are gendered lead to concerns over women's prioritization of work, which typically disadvantages them in evaluations (Correll et al., 2007; Rivera & Tilcsik, 2016). And racial stereotypes about 'laziness' suggest that

6

Black workers are less committed and exert less effort than their White counterparts (Kluegel, 1990), which contribute to workplace disadvantages. For instance, concerns about whether a Black worker will be reliable can impede their chances of being referred for a job opening (Smith, 2005).

Beyond leading to more favorable evaluations for White men, positive performance expectations also make their successes stand out (Correll & Ridgeway, 2003). Whereas contributions from White and male candidates are perceived to be reflective of their effort and ability, those from marginalized candidates are commonly attributed to luck, and to the help of others, among other circumstantial factors (Heilman & Haynes, 2005). To overcome inequities in evaluations, women and racially marginalized groups must thus do more than White and male candidates—by demonstrating even stronger performances, acquiring additional signals of competence, and/or clarifying their distinct contributions. Accordingly, double standards theory—a status-based theory of inequality—shows evidence that evaluators prioritize ascriptive characteristics even when they have better proxies for valued traits (Foschi, 2000). Often, the exact same behaviors and signals of ability are discounted for women and members of marginalized racial groups, and only those who demonstrate the very best performance are evaluated on par with their similar White-male counterparts (Abraham & Botelho, 2024; Botelho & Abraham, 2017; Correll et al., 2020; Leslie et al., 2017).

One might conclude that marginalized candidates can overcome common inequities by demonstrating that they do not fit neatly into their proverbial stereotype 'box.' Yet, this is not generally the case. Attempts to counteract gendered and racialized stereotypes often lead to even greater disadvantages because stereotypes are not only descriptive but also prescriptive (see Ridgeway & Correll, 2006 for review related to gender). These beliefs tend to reinforce existing

social structures, and when they are violated, those at the top of the hierarchy can experience status threat (Rudman et al., 2012). Thus, when members of lower-status groups like women and racial minorities act in a manner that contradicts stereotypes it is often perceived negatively by others (Kmec, 2008; Rudman & Phelan, 2010; Williams, 1992).

For example, when women and Black candidates act agentically, displaying ambition and assertiveness at work, they often encounter backlash. As described above, women are expected to be modest and communal (Rudman, 1998) while Black people are expected to be unmotivated and unreliable (Rudman & Phelan, 2010). Behaviors that violate these stereotypes pose a potential threat to high-status others, especially when coupled with traits like competence (Moss-Racusin & Rudman, 2010). Thus, although there are often penalties for being a working mother, women who are ambitious, childless, and single can still face barriers to promotion because they violate gender stereotypes (Merluzzi & Phillips, 2016). Similarly, when Black people demonstrate ambition and assertiveness, they receive lower job performance ratings (Rudman & Phelan, 2010; Wayne et al., 2023) and are less likely to be promoted (Wayne et al., 2023). Low-status demographic groups thus encounter a double bind; they are generally perceived as poorer performers, and when they violate these expectations, they encounter pushback (Magee & Galinsky, 2008).

In sum, though evaluators may explicitly favor male and White candidates (Becker, 1957), mounting evidence suggests that demographic characteristics more commonly shape evaluations in implicit ways (Devine, 2001; Dovidio & Gaertner, 2000; Swim et al., 1995). Gender and racial beliefs are often activated automatically simply by learning of someone's demographic characteristics (Correll & Ridgeway, 2003, p. 31). Even the mere awareness of

these beliefs—and particularly awareness of how they shape evaluations—can license evaluators to rely more heavily on candidates' demographic characteristics (Tilcsik, 2021).

Unifying Framework on the (Re)Production of Inequality in Evaluations: Three Core Drivers

In the remaining sections, we develop our framework outlining three categories of drivers central in the production and reproduction of inequality in evaluations: (i) How Prevailing Beliefs Affect Inequality, (ii) How the Design and Structure of Evaluation Processes Shape Inequality, and (iii) How Evaluator Characteristics Influence Inequality. Each of these drivers transcend types of evaluations and thus are critical for comprehensively understanding when evaluative disparities will emerge. Note we do not contend that these uniformly either increase or decrease inequality in evaluations; rather, they have the potential to do both, depending on how they manifest in the evaluation process. Figure 1 provides a visual representation of the framework with an overview of the main categories of drivers that can produce inequality in evaluations and specifies various subcategories for each driver. Table 1 provides further detail about each of the drivers, including the fundamental concepts, influence on inequality, common empirical contexts, and key citations.

Figure 1. Three Categories of Drivers in the (Re)Production of Inequality in Evaluations

PREVAILING BELIEFS IN EVALUATIONS

Beliefs Emerging From Identity-Based Social Movements

Meritocratic Beliefs

Gendered and Racialized Beliefs About Suitability

DESIGN AND STRUCTURE OF EVALUATION PROCESSES

Features of the Candidate Pool

Evaluative Criteria, Processes, and Tools

Multistage Evaluation Processes

Presence of Relevant Audiences

(RE)PRODUCTION OF INEQUALITY IN EVALUATIONS

Demographic Characteristics

Experitse, Experience, and Achieved Status

EVALUATOR CHARACTERISTICS

How Prevailing Beliefs Affect Inequality

The first driver of inequality in our framework outlines how shared beliefs permeate evaluations in ways that can perpetuate or mitigate evaluative inequalities. Evaluations are fundamentally social and, thus, inherently shaped by prevalent schemas about value and deservingness across groups (Becker, 1957; Bourdieu & Passeron, 1990; Lamont, 2012). Accordingly, the beliefs and their effects in shaping evaluative outcomes vary considerably across time and context. In this section we discuss how dominant beliefs—as well as societal shifts in these beliefs—shape the propensity for evaluative inequalities. *How do beliefs affect gender and racial inequality in evaluations?* The groundwork is set by first exploring gendered and racialized perceptions about who is best suited for various occupations and industries, and how these perceptions impact evaluative outcomes. Then we consider how meritocracy—a prevalent belief in professional

contexts that ubiquitously shapes evaluations—affects inequality in professional contexts.

Finally, we discuss how identity-based social movements can impact evaluative outcomes, altering perceptions of attributes linked to gender and race, as well as highlighting the significance of workplace equality. Put differently, we consider variation in the prevalence of inequality in evaluations as a function of (i) gendered and racialized beliefs about suitability, (ii) meritocratic beliefs, and (iii) beliefs emerging from identity-based social movements.

Gendered and racialized beliefs about suitability

Prevailing beliefs related to who is best suited for certain types of occupations and tasks shape the pervasiveness of inequality in evaluations. Although the role of the worker in Western societies is generally linked to notions of masculinity (Acker, 1990; Ridgeway, 2009) and to perceptions of Whiteness (Ray, 2019), gendered and racialized expectations are not constant across all professional contexts. The salience of demographic characteristics—and thus the propensity for associated beliefs and stereotypes to shape evaluations—is most rampant in contexts traditionally associated with a particular gender or racial group (Ridgeway, 2009). Gender- or race-typing usually emerges when there has historically been a large numerical representation of people from that group in a given discipline, industry, occupation, or position (Cejka & Eagly, 1999). For instance, whereas industries such as finance and technology—fields persistently dominated by men—are male-typed, nursing and human resources—where women numerically dominate—are female-typed (Ridgeway, 2011). Likewise, fields such as medicine, business, and engineering where there are many White male workers, are White-typed (King et al., 2006). And low-status work like manual labor, customer service, and elderly care tends to be Black-typed because Black workers are overrepresented in these fields (Dupree et al., 2021).

A primary consequence of this demographic typing is the formation of beliefs about who 'fits' or is better suited for certain work (Kaufman, 2002). Specifically, those candidates whose demographic attributes match the gender- or race-typing of a given profession are expected to be more qualified than candidates whose attributes do not align. Candidates who do not fit the prototypical mold may thus be penalized with worse performance outcomes (Jensen et al., 2018; Smith-Doerr et al., 2019).

Accordingly, the disadvantages commonly facing women and racially marginalized groups tend to be most prevalent for high-status opportunities because these are often the very contexts most tied to Whiteness and masculinity (Kulich et al., 2011). Extensive research documents inequality in hiring based on gender (Doering & Thébaud, 2017) and race typing (Wingfield & Chavez, 2020) for higher-status jobs. Another study finds that while most demographic groups benefit from applying to high-status jobs when they have high-quality resumes, Black professionals are more negatively evaluated for high-status jobs, no matter their credentials (King et al., 2006). Conscious of these negative professional stereotypes, Black candidates sometimes strive to reduce potential hiring penalties by engaging in "résumé whitening," altering elements of their résumé that may signal that they are non-White, such as omitting affiliations with race-specific groups (Kang et al., 2016).

Marginalized groups in male and White-typed work settings are also less likely to receive resources, like research grants or venture funding. For instance, as entrepreneurship has been dominated by men and is linked to expectations that masculine traits are critical for success, female-led ventures tend to struggle with fundraising (Huang et al., 2021; Snellman & Solal, 2023). Similarly, female-managed mutual funds receive lower capital inflows than male-managed funds although there are no gender differences in fund performance (Niessen-Ruenzi &

Ruenzi, 2019). Consistent with these findings, women and racially marginalized groups have been more successful in making inroads in more gender-neutral industries, where gendered and racialized expectations are more muted (U.S. Bureau of Labor Statistics, 2004).

Unlike the advantages enjoyed by men in male and White-typed industries, women and racially marginalized groups are not commonly advantaged in industries in which they are significantly represented. In fact, there is some suggestive evidence that White men are still favored in these contexts (Kmec, 2008). For example, in fields such as nursing and elementary school teaching, men often receive promotions to positions of authority at a higher rate, in accordance with expectations that they should occupy high-status roles. This trend is commonly known as the "glass escalator," a term that contrasts with the "glass ceiling" women encounter when striving for leadership positions (Williams, 1992). It is largely limited to White men because expectations do not dictate that Black men should occupy high status, high-paying roles (Wingfield, 2009).

Irrespective of industry or occupation, persistent advantages for White men exist largely because they tend to be considered the best-suited candidates for leadership positions. The prototypical leader is agentic, competent, and powerful—all attributes associated with White masculinity (Rosette et al., 2008). Men are both *less* likely to be seen as accountable for team failures and accorded *more* responsibility for team successes than are women and racially marginalized groups with the same levels of performance. For example, Black people are stereotyped as athletic (Niven, 2020) and have come to dominate many professional American sports (Turco, 2010). Still, racial disparities persist in leadership positions across those sports, with White men still holding many of the most lucrative and authoritative roles both as coaches and players (Marquez-Velarde et al., 2023).

Even within the same jobs and occupations, specific elements and tasks are at times demographically typed. For instance, whereas male leaders are expected to value and be committed to organizational success, female leaders are expected to value socially- and diversity-focused organizational initiatives. Therefore, organizational initiatives are valued differently based on the gender of the leader (Abraham & Burbano, 2022; Bode et al., 2022). Relatedly, among Travel Security Agents (TSA), pat-downs of travelers—particularly female travelers—are seen as a more suitable task for women than for men. Consequently, female agents do more of this type of undesirable work than their male peers, which places an unfair burden on these women because it is both physically and emotionally laborious (Chan & Anteby, 2016).

Given this evidence, it is important to continue investigating the multifaceted effects of gendered and racialized suitability beliefs across professional contexts. A core challenge is that many of the gendered and racial associations that we discussed are historically embedded, resulting in path dependencies. Demographic typing essentially entrenches inequities and segregation over time because candidates whose demographic attributes match those associated with the context are more likely to be evaluated favorably (Heilman, 1983; Kanter, 1977). Thus, we need to not only identify these associations but also develop our understanding of how they are shaped within institutionalized settings. Researchers can do this by focusing on why suitability beliefs affect job opportunities, performance evaluations, and access to resources. Furthermore, reducing inequality requires that we consider how these associations with race and gender can be dismantled.

Future research should consider the nested structural levels of gendered and racialized work: Existing work shows that disciplines, industries, occupations, organizations, positions, and even tasks can be demographically typed. Individuals are not in jobs devoid of an industry; rather, these facets are nested. For example, an industry like biotechnology is male-typed, but the

position of human resources administrator at a biotech firm is female-typed. How might the gender typing across these levels intersect to shape evaluative outcomes? Given that workplaces and other professional contexts are complex and nested, it is necessary to deepen our understanding of how inequality is produced in these real-world environments.

Meritocratic beliefs

Prevailing gender- and race-based beliefs about the nature of work are extremely sticky and are not easily overcome (Melamed et al., 2019; Ridgeway & Erickson, 2000). However, in professional contexts—and particularly those in the US—there are beliefs that evaluations *should* be fair and impartial, which may counteract people's prejudices. In particular, fundamental support for meritocracy could mitigate the disadvantaging effects of gender and race in evaluations. The concept of meritocracy emphasizes that individuals' outcomes and access to resources be based on their capabilities, contributions, and performance rather than on their race, gender, class, familial connections or other personal attributes that are not directly tied to ability or effort (Castilla & Benard, 2010). People generally believe the allocation of resources and rewards is—and ought to be—fair and reflective of meritocratic processes (Lerner, 1980) and these beliefs have only been strengthening over time (Mijs, 2021). Thus, evaluators may strive to act in accordance with these principles.

Meritocratic beliefs and expectations have permeated many professional and organizational contexts in recent years (Castilla & Benard, 2010). This is evident when we consider the stark rise in organizational statements related to eradicating workplace discrimination and commitments to equality over the past two decades (Kang et al., 2016). Such organizational efforts to *be*—or at least to convey a commitment to *becoming*—meritocratic are in part motivated by pressures from key stakeholders. Policymakers, investors, and employees

increasingly expect merit-based processes and call attention to instances where organizations do not demonstrate such a commitment. For instance, US Equal Employment Opportunity Commission (EEOC) regulations outlaw differential treatment of job applicants or employees based on their demographic or social identity (EEOC, 2023). In response, most employers communicate their commitment to equal opportunity employment practices.

Analyses of gender and racial inequality in professional contexts—such as in pay and access to employment opportunities—reveal some assuring trends that suggest having meritocratic principles at the forefront may be somewhat effective in reducing inequality in evaluations (England et al., 2020). Nonetheless, sociologists caution against concluding meritocratic beliefs and practices are a panacea for eradicating inequality (Castilla & Benard, 2010; Conzon, 2023). In fact, merit-based practices may instead reproduce or exacerbate inequalities. For example, in organizations espousing meritocratic policies, managers still favored men over equally performing women when translating performance ratings into rewards (Castilla & Benard, 2010). Furthermore, the gender gap in bonus pay has been found to be greater in workplaces *with* a merit-based pay system compared to those without such systems (Mun & Kodama, 2022). Relatedly, organizations claiming to be committed to Diversity, Equity, and Inclusion (DEI)—which suggests they strive to be meritocratic—behave no more equitably in their evaluations than organizations not making such claims (Kang et al., 2016).

The strength and nature of meritocratic beliefs can also *reinforce* beliefs that existing inequalities are warranted. System Justification Theory is a social psychological theory that proposes individuals have a need to perceive the existing social order as fair and legitimate. Resultantly, they are predisposed to believe that when there are gender or racial differences in evaluative outcomes, these outcomes reflect individual differences in effort and quality (Jost,

2020). This is thought to explain a general lack of concern with rising levels of inequality in US society (Mijs, 2021). Consistent with this perception, majority group members often dismiss discrimination claims, believing equitable evaluative processes are already in place (Kaiser et al., 2013). Contrary to what might be expected, it is not only advantaged groups that hold this view that existing systems are already meritocratic. Recent research suggests that even those most disadvantaged believe that processes are generally meritocratic, which further reinforces the dominant social ordering (Batruch et al., 2023).

Perceptions of meritocracy, particularly within the context of work, are an important determinant of inequality. Scholars have a general understanding of how meritocracy is interpreted in evaluations; sometimes, it might increase fairness and equity, but it can also lead evaluators to perceive the world through "rose-tinted glasses" and thus to overlook inequalities. Meritocracy, however, is a culture-laden concept, central to American conceptions about work (Hochschild, 1995). Future research could examine meritocracy across cultural contexts, as meritocracy is idealized less in some countries than it is in the US. More importantly, we should continue research aimed at aligning the perceived and actual fairness of merit-based systems, to avoid the distorting impacts of assumed meritocracy. For instance, future research could examine interventions that simultaneously espouse meritocratic processes and reveal the current level of inequality in outcomes. Pairing meritocratic goals with information about persistent inequality may help overcome misconceptions that equality has been achieved.

Beliefs emerging from identity-based social movements

Much like meritocratic beliefs, social movement activism can also inform evaluative processes in professional contexts as it increases awareness of inequality and discriminatory practices. Identity-based social movements—centered on demographic characteristics (e.g., gender or

race)—have become increasingly prevalent and influential in shaping institutional evaluation processes since the 1990s (Davis et al., 2022). The Civil Rights movement of the 1950s and 1960s has given rise to other social movements by providing a model for how to successfully advocate for the rights of marginalized groups (Skrentny, 2014). Subsequently, collective action in the US aimed at leveling the playing field, such as the Black Lives Matter and #MeToo movements of the past decade, has increased the visibility and perceived severity of inequality. These movements have the potential to reduce disadvantages for historically marginalized groups. Following major Black Lives Matter protests, for instance, Black women job seekers encountered less hiring discrimination, receiving similar job callbacks to White men (Chavez et al., 2022).

Evaluators need not be personally persuaded by movement activism for change to occur since there can be severe reputational and business costs to foregoing change. Social movements are interwoven with webs of actors both within and outside of firms that can hold evaluators accountable and thus drive behavioral changes (Briscoe & Safford, 2008). When individuals and organizations have ties to scandal, they become more cognizant of reputational risks connected to discriminatory behavior in evaluations, and they may strategically aim to appear more equitable (McDonnell et al., 2021). For example, after the Harvey Weinstein scandal, film producers previously associated with Weinstein made more equitable hiring decisions, increasing their likelihood of working with female writers by 35% (Luo & Zhang, 2022). Thus, while these changes may reflect underlying reductions in evaluator's biases, these could also be fueled by purely exogenous forces.

Similar reputation-based drivers may instead lead social movements to have the unexpected effect of *increasing* inequitable evaluations. Media coverage of the #MeToo

movement, for instance, negatively impacted the evaluations of prospective female board directors at major US firms. Incumbent directors—who were predominantly White men—were concerned about interacting and working with women and thus favored appointing male directors (Bednar et al., 2022). Evaluators' heightened concerns about reputational costs may drive them to avoid interactions with women and racially marginalized groups altogether. Evaluators are also more likely to demonstrate bias against members of marginalized groups following events that reify underlying racialized or gendered beliefs. Though not focused on social movements, community exposure to violent crimes perpetrated by Black people augmented negative biases, leading to disadvantages for Black job seekers (Gorbatai et al., 2023; Mobasseri, 2019).

Recent scholarship thus indicates that identity-based societal events alter the way organizations, and the evaluators within them, think about gender and racial inequality. As a result, these events have the potential to reshape decision-making in evaluations; however, the effects on inequality are divergent. One possible explanation for these divergent findings is that evaluators' responses to societal events depend on the types of threats they face. These events can impose personal threats to the evaluator (e.g., concerns about being accused of sexual harassment), which can in turn lead to an uptick in discriminatory behavior. In contrast, in cases where evaluators are most concerned with appearing unbiased, these events may instead reduce inequality.

Future research is necessary to empirically examine the conditions under which identity-based movements reduce inequality in evaluative outcomes. Counter movements can, for instance, increase inequality. Social movements that advocate zero-sum outcomes may be especially likely to exacerbate inequality in comparison to those that claim benefits for all. Another promising avenue is exploring the role of social media and digital activism in

amplifying or mitigating these movements' impacts on organizational decision-making. With the advent of computational tools for social science research, mapping the detailed development of movements is now more feasible and can be incorporated to better identify responses to emergent movements.

How the Design and Structure of Evaluation Processes Shape Inequality

Beyond gendered and racialized beliefs affecting inequality, evaluations operate within specific systems that guide decision-making. The second category of inequality drivers in our framework focuses on these systems. Specifically, we consider how the structure of evaluations directly shapes the likelihood of inequality. By this we mean the design and intended goals of evaluation processes, as determined by their creators. *How do structural features of evaluative systems exacerbate or mitigate the propensity for gender or race to shape assessment outcomes?*

Since there are so many possible design choices, we cannot provide an exhaustive review of every potential consideration. Instead, our goal is to introduce the structural elements in professional evaluation processes that are most likely to affect the degree of inequality. We highlight design choices that constrain evaluators' objectivity, increasing the likelihood that they rely on commonly held gendered and racialized beliefs in their assessments. We focus on the following four structural elements: (i) features of the candidate pool, (ii) multistage evaluation processes, (iii) evaluative criteria, processes, and tools, and (iv) the presence of relevant audiences. This approach uncovers both the predominant structural drivers of inequality in evaluations and the potential levers for redressing it.

Features of the candidate pool

People are often evaluated within a group, rather than individually. When evaluators assess a single candidate, the focal question is: *Is this candidate sufficiently high-quality?* However, when

people are evaluated as part of a larger group, the focal question shifts to: *Which candidate is the highest quality?* For instance, in gig work, platform users compare gig workers offering similar services to select who to hire, and in promotion decisions, managers compare employees to determine who to reward. When we consider evaluations that occur in groups, features of the candidate pools have been shown to affect evaluations, impacting the likelihood that evaluators rely on gender or race in their assessments of candidates.

The size of the candidate pool is one such feature. As the number of candidates in a consideration set increases, evaluators are at risk of choice overload (Iyengar & Lepper, 2000). An early study on this topic, for instance, showed that people presented with four times more options were ten times (30% compared to 3%) less likely to make a choice (Iyengar & Lepper, 2000). Within the context of evaluations, as the number of candidates increases, it is generally harder for evaluators to examine candidate profiles carefully and make informed choices (Sherf et al., 2019). For example, venture capital firms receive many pitches from startups interested in receiving funding and employees at the firm cannot dedicate the requisite time to assess each startup's quality in detail. As such, they often end up focusing their attention on only a small portion of available options, ignoring many high-quality alternatives (Brenčič, 2014).

The constraints placed on evaluators by larger candidate pools lead them to seek shortcuts in their assessments, which include relying on gender- and race-based inferences. Beliefs and expectations associated with gender and race offer cognitive shortcuts for redressing general uncertainty about candidates (Ridgeway, 2011, p. 93). Accordingly, when candidate pools are especially large, it increases the likelihood that ascriptive characteristics are used to assess which candidates are deserving of further attention. Botelho and Abraham (2017) provide direct evidence of this point: Evaluators rely on gender as a sorting heuristic, which in turn

21

disadvantages female candidates. Further substantiating that gendered beliefs facilitate evaluations when there are high volumes of candidates, this gender-based sorting emerges even when evaluators have direct access to clear and objective information about candidate performance. Constraints on an evaluator's time similarly lead them to rely on ascriptive characteristics (Sherf et al., 2019). For instance, recruiters are more likely to base their outreach efforts on candidate gender, such that they favor male over female prospective applicants, when they have high workloads (Lane et al., 2023).

The demographic composition of the candidate pool can also affect the propensity for inequality to emerge. Demographic atypicality of an applicant pool—meaning the extent to which the demographic composition of the pool contradicts what is expected for a given work context—is one particularly meaningful aspect of pool composition. With respect to gender, a greater presence of female candidates can reduce common gender-based inequality in male-dominated contexts. For example, female applicants for traditionally male-typed jobs are evaluated more positively when there is a larger share of other women also applying (Robbins & DeNisi, 1993). Having a critical mass of atypical candidates guides evaluators to make more equitable judgements, such that they benchmark their expectations differently, because atypical demographic characteristics (e.g., being female) become less distinctive (Leung & Koppman, 2018).

Furthermore, these potential inequality-reducing benefits do not emerge uniformly from pools with more atypical applicants. If demographically atypical hires are perceived as token members of their demographic groups, larger numbers of these candidates can create perceptions of threat among dominant demographic group members (Campero & Fernandez, 2019). Candidate pools for White-typed jobs with a greater share of racially marginalized groups are

thus shown to sometimes yield an even *greater disadvantage* for those groups: The likelihood that Black workers will be hired for demographically atypical jobs decreases when they make up a greater portion of applicants (Huffman & Cohen, 2004).

We encourage future research on the dynamics of candidate pools in shaping inequality in evaluations. It is important to note that we are not claiming that small pools or pools with "X" percentage of atypical candidates are uniformly better than others for reducing inequality. Identifying how to minimize inequality by engineering candidate pools will likely depend on a host of contextual factors and there is no one-size-fits-all solution. It is not possible to change the historical composition of candidate pools, which likely affect the evaluation of current candidate pools. However, researchers can consider how providing evaluators with information about the composition of candidate pools, and the percentage of atypical candidates in those pools, affects the emergence of inequality. Similarly, it is often possible to reduce candidate pool size (or at least the number of candidates any evaluator is considering), which may lead to less reliance on candidate gender and race and thus minimize inequality. Future research could test what pool sizes and how much information on each candidate, is most effective for reducing inequality across types of evaluations.

Multistage evaluation processes

Evaluation processes also vary with respect to whether they are limited to a single stage or span multiple stages. Evaluations in organizational contexts are frequently multistage (Castilla, 2008), as this allows for greater perceived evaluative rigor, which is common when evaluations are formalized and institutionalized (Castilla & Benard, 2010). For instance, the startup investment process for venture capital firms typically involves sequential evaluations, starting with an initial screening of startup materials, followed by an assessment of a pitch, and concluding with

multiple face-to-face interviews/discussions. In such multistage evaluations, key differences across these stages may directly affect the propensity for inequality to emerge.

Despite the prevalence of multistage evaluations, this feature has not been often been the explicit focus of existing scholarship on inequality.[2] Rather, research has commonly documented inequality in one stage of multistage processes: at the initial screening of job applicants in hiring processes (Campero & Fernandez, 2019), for example, or at the investment stage for entrepreneurs (Brooks et al., 2014). These examinations may inadvertently mask stage-specific differences and thus lead to a more limited theories about the conditions under which inequality varies across stages. Because these stages of evaluation tend to differ in important ways—most notably in terms of the candidate pool and the evaluators—the likelihood for inequality to emerge is also likely to vary across stages.

Notably, nearly all multistage evaluations funnel such that the pool of candidates becomes smaller in subsequent stages and evaluators commonly have access to additional, relevant information in later stages of evaluation. For instance, consider the selection process for prestigious awards and grants. Researchers submit an initial short proposal that evaluators use to determine who will advance to later stages of the evaluation process. The smaller groups of candidates who are selected then submit more detailed proposals, often including a response to reviewer comments and an exposition about their research. As we detailed above, both smaller candidate pools and greater certainty about candidate quality which are common in later stages of evaluation indicate there will be a reduction in inequality. Botelho and Abraham's (2017) examination across two stages suggests that this is the case: Though evaluators used gender as a

---

[2] There are some additional papers that examine multistage evaluations including Goldin and Rouse (2000) and Castilla and Benard (2010), but the main takeaways for these papers are not specifically about the multistage nature of evaluations. It is considered as a feature of the evaluative setting—primarily based on the notion that multistage evaluations are common—rather than a factor associated with variations in the emergent degree of inequality.

"sorting heuristic" in the first stage, these same evaluators did not rely on gender in the following stage of evaluation where there were fewer candidates to assess and more information available about each candidate. However, evaluators in their context remain constant across both stages, which is not always the case in professional settings. It remains unclear whether inequality always decreases in later stages of evaluations (Botelho & Chang, 2023).[3]

It is common for evaluators to change across stages of evaluation, which may also affect the propensity for inequities to emerge. In the context of hiring, for instance, intermediaries and recruiters tend to screen potential candidates (Fernandez-Mateo & King, 2011). Although these individuals have expertise in recruiting, they are typically less knowledgeable about the specific domain for which they are hiring. It is often the case that screening-stage evaluators are not domain experts. As we discuss below (see subsection, Evaluator expertise, evaluation experience, and achieved status), lack of expertise among early-stage evaluators may inhibit their ability to assess candidates objectively. Furthermore, the general underrepresentation of women and racially marginalized groups in high-ranking organizational positions suggests that it may often be less common to have evaluators from marginalized groups in later stages. We discuss the implications of these evaluator characteristics further in the third and final category of drivers of inequality (see section, How Evaluator Characteristics Influence Inequality) but shifts in who is evaluating candidates may lead to marked differences in observed race and gender inequality across stages of evaluation.

Although research alludes to the fact that most evaluation processes are multistage, it is rare that the relationship between inequality and specific evaluative stages is directly examined.

---

[3] Botelho and Chang (2023) discuss that this may not always be the case: They identify no gender difference in the initial callback stage, but anecdotal evidence from interviews with recruiters suggests that gender inequality occurs in subsequent hiring stages.

This is likely because data across stages are difficult to access. Current scholarship has barely scratched the surface in understanding inequality in multistage evaluations. It does not provide concrete answers and often does not even offer speculative trends as to whether and how multiple stages impact inequality in evaluations. There are thus many opportunities for future research to empirically identify variations in inequality across stages and develop more comprehensive theories. Notably, future research should unpack the association between multistage evaluations and inequality when new evaluators are introduced across stages. How do evaluators coordinate from one stage to the next and how are each of their assessments factored into final evaluative outcomes?

Evaluative processes, criteria, and tools

Although there are a multitude of ways that evaluation processes differ, inequality scholars have disproportionately focused on the effects of formalizing evaluative systems for inequality (Baron et al., 1986; Elvira & Graham, 2002). Lack of clarity about how one ought to assess candidates is often associated with more inequitable outcomes. As we have explained, the goal of evaluations is typically to differentiate candidates based on merit and then select the highest-quality candidate(s) (Alon & Tienda, 2007). One challenge that can prevent evaluators from achieving this goal is that definitions of what it means to be the "best" or "highest quality" are not always explicit—and at times are not even agreed upon. Whereas one evaluator may consider quantitative skills as paramount in judgments of performance, another evaluator may believe communication skills to be most important. Therefore, uncertainty about how to evaluate can result in stark differences in assessment criteria and outcomes for the same candidate within the same evaluation system.

When evaluators lack clarity on how to make assessments, they are also apt to prioritize more subjective information about candidates and to rely on personal beliefs, including those tied to gender and race (Ridgeway, 1997, 2011). Evaluators are more likely act on their biases— either implicitly or explicitly (Castilla, 2015). In hiring, for instance, evaluators frequently prioritize whether a job candidate will fit in and be committed to the hiring firm (Chatman, 1991; Rivera, 2012). Although assessing candidates for fit is reasonable and expected, the subjective nature of these fit assessments can lead to disadvantages for candidates from marginalized groups. Consistent with this, research on tokenism has shown that women and racially marginalized group members are less likely to be deemed a cultural fit in various settings and thus experience worse outcomes (Turco, 2010).

It is now well established that formalization of criteria through objective metrics and of processes through rules can significantly lessen uncertainty and is thus an effective strategy for reducing inequality (Anderson & Tomaskovic-Devey, 1995; Bohnet et al., 2016; DiPrete et al., 2010). Highly-structured organizational compensation systems and formalized employment relations—like collective bargaining—are demonstrated to reduce gender pay disparities (Elvira & Graham, 2002). Similarly, structured job postings and career progression ladders lead to less inequality in hiring and promotion (Dobbin et al., 2015). And some evidence suggests that anonymizing applicant demographic characteristics and information that can give away their demographic background (e.g. applicant name, participation in an ethnicity or race-based professional association, place of origin), can significantly reduce biases in evaluations (Goldin & Rouse, 2000; Younkin & Kuppuswamy, 2018). However, research has also found evidence that removing this information does not attenuate disparities because evaluators key in on other

related information (e.g., Behaghel et al., 2015). Thus, there may be conditions where anonymization is more effective than others.

Beyond providing evaluators guidance, formalization also increases evaluator accountability. Evaluators are more likely to base their assessments of candidates on clearly defined criteria to avoid the risk of their personal biases being detected (Castilla, 2015; Chung et al., 2020).[4] As might be expected, research suggests that marginalized groups prefer standardized evaluations schemas (Berger et al., 2020; Fernandez & Sosa, 2005). Managers, in contrast, sometimes oppose enforced top-down initiatives that increase their accountability and constrain their ability to freely assess employees (Dobbin et al., 2015).

However, formalization does not always decrease inequality. Evaluation processes will curtail inequality depending on the specific criteria that are formalized (Dencker, 2008). Institutionalized systems tend to be shaped by bureaucratic politics and determined by high-status actors (Bridges & Nelson, 1989). These high-status actors are motivated to reinforce existing systems that reify their status positions. Thus, even if evaluative criteria are clear and structured, they may reinforce existing systems that perpetuate inequalities (Baron, 1984; Ray, 2019). And at times discretion, rather than formalization, fosters greater equity, by permitting evaluators who support change to counteract the deleterious effects of inequitable systems (Abraham, 2017).

Furthermore, even reliance on seemingly meritocratic criteria as proxies for valued traits, like results on standardized tests or credit scores, can reinforce racial and gender inequalities (Kiviat, 2019; Burrell & Fourcade, 2021). For example, high credit scores—calculations based

---

[4] Accountability can both increase or decrease inequality, dependent upon who an evaluator is accountable to, and the preferences of that authority. In relation to formalization specifically, accountability reduces inequality when we assume formalization is generally better than no formalization. It provides clear metrics with which to evaluate candidates, compelling evaluators to not simply rely on their biases when making evaluative decisions.

on credit history that are meant to be standardized—are not equally accessible to everyone. Systemic racism makes it harder for an equally skilled and qualified Black person to receive as high of a score as a White person (Kiviat, 2019). Thus, evaluations with formalized processes that prioritize biased criteria can reproduce inequalities as biased metrics can give the (false) impression of evaluative rigor (Espeland & Sauder, 2007).

Alongside defining the processes and criteria used to assess candidates, designers of evaluative systems also generate tools and devise metrics to capture assessments of candidates. These quantitative metrics tend to be regarded as "neutral and objective" and thus more legitimate (Mazmanian & Beckman, 2018). Though inequality scholars have paid less attention to the relationship between the specific tools used in evaluations and inequality, there is reason to believe these matter (Rivera & Tilcsik, 2019). Research in psychometrics shows that the specific design of survey measures, for instance, affects the reliability of those measures (Rust & Golombok, 2009).

Inequality in quantitative assessments that commonly dictate access to valued resources often vary based on the scale used. Some assessments rely on binary scales (e.g., whether to hire, invest, or promote) while others use multipoint scales (e.g., rating a candidate from one to five). Rivera and Tilcsik (2019) provide direct evidence that the design of the rating scale contributes to inequality: Under an original 10-point scale, female professors received worse student evaluations than male professors, but this gap was significantly reduced when the university changed to a 6-point scale. They discuss that the number 10 evoked perceptions of "brilliance," which are salient in scientific fields, and are less likely to be applied to women.

The criteria, processes, and tools that are formalized in evaluative systems define how evaluators make assessments. Though formalization may accentuate inequality in evaluative

outcomes, most evidence suggests formalization diminishes it. The key to the equalizing effects of formalization lies in thoughtfully considering formalized elements to ensure they are unbiased. Accordingly, future work should further explore the ways evaluators use specific tools and measures to maintain existing social hierarchies. Quantitative metrics—such as scale-based ratings and numerical rankings—are generally assumed to be objective and are therefore used to justify differences in performance-based rewards (Kellogg et al., 2020). As performance-based rewards continue to dominate workplaces, it is necessary to consider how they can be inherently biased with direct consequences for race and gender inequality. Furthermore, future research could also explore how to evaluate performance given the shortfalls of quantitative metrics by, for instance, pairing quantitative and qualitative evaluative processes (see Belliveau, 2012 for an example of existing research taking such an approach).

Presence of relevant audiences

Beyond the candidate and the evaluator involved in a focal evaluation, there are commonly audiences—individuals or groups—that may also shape the propensity for gender and racial inequality to emerge. Audiences are typically not directly evaluating the candidate—and may never evaluate them—but evaluators often consider audiences' judgements and preferences when making their assessments. Social psychologists and sociologists differentiate between two kinds of evaluative behavior that vary depending on whether audiences are present. In cases where evaluators are not at all concerned with audiences, they are free to base their assessments on their own perceptions of quality. By contrast, the presence of audiences lead evaluators to deduce and prioritize audience preferences in assessments (Ridgeway & Correll, 2006). Evaluators will consider how their own assessments align with the expectations of relevant others, such as colleagues, customers, and future evaluators (Becker, 1957; Correll et al., 2007).

The presence of audiences can have divergent effects on inequality in evaluative outcomes. On the one hand, audiences can promote fairness in evaluations by holding evaluators accountable and encouraging meritocratic decision-making (Castilla, 2008; 2015). This is consistent with research showing that people who need to justify their decision to others are more thorough in their decision-making and resist biases (Tetlock, 1983). Speaking to this point, evaluative disparities are significantly reduced in performance-reward systems where managers are held accountable for their assessments of employees (Castilla, 2015; Chung et al., 2020). Similarly, when evaluators expect their assessments will be judged by relevant audiences, they are less likely to let candidates' demographic characteristics influence their evaluations (Botelho & Gertsberg, 2022).

On the other hand, the presence of audiences can exacerbate inequality. This is especially likely when evaluators are making an assessment on behalf of someone else. Inherent uncertainty about others' beliefs leads decision makers to generate inferences about audiences' likely preferences linked to gender and race (Beckman & Phillips, 2005; Melamed et al., 2019). Evaluators are likely to assume that people hold dominant gender and racial biases because these are so pervasive (Abraham, 2020; Correll et al., 2017; Fernandez-Mateo & King, 2011). For instance, managers are inclined to make hiring and promotion decisions based on their assumptions of customer preferences, especially when employees work in customer-facing roles (Becker, 1957; Neumark et al., 1996). These assumptions are especially prevalent because it is often nearly impossible to definitively know others' beliefs related to gender and race. As a result, women and racial minorities can be disadvantaged in evaluations where evaluators are weighing the preferences of relevant audiences.

Relatedly, inferences that others hold dominant gender and racial beliefs are a common explanation for why racially marginalized groups receive fewer connections to valuable others, including potential employers. People may avoid giving job referrals to members of marginalized groups because they are worried doing so may reflect poorly on them, specifically by negatively affecting the potential employer's perception of them for making the connection (Smith, 2005). Professional staffing firms that function as labor market intermediaries similarly anticipate the gender preferences of their clients (potential employers). Fernandez-Mateo and King (2011) introduce the concept of "anticipatory gender-sorting" to explain how initial screeners at staffing firms favor male candidates because they expect employers prefer male job candidates. Inequality-producing effects of audiences can emerge *even* when evaluators do not hold negative gender or racial biases themselves (Walker et al., 1988). For example, entrepreneurs are more likely to connect their male peers to potential clients because they assume these clients likely prefer to work with men over women (Abraham, 2020). However, these same entrepreneurs do not exhibit biased behavior when choosing with whom to work directly.

Audiences also perpetuate inequality by shaping other general design elements of evaluations. Evaluators who tend to be most concerned about audiences' preferences are typically involved in the early stages of evaluations, which means that audiences are particularly important in shaping the gender and/or racial composition of candidate pools. These early-stage evaluators have been shown to steer candidates into different opportunities, sorting female job seekers into lower-paying female-typed jobs (DiPrete & Soule, 1988; Fernandez-Mateo & Fernandez, 2016). Lower-status evaluators are also most apt to be concerned with the impressions of others. For example, law firm associates—who are relatively lower status in their firms—are apt to perceive their direct superiors as an important audience. Evidence suggests that

when these supervisors are liberal rather than conservative, there is more gender parity in associate earnings, irrespective of the associate's own preferences (Briscoe & Joshi, 2017; Carnahan & Greenwood, 2018).

Audiences are often present when considering evaluations in professional contexts. Although evaluators are making their own assessments, research clearly articulates that they are doing so with the audience in mind. In some cases, audiences serve as oversight, helping mitigate inequality but in others they magnify inequality. We have surmised that the precise role—or perceived role—of the audience is a key dimension helping to reconcile these divergent findings. However, there is room to expand our understanding of the role of audiences in generating inequality. Future research should further explore how the relationship between the evaluator and the relevant audience shapes the production of inequality in evaluations. In some cases, an evaluator is supervised (or accountable to) the relevant audience (e.g., a manager), while in other instances, evaluators and audiences have the same social standing (e.g., coworkers). How might the precise position of the audiences with respect to the evaluator affect resultant inequalities? Another avenue for future research lies in considering whether audience communication about preferences influences evaluator decision-making. Some work suggests that when customers communicate preferences for more diverse workplaces, managers rely on this direct information rather than general assumptions about biases such that they hire more diverse pools of employees (Pedulla et al., 2023).

How Evaluator Characteristics Influence Inequality

The third category of drivers we consider in this framework relates to the distinct role of evaluators, and how characteristics of evaluators are associated with producing varying degrees of inequality in evaluations. Although the design and structure of evaluative processes constrain

evaluator discretion, individuals—or groups of individuals—are ultimately responsible for the implementation of these processes. In the case of promotion decisions within organizations, for instance, evaluations of candidates are conducted by a committee, direct supervisors, peers, or a combination of these. Broadly, research indicates that evaluative outcomes for the same work may differ simply given differences in who is making an assessment (Botelho, 2024; Kovács & Sharkey, 2014; Li, 2017). This is because evaluative judgements depend heavily on an evaluators' own identity and beliefs (Abraham, 2017; Carnahan & Greenwood, 2018; O'Reilly & Chatman, 1996; Rivera, 2012). These factors not only determine which specific criteria evaluators prioritize when assessing others but also guide their interpretation of candidate behaviors and attributes (Ridgeway & Correll, 2006). In other words, *who* is evaluating can determine whether and to what extent inequality arises.

Building on this perspective, we consider the following question: *How do evaluators' characteristics influence the degree of gender or racial inequality in evaluative outcomes?* We specifically identify two broad categories of evaluator characteristics: (i) demographic characteristics, and (ii) expertise, experiences being evaluated, and achieved status (which are non-demographic characteristics).

Demographic Characteristics: Evaluator Gender and Race

Demographic similarity between evaluators and the individuals they are assessing can profoundly affect inequality. Observable demographic characteristics, such as race or gender, significantly influence preferences, primarily due to identification processes by which individuals classify people based upon visible attributes (Ridgeway, 1991). Categorizations are ubiquitous in social interactions, as people are continuously defining who they are relative to

others (Ashforth & Mael, 1989; Reskin, 2000).[5] Thus, people often see their in-group positively, as they desire to see themselves positively (Gorman, 2005); they exhibit in-group favoritism (Tajfel & Turner, 1979). Attraction to similar others is also partially grounded in the belief that people who share our observable attributes or characteristics will also share more fundamental, valued attributes, regardless of whether this is true (Milliken & Martins, 1996). Essentially, individuals assume that those who are demographically similar will also be alike in other significant, albeit less observable, ways.

This propensity to engage in demographic categorization becomes even more pronounced among those in positions of power (Jost & Banaji, 1994). As gatekeepers governing access to resources and opportunities (Lamont, 2012; Rivera, 2015; Smith, 2005), evaluators may have a vested interest in providing access to demographically similar others. This practice contributes to the reproduction of gender- and race-based disadvantages, as those in positions of authority are commonly White men who tend to favor other White men. Consequently, the historical underrepresentation of women and racially marginalized groups in these positions perpetuates inequality.

However, demographic similarity can also lead to more equitable outcomes when marginalized groups favor and support same-race and same-gender candidates who would otherwise be at greatest risk for disadvantage. Research on homophily shows that demographically similar people tend to connect at a higher rate than would occur by chance (Ibarra, 1992; McPherson et al., 2001). For instance, part of the gender pay gap in medicine is explained by doctors referring patients to other same-gender physicians. Since there are more

---

[5] Note that these demographic categories are socially constructed such that they vary across countries and cultures. Two people who are the same race in the US, for instance, might not be in another country.

male physicians to begin with, this leads to more patient referrals to other male physicians (Zeltzer, 2020).

A similar trend can be observed in other professional environments, illustrating the widespread influence of demographic similarity on workplace equity. For example, the gender pay gap—a pervasive disparity in gender-based earnings—is often linked to the sustained underrepresentation of women in managerial positions (Bridges & Nelson, 1989; Ely, 1995). Scholars thus show that there tends to be greater equity in settings with a larger proportion of women in positions of authority (Cohen & Huffman, 2007). Women and racially marginalized groups tend to receive more favorable performance assessments from demographically similar evaluators across various aspects of workplace evaluations including hiring, promotions, and salary decisions (Cohen & Broschak, 2013; Cohen et al., 1998; Gorman, 2005). Similarly, recent work shows that subordinates evaluating their superiors also tend to favor those who are demographically similar and are more likely to recognize their achievements (Stainback et al., 2016).

Demographic similarity not only fosters preferences but also suggests a common understanding of social obstacles. Evaluators and candidates from the same marginalized groups have likely faced similar biases and inequities. As a result, marginalized evaluators tend to be mindful of potential biases in other evaluators' judgments; they may scrutinize other's assessments of candidates who share their social identity. In some cases, this awareness prompts evaluators from marginalized groups to advocate on behalf of ingroup members (Cohen & Huffman, 2007). Greenberg and Mollick (2017) offer evidence of this in the context of donation-based crowdfunding. They find that female investors are most apt to advocate on behalf of female founders, and this trend is most prevalent in industries where gender biases tend to be

most pronounced. This has tangible benefits, increasing women's access to funding in male-dominated fields. While women are often penalized for self-promotion in professional contexts, they do not face sanctions for advocating for others (Amanatullah & Morris, 2010). Consequently, there may be downstream costs to those helped by demographically similar evaluators; recent work shows that female founders who received investments from female investors were perceived to be lower quality by male investors (Snellman & Solal, 2023).

Furthermore, maintaining demographic similarity with evaluators does not always advantage marginalized candidates. When evaluators are concerned about their own status position and reputation, they are less apt to act on preferences for similar others. People's behaviors are generally motivated by a desire to improve their social standing, and thus their actions do not purely reflect personal preferences and beliefs (Ridgeway et al., 1998, 2009). Female and racially marginalized evaluators' concerns about their admittedly more fragile status position means that these evaluators are at times unwilling or unable to support in-group candidates. In line with this, research demonstrates how marginalized groups exhibit the same biases as their majority group counterparts when their own social position is threatened (Duguid et al., 2012). For example, though female managers compensate their lowest-ranking subordinates more equitably, they generate the same gender pay gaps as male managers among their higher-ranking subordinates with whom they ostensibly face greater competitive threat (Abraham, 2017). Relatedly, high-performing female supervisors allocate lower salaries to low-performing female employees when they anticipate that their association with a low performer of the same gender will negatively impact their own work reputations (Srivastava & Sherman, 2015).

An evaluator's lower-status position can especially disincentivize them from showing support for demographically similar candidates when they are a numerical minority. In these instances, individuals from marginalized groups are perceived as 'tokens' with access to only a fixed number of allotted opportunities (Ely, 1995; Rudman, 1998). Being a token compels lower-status evaluators to dismiss in-group solidarity to maintain or enhance their own positions. If there are only so many opportunities afforded to their sub-group, they aptly infer that maintaining demographic distinctiveness is an advantage (Kirgios et al., 2020).

Moreover, people generally maintain existing social hierarchies, even when they are disadvantaged within these hierarchies. Predictably, those in higher-status positions tend to exhibit preferences that support the status quo (Ridgeway & Correll, 2006). But individuals who occupy lower-status positions also face pressure to conform to the status quo, even if doing so requires that they show deference and alter their personal opinions (Ridgeway et al., 1998). Evaluators from lower status groups—including women and those from racially marginalized groups—are no exception; they often feel compelled to conform to other's preferences when there are status differentials among evaluators (Ridgeway & Nakagawa, 2017). Consistent with this, Black raters on interview panels evaluate White candidates more favorably than Black candidates when there are also White evaluators, but not when all interviewers are Black (Shen et al., 2022). Evaluators from marginalized groups are more likely to express in-group preference when they are surrounded with like-minded others and group consensus about race and gender prejudices is weak (Ridgeway & Correll, 2006). The structural hierarchies in evaluator groups thus play an important role in determining evaluative outcomes by shaping the extent to which women and racial minorities deviate from patterns that reproduce inequality.

Overall, extensive work documents how demographic characteristics of evaluators impact their evaluative decisions. Women and racially marginalized groups evaluating demographically similar others encounter a unique tension between favoring those in their own gender or racial group and maintaining their own status position. There are numerous opportunities for future research that would deepen our understanding of how evaluators' demographic attributes shape their propensity to reproduce inequality. A particularly useful direction is to further explore when shared demographics mitigate versus accentuate inequality in evaluations. It seems that the social standing of underrepresented evaluators within the local context is a key predictor of whether they show support for similar others. Though members of marginalized groups are de facto perceived to be—and thus likely to perceive themselves as being—lower status, this varies across contexts. Uncovering the precise factors that overcome marginalized evaluators' concerns about status loss would thus be informative.

Relatedly, future work should focus on how women and racially marginalized individuals in decision-making roles can indirectly reduce inequality. For instance, it is plausible that exposure to evaluations by others can influence an evaluators' own assessments of a candidate. Increasing exposure to evaluations from in-group members might have the potential to de-bias assessments.

Evaluator expertise, evaluation experience, and achieved status

Research on evaluator characteristics and inequality has predominately focused on evaluator demographics, but scholars have recently begun to consider how other evaluator characteristics influence inequality. Three non-demographic characteristics in particular stand out: expertise based on prior experiences (with both having been an evaluator and having acquired domain-

specific knowledge), prior experiences as a candidate (having been previously evaluated) and achieved status stemming from past accomplishments.

First, evaluators with relevant expertise—stemming both from having previously been evaluators and from more general experience in a given domain—may be more objective in their evaluations because their expertise provides them with clarity about how to make assessments. People are generally motivated to satisfy the expectations associated with their roles and positions within organizations. That is, if they are tasked with being an evaluator, they will strive to evaluate effectively, for instance by seeking local cues for information on how to fulfill their role (Golden-Biddle & Rao, 1997). One source of such information comes from an individual's experience evaluating candidates, *especially within the same evaluation system*. These experiences offer the most relevant insights into what ought to be rewarded within that system. Though not looking at inequality per se, recent research shows that people update their approaches to evaluating as they gain experience within a given system: Initially, inexperienced evaluators tend to be harsher in their assessments, but as they gain experience they develop benchmarks that allow them to more accurately determine candidate quality (Bian et al., 2022).

Beyond this, other forms of experience in an evaluative system can impact decision-making processes, and shape expertise. Direct experience with a candidate from a marginalized group is linked to reduced biases when evaluators subsequently interact with that same candidate (Pager & Karafin, 2009; Sterling & Fernandez, 2018). Evidence that a candidate from a marginalized group is a higher performer can contradict and ultimately counteract negative biases an evaluator may hold about the candidate and thus lead them be to more equitable.

More commonly, expertise refers to an individual's general depth of knowledge acquired from their full range of experiences in an area. Research has shown that individuals lacking

domain-specific expertise are likely to turn to the assessments of others, essentially prioritizing the wisdom of the crowd over their own judgements of candidate quality (Botelho, 2024; Salganik & Watts, 2008). Although research has not looked at the relationship between domain-specific expertise and inequality, expertise has been tied to increased accuracy in evaluator's reviews of scientific grant proposals (Li, 2017) and to evaluators being less influenced by the evaluations of others when assessing candidates (Botelho, 2024).

Expertise offers clarity about how one ought to evaluate candidates and such clarity tends to reduce inequality (Anderson & Tomaskovic-Devey, 1995; DiPrete et al., 2010; see also subsection Evaluative processes, criteria, and tools). Experts—both those with extensive experience conducting evaluations and those with deep domain-specific knowledge—are better equipped to identify highly capable, or skilled, candidates in that domain. Thus, expertise empowers evaluators to ignore factors like candidates' gender and race enabling them to be more objective in their assessments.

Second, evaluators also come with varied prior experiences having been evaluated; in other words, prior experiences as a candidate will differ across evaluators. The prevalence of evaluations in professional contexts means that everyone experiences being evaluated at some point. Furthermore, evaluators are likely to have been assessed through the same evaluation system; for example, a hiring manager has been considered as a job candidate previously, and an evaluator reviewing grant applications has likely submitted a grant proposal for consideration. Evaluators draw on these experiences when assessing others, particularly within the same evaluative systems. This may both exacerbate or mitigate inequality. When evaluators' prior assessments do not correctly recognize their performance or contributions, for instance, they tend to evaluate others in kind, reproducing their own experiences as candidates (Abraham et al.,

41

2024). Interview evidence suggests that managers who have primarily experienced negative assessments throughout various assessments in their careers describe attempts to be objective when evaluating others (Castilla & Ranganathan, 2020).

Third, evaluators' achieved status also shapes the ways they evaluate others. Unlike status stemming from a person's ascribed characteristics like gender or race, achieved status refers to the social esteem a person derives from their performance, actions, and affiliations (Podolny, 2005). There is a vast organizational and sociological literature on the cumulative advantages of status, which contends that high-status individuals generally receive more credit, attention, and resources (Podolny, 2005). High-status evaluators are aware that their positive assessments are sought after and valuable to candidates (Podolny, 2005; Stuart et al., 1999). As a result high-status evaluators are likely to view their assessments as especially valid, which may affect how they approach the evaluation process, and how they incorporate gender and race into their evaluations (Botelho & Gertsberg, 2022).

Unlike status based on demographic characteristics, achieved status is sometimes unstable, meaning individuals can lose or gain status in subsequent periods. Although seldomly discussed, this is especially true when considering status within professional contexts, such as within organizations, where hierarchies are ever evolving with personnel changes and shifts in leadership. More generally, the halo effect from most awards and recognition likely dissipates over time. Thus, when maintaining one's status position is less certain, evaluators may be sensitized to the potential for status loss in ways that shape their evaluations (Botelho & Gertsberg, 2022). Given the general expectation that evaluation processes ought to be meritocratic (Castilla, 2008; Castilla & Benard, 2010; Mijs, 2021), evaluators are apt to assume

that objectivity is requisite for being considered a "good" evaluator—one worthy of earning and maintaining status distinctions.

There is now more work on how evaluator characteristics beyond demographics, including expertise, experience being evaluated, and achieved status, shape their behavior, but much more is needed to link these findings to inequality in evaluations. Scholarship on expertise has begun to show that an evaluator's domain expertise is related to their accuracy and tendency to be influenced by the assessments of others affecting stratification (Botelho, 2024); future work can further discern how domain expertise relates to an evaluator's reliance on gender and race. While individuals in positions of power and with experience (e.g., managers) commonly exhibit biases, experience does not always lead to expertise. Thus, it may be that domain expertise—and not simply broader experience evaluating others—leads evaluators to rely less on gender and race in their assessments.

Furthermore, researchers have yet to consider the effects of prior experiences being evaluated through a gendered or racialized lens. There is little doubt that evaluators from marginalized groups endure more unjust outcomes when they are being assessed, and that these negative experiences might shape their own evaluative judgments. One could, for instance, presume that negative experiences being evaluated might impact how evaluators take into account procedural fairness and distributive justice when evaluating others (Brockner & Wiesenfeld, 2019). The potential impact of these experiences is unclear: On the one hand, marginalized group members may reproduce the unjust outcomes they have experience, but on the other hand they may strive to remedy these inequities. Future research is thus needed to understand how evaluators from marginalized groups shape inequality.

Discussion and Conclusion

Evaluations are increasingly prevalent in determining the allocation of resources and rewards across professional contexts (Botelho, 2024; Botelho & Abraham, 2017; Bowers & Prato, 2018; Rivera & Tilcsik, 2019). As their use continues to increase and evolve, it is more important than ever to consider how these processes produce and reproduce inequality across professional contexts. A substantive and effective redressal of systemic inequality requires a rigorous, scholarly dissection of the complex drivers of inequality in evaluation processes. This review aims to provide such a redressal by generating an integrative framework. Our framework (depicted in Figure 1, with Table 1 providing further detail) identifies the overarching drivers associated with variations in inequality and allows us to unpack when and how these drivers attenuate or sustain observed inequality.

We introduce three core categories of drivers. First, prevailing beliefs—such as the increased prominence of meritocratic beliefs and beliefs stemming from identity-based social movements—affect the propensity for a candidate's gender or race to shape evaluations. Second, the design and structure of evaluation processes shape inequality through various features, including the degree to which evaluative processes are formalized as well as the specific tools used to interpret candidate quality. Third, characteristics of evaluators themselves can influence who they evaluate generously and who they evaluate more harshly. Below, we draw on the framework to expand on future research directions and consider implications for how the three core drivers interact, jointly producing inequality in practice.

Research Directions

Throughout the paper, we offer a roadmap highlighting promising directions for future research that directly relate to each of these three categories of drivers. Here, we further this discussion,

44

providing an overview of three avenues for future research that we see as most central. Specifically, researchers should (i) focus more attention on racial inequality in professional contexts as existing work has prioritized gender, (ii) deepen theory on how structural changes to evaluative systems affect outcomes, and (iii) consider cross-cultural examinations of how evaluations produce inequities.

First, there is a relative lack of attention on racial inequality in professional contexts. Noticeably more scholarly attention has been devoted to gender than to race. In our initial pass at identifying papers on inequality in evaluations within professional contexts, 53% were on gender versus 22% on race.[6] One possible explanation for this discrepancy in scholarly attention relates to the persistent underrepresentation of racially marginalized individuals, particularly Black workers, in "white-collar" professions: Black workers represent only 8% of the white-collar workforce and, similarly, only 7% of managers—a very low percentage given their share of the US population (Gee, 2018; McKinsey, 2021). Likewise, Black people in the US comprise only 3% of academics (National Center for Education Statistics, 2023) and 3% of business owners (Leppert, 2014).

Another possible explanation is that while stereotypes, norms, and status beliefs about gender apply broadly across populations, the same is not true for race. Much of the research on inequality draws on our understanding of gender beliefs (e.g., women as communal) and gendered social roles (e.g., women as caretakers). There is much greater ambiguity on how

---

[6] The initial set of relevant papers for this review were identified using Web of Science. We searched for all articles published—across all years—that contained the following search terms in the title, abstract, or keywords: "evaluation/rating" AND "inequality" or "gender/race" AND "inequality/bias." We then (1) targeted research published in top journals in management, organizational theory, and sociology, and (2) examined highly cited works to ensure our process had not missed any significant citations or outlets. This search left us with 132 articles. We coded papers as discussing race and/or gender by searching these two terms as well as synonyms in the title, abstract, and/or keywords for each paper. Reading these papers led us to other related research that was not captured in our initial search.

racialized beliefs and race-typed work shape evaluations of marginalized groups. For instance, common beliefs about Asians at work differ from those about Black people and Latinos (Zou & Cheryan, 2017). Thus, research on evaluations does not broadly inform inequality in evaluations across these groups. Future work should nonetheless aim to unpack the influence of racialized beliefs and race-typed professional contexts on assessments of marginalized candidates. Continuing to overlook the unique barriers facing racially marginalized groups will prevent us from remedying race-based inequalities in evaluations.

In the same vein, though not addressed in this paper, other forms of discrimination (e.g. disability, sexuality, age-based) have received even less attention, with only a small number of papers considering the specific social processes that uniquely shape them (Martin & North, 2022; Mishel, 2020; Rivera & Tilcsik, 2023). Notably, the underlying stereotypes that result in these different types of inequality vary widely and will intersect differently with the three core drivers of inequality we outline. The gendered nature of work will, for instance, uniquely shape the experiences of gay men, who may often be evaluated more positively for female-typed work (Pedulla, 2016).

Second, future research should explore the effect of emerging technologies and technology-mediated evaluations on inequality. Monumental shifts in work environments and professional infrastructures are in full force, particularly with the expansion of the gig economy, online labor markets, and the solidification of remote work. These changes affect all aspects of our work lives, including evaluations, which are increasingly conducted in the digital sphere or via technological interfaces. Even in workplaces where evaluators are assessing coworkers with whom they have in-person interactions, the actual assessments often happen on digital platforms—increasingly mimicking the types of evaluations commonly used in product markets.

In some cases, the entire experience with candidates is facilitated via technology. Assessments of workers in the gig economy, grant applicants, and entrepreneurs often occur online and without any direct personal interaction between evaluators and candidates.

Thus, we call for research on inequality in these technology-mediated evaluative systems. One possibility is that such evaluations may be more meritocratic as evaluator's assessments are not as heavily influenced by social ties (or in-person interactions), which may also make candidates' demographic attributes less salient. However, in these virtual cases evaluators often have less granular information with which to evaluate candidates and may thus be more reliant on gender and race. Additionally, evaluators may feel less accountable for their actions in technology-mediated evaluations, which could also lead to the reproduction of inequality.

Relatedly, emerging technologies, like artificial intelligence are increasingly dominating various aspects of evaluative processes (Kellogg et al., 2020; Rahman, 2021). For instance, recruiters sometimes narrow applicant pools through algorithmic assessment. The effects of these practices on inequality remain mixed (Cowgill & Tucker, 2023). On the one hand, algorithmic decision-making can reflect human biases and exacerbate inequality in evaluations.[7] On the other hand, distancing humans from assessments may curtail the propensity for gender and racial beliefs to shape these evaluations. Examinations of the many potential indirect effects of these technological shifts are also warranted. For instance, as automation and new technologies diminish the value of evaluators' expertise, one might expect experts to experience greater perceptions of threat, which could affect evaluative inequalities in divergent ways. On the one hand, this threat could lead evaluators, particularly those from majority groups, to protect their

---

[7] Algorithms are often trained on historical data, which may contain biases present in human decision-making processes. If historical decisions were influenced by factors like race or gender, the algorithm may learn and perpetuate these biases, leading to discriminatory outcomes.

status position by disadvantaging those from marginalized groups. On the other hand, this threat could have an inequality-reducing effect in cases where evaluators know they ought to be unbiased. Future research is thus needed to consider the distinct implications of technology-mediated evaluations and to inform how these systems can be designed to minimize inequality.

Third, the review focuses on evaluations in mostly American professional contexts. Perceptions about gender and race vary significantly by geographic area and are often intertwined with cross-cultural and ethnicity-based differences. While a direct examination of these variations is beyond the scope of our paper, there is a significant opportunity for future research. More generally, our understanding of inequality in evaluations primarily lies within Western contexts. Partially, this may be a product of population-level demographic differences, seeing as Eastern countries tend to have more demographically homogenous populations (Fisher, 2021). But more uniform national populations still evince ethnicity-, race-, and gender-based biases, and many of these are central, albeit unique, to these countries. Thus, at the most basic level, we recommend exploring how cultural differences impact evaluative decisions. For instance, extensive work in social psychology documents how behaviors like agency and independence are far more valued in Western than Eastern contexts. Demonstrating communality and respect for social hierarchies holds more esteem in Eastern contexts (Gelfand et al., 2011). Therefore, valued candidate characteristics will differ by country and can vary based upon gender and other demographic characteristics. Seeing as evaluations are entrenched in social and normative beliefs systems, future scholarship should explore the ramifications of different cultural schemas for evaluations.

The Bigger Picture: Interdependencies Across Drivers and Implementing the Framework

Beyond future research directions, it is critical to recognize that the three categories of drivers in the framework are not necessarily exclusive, and there are important interdependencies among them. Often, efforts to reform evaluative systems serve more as symbolic gestures of support for DEI rather than providing evaluators with guidance on how they ought to assign value to candidates. As a result, polices are implemented with little evidence of their actual effectiveness, instead providing only the illusion of change (Kalev et al., 2006; Kang et al., 2016; Kang & Kaplan, 2019). Our framework offers an actionable, research-backed blueprint for policymakers, organizational architects, and social stakeholders, that accounts for the complex social processes that perpetuate inequality. Practitioners, for instance, should consider characteristics of the evaluators—such as gender and expertise—when determining how evaluation processes are structured, as processes will be implemented differently depending upon evaluator characteristics. Implementing new design processes aimed at reducing inequality might be futile if an evaluator is not on board with the underlying motivation for the changes.

A key takeaway is that evaluations are inherently multilevel processes: Evaluators tend to operate within specific evaluative systems, which are often embedded within organizations, and these organizations exist within broader societal contexts. Thus, it is challenging to draw clear lines between the categories of drivers. When are evaluative criteria within a system not reflective of broader societal beliefs? When are evaluative processes not shaped by the evaluators designing these processes? Understanding the connections across these three areas is therefore critical for reducing inequality in practice. Effectively redressing inequality within any given system requires consideration of the interplay among these drivers—both structural and social—for shaping decision-making.

The greatest opportunity to reduce inequality in evaluations lies in considering the design and structure of evaluative processes embedded within social systems and utilized by evaluators with diverse backgrounds. For instance, companies could further regulate the use of intermediaries (e.g., search firms) for hiring seeing as these evaluator types have been found to discriminate against marginalized groups, even when the hiring firm espouses meritocratic values. Gendered and racialized beliefs are stubbornly sticky and trying to alter them is often far more complex than redesigning evaluative systems to account for them. It is only recently that scholars have begun to investigate the effects of design and structural processes in evaluations for inequality (e.g., connections across multiple evaluative stages, evaluation scales). With consideration for the other drivers, redesigning evaluation processes will likely be the most effective at reducing inequality and will likely be the easiest for firms to implement. It may be the case that these redesigns will be most effective in contexts where evaluators are already apt to be more equitable. However, with thorough consideration of prevailing beliefs and evaluator characteristics, there is the potential that these designs can reduce inequality even when this condition is not met.

Table 1. Foundational Concepts in the Theoretical Framework

		How Prevailing Beliefs Affect Inequality			
Underlying Drivers	**Foundational Concepts**	**Description**	**Influence on Inequality**	**Common Empirical Contexts**	**Key Citations**
Gendered and Racialized Nature of Professional Contexts	**Occupational Stereotypes and Fit**	Stereotypes and historical numerical representation lead to gender- or race-typing of specific fields, occupations, jobs, and tasks.	Demographic typing exacerbates inequality by creating expectations about which demographic groups are perceived as most suitable for specific work. High status, profitable work is commonly White- and/or male-typed, leading to marginalized group members being perceived as atypical and thus disadvantaged in those settings.	Evaluations across demographically-typed work contexts, including finance, technology (male-typed); nursing, human resources (female-typed).	Cejka & Eagly, 1999; Ridgeway, 2011; 2009; Williams, 1992; Wingfield & Chavez, 2020
	Leadership Prototypicality	Stereotypes of the ideal leader as a White man: intelligent, agentic, and confident.	Leadership evaluations typically favor White men over others, leading to disparities in perceptions of leadership capabilities, which in turn produce inequality in terms of who holds leadership positions.	Leadership evaluations across demographically-typed contexts, including business and sports	Marquez-Velarde et al., 2023; Rosette et al., 2008
Beliefs about Meritocracy	**Organizational Commitment to Meritocracy**	Organizations and management systems that rely on the notion that outcomes, such as wages, promotions, bonuses, and awards, should be based on candidate capabilities and performance.	Meritocratic beliefs *can* reduce inequality in evaluations but often also lead evaluators to assume that meritocracy has been achieved and thus to overlook systemic inequalities.	Wage determination, bonus pay systems, performance evaluations	Castilla & Benard, 2010; Mijs, 2021; Mun & Kodama, 2022
Shifts in Beliefs Resulting from Identity-based Social Movements	**Reputational Risks and Stigma Spillovers**	Shifts in evaluators' behaviors and outside perceptions following scandals or societal movements.	Evaluators alter behaviors to avoid reputational damage, which produces mixed results: It can lead them to either mitigate inequality or reinforce inequality (e.g., avoid marginalized group members).	Hiring post-Weinstein scandal, evaluations post-#MeToo movement	Bednar et al., 2022; McDonnell et al., 2021; Luo & Zhang, 2022
	Societal Events and Gendered / Racialized Beliefs	Societal events, such as crimes or protests, connected to gender and race shape evaluators' behaviors and perceptions.	Societal events can either reify biases, thereby increasing inequality, or raise awareness about biases, thereby reducing inequality, depending on the particular event.	Job evaluations post-community exposure to crimes, hiring post-Black Lives Matter protests	Chavez et al., 2022; Gorbatai et al., 2023; Mobasseri, 2019

Table 1. Foundational Concepts in the Theoretical Framework (continued)

How the Design and Structure of Evaluation Processes Shape Inequality					
Underlying Drivers	**Foundational Concepts**	**Description**	**Influence on Inequality**	**Common Empirical Contexts**	**Key Citations**
Candidate Pool Features	**Evaluator Cognitive Overload**	Challenges affecting evaluators when assessing large candidate pools and facing time/resource constraints.	Increased reliance on cognitive shortcuts, like stereotypes, when evaluators face cognitive overload reproduce inequality.	Venture capital funding, job promotions, gig work platforms	Brenčič, 2014; Iyengar & Lepper, 2000; Sherf et al., 2019
	Demographic Composition Dynamics	Makeup of candidate pools (i.e. more atypical candidates).	Higher presence of atypical candidates (e.g., female candidates for male-typed jobs) may both reduce or reinforce inequality.	Hiring for male-dominated jobs/industries, applications for atypical roles	Huffman & Cohen, 2004; Leung and Koppman, 2018; Robbins & DeNisi, 1993
Multistage Evaluation Processes	**Sequential Decision-Making**	Changes in context, evaluators, and criteria across evaluation stages.	Early stage evaluators' lack of expertise and reliance on audience inferences can introduce biases. In later stages with fewer candidates, more information available on candidates, and more experienced evaluators, there is less reliance on biases and thus less inequality.	Hiring processes, startup investment evaluations, prestigious awards and grants	Botelho & Abraham, 2017; Botelho & Chang, 2023; Castilla, 2008
Evaluative Criteria, Processes, and Tools	**Formalization vs. Discretion**	Extent to which evaluation criteria and processes are standardized or left to evaluator discretion.	Formalized criteria reduce uncertainty and inequality; however, formalization of biased criteria can exacerbate bias, making discretion a path to equitable outcomes.	Compensation systems, hiring, promotion processes, performance evaluations	Abraham, 2017; Castilla, 2015; Elvira & Graham, 2002
	Quantitative Metrics	Use of numerical ratings and scales vs. qualitative metrics (e.g., written feedback).	Quantitative metrics can reduce *or* perpetuate inequitable evaluative outcomes depending on their design.	Evaluations in educational contexts, customer ratings of workers, evaluations on platforms	Kiviat, 2019; Mazmanian & Beckman, 2018; Rivera & Tilcsik, 2019
Presence of Relevant Audiences	**Evaluator Accountability and Oversight**	Audience provides oversight in evaluative decision-making.	The need to justify decisions to others reduces inequality based on the belief that others expect evaluations to be unbiased.	Performance-reward systems, hiring and promotion decisions, peer review evaluations	Botelho & Gertsberg, 2022; Castilla, 2015; Chung et al., 2020
	Evaualuator Infers Audience Preferences	Evaluator strives to make a selection for an audience (i.e., needs approval from audience).	Assumptions about audience preferences and biases can exacerbate inequality.	Referrals, network-based hiring practices, and hiring through staffing firms	Abraham, 2020; Fernandez-Mateo & Fernandez, 2016; Fernandez-Mateo & King, 2011; Smith, 2005; 2008

Table 1. Foundational Concepts in the Theoretical Framework (continued)

How Evaluator Characteristics Influence Inequality					
Underlying Drivers	**Foundational Concepts**	**Description**	**Influence on Inequality**	**Common Empirical Contexts**	**Key Citations**
Demographic Characteristics: Gender and Race	**Ingroup Preference and Homophily**	Evaluators have preferences for demographically similar others.	Demographic similarity can lead to more equitable outcomes but can also reinforce inequality if evaluators are predominantly White/male.	Patient referrals, performance assessments, gender pay gap, hiring and promotions	Cohen & Broschak, 2013; Cohen & Huffman, 2007; McPherson et al., 200; Tajfel & Turner, 1979
	Marginalized Evaluators	Role of evaluators from marginalized groups in evaluative decision-making.	Awareness of systemic biases can lead to advocacy, but concerns about one's own status can lead evaluators to reproduce inequality.	Crowdfunding for female entrepreneurs, salary allocation by female supervisors	Abraham, 2017; Amanatullah & Morris, 2010; Duguid et al., 2012; Greenberg & Mollick, 2017; Snellman & Solal, 2023
Evaluator Expertise, Evaluation Experience, and Achieved Status	**Domain Expertise**	Relevant expertise of evaluator in the domain in which they are making an assessment.	Expertise may lead to more objective evaluations, reducing reliance on gender and race, and ultimately mitigating inequality.	Scientific grant proposal reviews, evaluations within the same system	Anderson & Tomaskovic-Devey, 1995; Bian et al., 2022; Botelho, 2024; Li 2017
	Maintaining Achieved Status	Evaluators are motivated to maintain/reinforce their social standing derived from their performance and affiliations.	High-status evaluators strive for objectivity to maintain their status, potentially reducing bias; conformity to status quo tends to reinforce inequality.	Professional recognition and awards, mixed demographic interview panels, salary decisions by female supervisors	Botelho & Gertsberg, 2022; Bowers & Prato, 2018; Ridgeway & Nakagawa, 2017; Shen et al., 2022
	Evaluator Experiences as a Candidate	Evaluator's draw on their own past experiences having been evaluated, both from the same and different evaluation systems.	Experiencing low evaluations can lead evaluators to make more equitable evaluations themselves while experiencing unfair evaluations can lead them to perpetuate inequalities.	Hiring managers evaluating job candidates	Abraham et al., 2024; Castilla & Ranganathan, 2020

References

Abraham, M. (2017). Pay Formalization Revisited: Considering the Effects of Manager Gender and Discretion on Closing the Gender Wage Gap. *Academy of Management Journal*, *60*(1), 29–54. https://doi.org/10.5465/amj.2013.1060

Abraham, M. (2020). Gender-role Incongruity and Audience-based Gender Bias: An Examination of Networking among Entrepreneurs. *Administrative Science Quarterly*, *65*(1), 151–180. https://doi.org/10.1177/0001839219832813

Abraham, M., & Botelho, T. L. (2024). Tuning Out Gender? Understanding the Heterogeneous Effects of Educational Status on Compensation with Implications for the Gender Pay Gap. *Working Paper.*

Abraham, M., Botelho, T. L., & Carter, J. T. (2024). Paying & Punishing it Forward: Misrecognized Evaluators Reproduce Unmeritocratic Peer Evaluations. *Academy of Management Proceedings*, *2024*(1), 14861. https://doi.org/10.5465/AMPROC.2024.141bp

Abraham, M., & Burbano, V. (2022). Congruence Between Leadership Gender and Organizational Claims Affects the Gender Composition of the Applicant Pool: Field Experimental Evidence. *Organization Science*, *33*(1), 393–413. https://doi.org/10.1287/orsc.2021.1442

Acker, J. (1990). Hierarchies, Jobs, Bodies: A Theory of Gendered Organizations. *Gender and Society*, *4*(2), 139–158.

Alon, S., & Tienda, M. (2007). Diversity, Opportunity, and the Shifting Meritocracy in Higher Education. *American Sociological Review*, *72*(4), 487–511. https://doi.org/10.1177/000312240707200401

Amanatullah, E. T., & Morris, M. W. (2010). Negotiating Gender Roles: Gender Differences in Assertive Negotiating are Mediated by Women's Fear of Backlash and Attenuated When Negotiating on Behalf of Others. *Journal of Personality and Social Psychology*, *98*(2), 256–267. https://doi.org/10.1037/a0017094

Anderson, C. D., & Tomaskovic-Devey, D. (1995). Patriarchal Pressures: An Exploration of Organizational Processes that Exacerbate and Erode Gender Earnings Inequality. *Work and Occupations*, *22*(3), 328–356. https://doi.org/10.1177/0730888495022003005

Ashforth, B. E., & Mael, F. (1989). Social Identity Theory and the Organization. *The Academy of Management Review*, *14*(1), 20–39. https://doi.org/10.2307/258189

Baron, J. N. (1984). Organizational Perspectives on Stratification. *Annual Review of Sociology*, *10*, 37–69.

Baron, J. N., Davis-Blake, A., & Bielby, W. T. (1986). The Structure of Opportunity: How Promotion Ladders Vary Within and Among Organizations. *Administrative Science Quarterly*, *31*(2), 248–273. https://doi.org/10.2307/2392790

Batruch, A., Jetten, J., Van De Werfhorst, H., Darnon, C., & Butera, F. (2023). Belief in School Meritocracy and the Legitimization of Social and Income Inequality. *Social Psychological and Personality Science*, *14*(5), 621–635. https://doi.org/10.1177/19485506221111017

Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/E/bo22415931.html

Beckman, C. M., & Phillips, D. J. (2005). Interorganizational Determinants of Promotion: Client Leadership and the Attainment of Women Attorneys. *American Sociological Review*, *70*(4), 678–701. https://doi.org/10.1177/000312240507000407

Bednar, M. K., Westphal, J. D., & McDonald, M. L. (2022). Birds of a Feather Flock (Even More) Together: An Intergroup Relations Perspective on How #MeToo-related Media Coverage Affects the Evaluation of Prospective Corporate Directors. *Strategic Management Journal*, *43*(11), 2313–2350. https://doi.org/10.1002/smj.3405

Behaghel, L., Crépon, B., & Le Barbanchon, T. (2015). Unintended Effects of Anonymous Résumés. *American Economic Journal: Applied Economics*, *7*(3), 1–27. https://doi.org/10.1257/app.20140185

Belliveau, M. A. (2012). Engendering Inequity? How Social Accounts Create vs. Merely Explain Unfavorable Pay Outcomes for Women. *Organization Science*, *23*(4), 1154–1174. https://doi.org/10.1287/orsc.1110.0691

Belz, A. P., Graddy-Reed, A., Hanewicz, I., & J. Terrile, R. (2022). Gender Differences in Peer Review of Innovation. *Strategic Entrepreneurship Journal*, *16*(2), 255–280. https://doi.org/10.1002/sej.1429

Berger, J., Osterloh, M., & Rost, K. (2020). Focal Random Selection Closes the Gender Gap in Competitiveness. *Science Advances*, *6*(47). https://doi.org/10.1126/sciadv.abb2142

Bian, J., Greenberg, J., Li, J., & Wang, Y. (2022). Good to Go First? Position Effects in Expert Evaluation of Early-Stage Ventures. *Management Science*, *68*(1), 300–315. https://doi.org/10.1287/mnsc.2021.4132

Bode, C., Rogan, M., & Singh, J. (2022). Up to No Good? Gender, Social Impact Work, and Employee Promotions. *Administrative Science Quarterly*, *67*(1), 82–130. https://doi.org/10.1177/00018392211020660

Bohnet, I., van Geen, A., & Bazerman, M. (2016). When Performance Trumps Gender Bias: Joint vs. Separate Evaluation. *Management Science*, *62*(5), 1225–1234. https://doi.org/10.1287/mnsc.2015.2186

Botelho, T. L. (2024). From Audience to Evaluator: When Visibility into Prior Evaluations Leads to Convergence or Divergence in Subsequent Evaluations Among Professionals. *Organization Science*. https://doi.org/10.1287/orsc.2017.11285

Botelho, T. L., & Abraham, M. (2017). Pursuing Quality: How Search Costs and Uncertainty Magnify Gender-based Double Standards in a Multistage Evaluation Process. *Administrative Science Quarterly*, *62*(4), 698–730. https://doi.org/10.1177/0001839217694358

Botelho, T. L., & Chang, M. (2023). The Evaluation of Founder Failure and Success by Hiring Firms: A Field Experiment. *Organization Science*, *34*(1), 484–508. https://doi.org/10.1287/orsc.2022.1592

Botelho, T. L., & Gertsberg, M. (2022). The Disciplining Effect of Status: Evaluator Status Awards and Observed Gender Bias in Evaluations. *Management Science*, *68*(7), 5311–5329. https://doi.org/10.1287/mnsc.2021.4161

Botelho, T. L., Gulati, R., & Sorenson, O. (2024). *The Sociology of Entrepreneurship Revisited*. https://doi.org/10.1146/annurev-soc-030222-014049

Bourdieu, P., & Passeron, J.-C. (1990). *Reproduction in education, society and culture* (Vol. 4). Sage. https://scholar.google.com/scholar?cluster=15230847337768080422&hl=en&oi=scholarr

Bowers, A., & Prato, M. (2018). The Structural Origins of Unearned Status: How Arbitrary Changes in Categories Affect Status Position and Market Impact. *Administrative Science Quarterly*, *63*(3), 668–699. https://doi.org/10.1177/0001839217727706

Brenčič, V. (2014). Search Online: Evidence from Acquisition of Information on Online Job Boards and Resume Banks. *Journal of Economic Psychology*, *42*, 112–125. https://doi.org/10.1016/j.joep.2014.02.003

Brezina, T., & Winder, K. (2003). Economic Disadvantage, Status Generalization, and Negative Racial Stereotyping by White Americans. *Social Psychology Quarterly*, *66*(4), 402–418. https://doi.org/10.2307/1519837

Bridges, W. P., & Nelson, R. L. (1989). Markets in Hierarchies: Organizational and Market Influences on Gender Inequality in a State Pay System. *American Journal of Sociology*, *95*(3), 616–658. https://doi.org/10.1086/229328

Briscoe, F., & Joshi, A. (2017). Bringing the Boss's Politics In: Supervisor Political Ideology and the Gender Gap in Earnings. *Academy of Management Journal*, *60*(4), 1415–1441. https://doi.org/10.5465/amj.2016.0179

Briscoe, F., & Safford, S. (2008). The Nixon-in-China Effect: Activism, Imitation, and the Institutionalization of Contentious Practices. *Administrative Science Quarterly*, *53*(3), 460–491. https://doi.org/10.2189/asqu.53.3.460

Brockner, J., & Wiesenfeld, B. M. (2019). Organizational Justice is Alive and Well and Living Elsewhere (But Not Too Far Away). In *Social Psychology and Justice*. Routledge.

Brooks, A. W., Huang, L., Kearney, S. W., & Murray, F. E. (2014). Investors Prefer Entrepreneurial Ventures Pitched by Attractive Men. *Proceedings of the National Academy of Sciences*, *111*(12), 4427–4431. https://doi.org/10.1073/pnas.1321202111

Burrell, J., & Fourcade, M. (2021). The Society of Algorithms. *Annual Review of Sociology*, *47*(1), 213–237. https://doi.org/10.1146/annurev-soc-090820-020800

Campero, S., & Fernandez, R. M. (2019). Gender Composition of Labor Queues and Gender Disparities in Hiring. *Social Forces*, *97*(4), 1487–1516. https://doi.org/10.1093/sf/soy097

Carnahan, S., & Greenwood, B. N. (2018). Managers' Political Beliefs and Gender Inequality among Subordinates: Does His Ideology Matter More Than Hers? *Administrative Science Quarterly*, *63*(2), 287–322. https://doi.org/10.1177/0001839217708780

Castilla, E. J. (2008). Gender, Race, and Meritocracy in Organizational Careers. *American Journal of Sociology*, *113*(6), 1479–1526. https://doi.org/10.1086/588738

Castilla, E. J. (2015). Accounting for the Gap: A Firm Study Manipulating Organizational Accountability and Transparency in Pay Decisions. *Organization Science*, *26*(2), 311–333. https://doi.org/10.1287/orsc.2014.0950

Castilla, E. J., & Benard, S. (2010). The Paradox of Meritocracy in Organizations. *Administrative Science Quarterly*, *55*(4), 543–676. https://doi.org/10.2189/asqu.2010.55.4.543

Castilla, E. J., & Ranganathan, A. (2020). The Production of Merit: How Managers Understand and Apply Merit in the Workplace. *Organization Science*, *31*(4), 909–935. https://doi.org/10.1287/orsc.2019.1335

Castilla, E. J., & Rissing, B. A. (2019). Best in Class: The Returns on Application Endorsements in Higher Education. *Administrative Science Quarterly*, *64*(1), 230–270.

Cejka, M. A., & Eagly, A. H. (1999). Gender-Stereotypic Images of Occupations Correspond to the Sex Segregation of Employment. *Personality and Social Psychology Bulletin*, *25*(4), 413–423. https://doi.org/10.1177/0146167299025004002

Chan, C. K., & Anteby, M. (2016). Task Segregation as a Mechanism for Within-job Inequality: Women and Men of the Transportation Security Administration. *Administrative Science Quarterly*, *61*(2), 184–216.

Chatman, J. A. (1991). Matching People and Organizations: Selection and Socialization in Public Accounting Firms. *Administrative Science Quarterly*, *36*(3), 459–484. https://doi.org/10.2307/2393204

Chavez, K., Weisshaar, K., & Cabello-Hutt, T. (2022). Gender and Racial Discrimination in Hiring Before and During the COVID-19 Pandemic: Evidence from a Field Experiment of Accountants, 2018–2020. *Work and Occupations*, *49*(3), 275–315. https://doi.org/10.1177/07308884221094539

Chung, K., Kim, K., & Lim, N. (2020). Social Structures and Reputation in Expert Review Systems. *Management Science*, *66*(7), 3249–3276. https://doi.org/10.1287/mnsc.2019.3347

Cohen, L., & Broschak, J. (2013). Whose Jobs Are These? The Impact of the Proportion of Female Managers on the Number of New Management Jobs Filled by Women versus Men. *Administrative Science Quarterly*, *58*(4), 509–541. https://doi.org/10.1177/0001839213504403

Cohen, L. E., Broschak, J. P., & Haveman, H. A. (1998). And Then There were More? The Effect of Organizational Sex Composition on the Hiring and Promotion of Managers. *American Sociological Review*, *63*(5), 711–727. https://doi.org/10.2307/2657335

Cohen, P., & Huffman, M. (2007). Working for the Woman? Female Managers and the Gender Wage Gap. *American Sociological Review*, *72*(5), 681–704. https://doi.org/10.1177/000312240707200502

Conzon, V. M. (2023). The Equality Policy Paradox: Gender Differences in How Managers Implement Gender Equality–Related Policies. *Administrative Science Quarterly*, *68*(3), 648–690. https://doi.org/10.1177/00018392231174235

Correll, S. J., & Benard, S. (2006). Biased Estimators? Comparing Status and Statistical Theories of Gender Discrimination. In *Advances in Group Processes* (Vol. 23, pp. 89–116). Emerald (MCB UP). https://doi.org/10.1016/S0882-6145(06)23004-2

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a Job: Is There a Motherhood Penalty? *American Journal of Sociology*, *112*(5), 1297–1338. https://doi.org/10.1086/511799

Correll, S. J., & Ridgeway, C. L. (2003). Expectation States Theory. In *Handbook of social psychology* (pp. 29–51). Kluwer Academic/Plenum Publishers.

Correll, S. J., Ridgeway, C. L., Zuckerman, E. W., Jank, S., Jordan-Bloch, S., & Nakagawa, S. (2017). It's the Conventional Thought That Counts: How Third-Order Inference Produces Status Advantage. *American Sociological Review*, *82*(2), 297–327. https://doi.org/10.1177/0003122417691503

Correll, S. J., Weisshaar, K. R., Wynn, A. T., & Wehner, J. D. (2020). Inside the Black Box of Organizational Life: The Gendered Language of Performance Assessment. *American Sociological Review*, *85*(6), 1022–1050. https://doi.org/10.1177/0003122420962080

Cowgill, B., & Tucker, C. (2023). Economics, Fairness and Algorithmic Bias. *Working Paper*.

Darity, W. A., Hamilton, D., Myers, S. L., Price, G. N., & Xu, M. (2022). Racial Differences in Time at Work Not Working. *ILR Review*, *75*(3), 552–572. https://doi.org/10.1177/00197939211063863

Davis, G. F., King, B., & Soule, S. A. (2022). *Do Social Movements Improve Corporate Behavior? A Discussion and Research Agenda*. Rutgers Business Review. https://papers.ssrn.com/abstract=4317261

de Vaan, M., & Stuart, T. (2022). Gender in the Markets for Expertise. *American Sociological Review*, *87*(3), 443–477. https://doi.org/10.1177/00031224221087374

Dencker, J. C. (2008). Corporate Restructuring and Sex Differences in Managerial Promotion. *American Sociological Review*, *73*(3), 455–476.

Devine, P. G. (2001). Implicit Prejudice and Stereotyping: How Automatic are They? Introduction to the Special Section. *Journal of Personality and Social Psychology*, *81*(5), 757–759. https://doi.org/10.1037/0022-3514.81.5.757

DiPrete, T. A., Eirich, G. M., & Pittinsky, M. (2010). Compensation Benchmarking, Leapfrogs, and the Surge in Executive Pay. *American Journal of Sociology*, *115*(6), 1671–1712. https://doi.org/10.1086/652297

DiPrete, T. A., & Soule, W. T. (1988). Gender and Promotion in Segmented Job Ladder Systems. *American Sociological Review*, *53*(1), 26–40. https://doi.org/10.2307/2095730

Dobbin, F., Schrage, D., & Kalev, A. (2015). Rage against the Iron Cage: The Varied Effects of Bureaucratic Personnel Reforms on Diversity. *American Sociological Review*, *80*(5), 1014–1044. https://doi.org/10.1177/0003122415596416

Doering, L., & Thébaud, S. (2017). The Effects of Gendered Occupational Roles on Men's and Women's Workplace Authority: Evidence from Microfinance. *American Sociological Review*, *82*(3), 542–567. https://doi.org/10.1177/0003122417703087

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science*, *11*(4), 315–319. https://doi.org/10.1111/1467-9280.00262

Duguid, M., Loyd, D. L., & Tolbert, P. S. (2012). The Impact of Categorical Status, Numeric Representation, and Work Group Prestige on Preference for Demographically Similar Others: A Value Threat Approach. *Organization Science*, *23*(2), 386–401.

Dupree, C. H., Torrez, B., Obioha, O., & Fiske, S. T. (2021). Race–Status Associations: Distinct Effects of Three Novel Measures Among White and Black Perceivers. *Journal of Personality and Social Psychology*, *120*(3), 601–625. https://doi.org/10.1037/pspa0000257

Eagly, A. H., & Koenig, A. M. (2008). Gender Prejudice: On the Risks of Occupying Incongruent Roles. In *Beyond common sense: Psychological science in the courtroom* (pp. 63–81). Blackwell Publishing.

EEOC. (2023). *Prohibited Employment Policies/Practices*. US EEOC. https://www.eeoc.gov/prohibited-employment-policiespractices

Ellemers, N. (2018). Gender Stereotypes. *Annual Review of Psychology*, *69*, 275–298. https://doi.org/10.1146/annurev-psych-122216-011719

Elvira, M. M., & Graham, M. E. (2002). Not Just a Formality: Pay System Formalization and Sex-Related Earnings Effects. *Organization Science*, *13*(6), 601–617. https://doi.org/10.1287/orsc.13.6.601.499

Ely, R. J. (1995). The Power in Demography: Women's Social Constructions of Gender Identity at Work. *The Academy of Management Journal*, *38*(3), 589–634. https://doi.org/10.2307/256740

England, P., Levine, A., & Mishel, E. (2020). Progress Toward Gender Equality in The United States has Slowed or Stalled. *Proceedings of the National Academy of Sciences*, *117*(13), 6990–6997. https://doi.org/10.1073/pnas.1918891117

Espeland, W. N., & Sauder, M. (2007). Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology*, *113*(1), 1–40. https://doi.org/10.1086/517897

Fernandez, R. M., Castilla, E. J., & Moore, P. (2000). Social Capital at Work: Networks and Employment at a Phone Center. *American Journal of Sociology*, *105*(5), 1288–1356. https://doi.org/10.1086/210432

Fernandez, R. M., & Sosa, M. L. (2005). Gendering the Job: Networks and Recruitment at a Call Center. *American Journal of Sociology*, *111*(3), 859–904. https://doi.org/10.1086/497257

Fernandez-Mateo, I., & Fernandez, R. M. (2016). Bending the Pipeline? Executive Search and Gender Inequality in Hiring for Top Management Jobs. *Management Science*, *62*(12), 3636–3655. https://doi.org/10.1287/mnsc.2015.2315

Fernandez-Mateo, I., & Kaplan, S. (2018). Gender and Organization Science: Introduction to a Virtual Special Issue. *Organization Science*, *29*(6), 1229–1236. https://doi.org/10.1287/orsc.2018.1249

Fernandez-Mateo, I., & King, Z. (2011). Anticipatory Sorting and Gender Segregation in Temporary Employment. *Management Science*, *57*(6), 989–1008. https://doi.org/10.1287/mnsc.1110.1333

Fisher, M. (2021, December 1). A Revealing Map of the World's Most and Least Ethnically Diverse Countries. *Washington Post*. https://www.washingtonpost.com/news/worldviews/wp/2013/05/16/a-revealing-map-of-the-worlds-most-and-least-ethnically-diverse-countries/

Fiske, S. T., & Stevens, L. E. (1993). What's so Special about Sex? Gender Stereotyping and Discrimination. In *Gender issues in contemporary society* (pp. 173–196). Sage Publications, Inc.

Foschi, M. (2000). Double Standards for Competence: Theory and Research. *Annual Review of Sociology*, *26*, 21–42. https://doi.org/10.1146/annurev.soc.26.1.21

Gee, M. (2018, February 28). Why Aren't Black Employees Getting More White-Collar Jobs? *Harvard Business Review*. https://hbr.org/2018/02/why-arent-black-employees-getting-more-white-collar-jobs

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., Aycan, Z., Boehnke, K., Boski, P., Cabecinhas, R., Chan, D., Chhokar, J., D'Amato, A., Subirats Ferrer, M., Fischlmayr, I. C., … Yamaguchi, S. (2011). Differences Between Tight and Loose Cultures: A 33-Nation Study. *Science*, *332*(6033), 1100–1104. https://doi.org/10.1126/science.1197754

Golden-Biddle, K., & Rao, H. (1997). Breaches in the Boardroom: Organizational Identity and Conflicts of Commitment in a Nonprofit Organization. *Organization Science*, *8*(6), 593–611. https://doi.org/10.1287/orsc.8.6.593

Goldin, C., & Rouse, C. (2000). Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians. *American Economic Review*, *90*(4), 715–741. https://doi.org/10.1257/aer.90.4.715

Gorbatai, A., Younkin, P., & Burtch, G. (2023). Collateral Damage: The Relationship Between High-Salience Events and Variation in Racial Discrimination. *Organization Science*, *34*(6), 2250–2271. https://doi.org/10.1287/orsc.2021.1527

Gorman, E. H. (2005). Gender Stereotypes, Same-Gender Preferences, and Organizational Variation in the Hiring of Women: Evidence from Law Firms. *American Sociological Review*, *70*(4), 702–728.

Greenberg, J., & Mollick, E. (2017). Activist Choice Homophily and the Crowdfunding of Female Founders. *Administrative Science Quarterly*, *62*(2), 341–374. https://doi.org/10.1177/0001839216678847

Heilman, M. E. (1983). Sex Bias in Work Aettings: The Lack of Fit Model. *Research in Organizational Behavior*, *5*, 269–298.

Heilman, M. E., & Haynes, M. C. (2005). No Credit Where Credit Is Due: Attributional Rationalization of Women's Success in Male-Female Teams. *Journal of Applied Psychology*, *90*(5), 905–916. https://doi.org/10.1037/0021-9010.90.5.905

Hochschild, J. L. (1995). *Facing Up to the American Dream: Race, Class, and the Soul of the Nation*. Princeton University Press. https://www.jstor.org/stable/j.ctt7rhtn

Huang, L., Joshi, P., Wakslak, C., & Wu, A. (2021). Sizing up Entrepreneurial Potential: Gender Differences in Communication and Investor Perceptions of Long-Term Growth and Scalability: Academy of Management Journal. *Academy of Management Journal*, *64*(3), 716–740. https://doi.org/10.5465/amj.2018.1417

Huffman, M. L., & Cohen, P. N. (2004). Racial Wage Inequality: Job Segregation and Devaluation across U.S. Labor Markets. *American Journal of Sociology*, *109*(4), 902–936. https://doi.org/10.1086/378928

In Corporate America, Black Senior Leadership Remains Scarce. (2021, March 11). McKinsey. https://www.mckinsey.com/featured-insights/sustainable-inclusive-growth/chart-of-the-day/in-corporate-america-black-senior-leadership-remains-scarce

Iyengar, S. S., & Lepper, M. R. (2000). When Choice is Demotivating: Can One Desire too Much of a Good Thing? *Journal of Personality and Social Psychology*, *79*(6), 995–1006. https://doi.org/10.1037/0022-3514.79.6.995

Jensen, K., Kovács, B., & Sorenson, O. (2018). Gender differences in obtaining and maintaining patent rights. *Nature Biotechnology*, *36*, 307–309. https://doi.org/10.1038/nbt.4120

Jost, J. T. (2020). *A Theory of System Justification*. Harvard University Press. https://www.hup.harvard.edu/books/9780674244658

Kaiser, C. R., Major, B., Jurcevic, I., Dover, T. L., Brady, L. M., & Shapiro, J. R. (2013). Presumed Fair: Ironic Effects of Organizational Diversity Structures. *Journal of Personality and Social Psychology*, *104*(3), 504–519. https://doi.org/10.1037/a0030838

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. https://doi.org/10.1177/000312240607100404

Kang, S. K., DeCelles, K. A., Tilcsik, A., & Jun, S. (2016). Whitened Résumés: Race and Self-Presentation in the Labor Market. *Administrative Science Quarterly*, *61*(3), 469–502. https://doi.org/10.1177/0001839216639577

Kang, S. K., & Kaplan, S. (2019). Working toward gender diversity and inclusion in medicine: Myths and solutions. *The Lancet*, *393*(10171), 579–586. https://doi.org/10.1016/S0140-6736(18)33138-6

Kanter, R. M. (1977). *Men and women of the corporation*. Basic books.

Kaufman, R. L. (2002). Assessing Alternative Perspectives on Race and Sex Employment Segregation. *American Sociological Review*, *67*(4), 547–572. https://doi.org/10.1177/000312240206700404

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at Work: The New Contested Terrain of Control. *Academy of Management Annals*, *14*(1), 366–410. https://doi.org/10.5465/annals.2018.0174

King, E. B., Mendoza, S. A., Madera, J. M., Hebl, M. R., & Knight, J. L. (2006). What's in a Name? A Multiracial Investigation of the Role of Occupational Stereotypes in Selection Decisions. *Journal of Applied Social Psychology*, *36*(5), 1145–1159. https://doi.org/10.1111/j.0021-9029.2006.00035.x

Kirgios, E. L., Chang, E. H., & Milkman, K. L. (2020). Going it Alone: Competition Increases the Attractiveness of Minority Status. *Organizational Behavior and Human Decision Processes*, *161*, 20–33. https://doi.org/10.1016/j.obhdp.2020.03.009

Kiviat, B. (2019). The Moral Limits of Predictive Practices: The Case of Credit-Based Insurance Scores. *American Sociological Review*, *84*(6), 1134–1158. https://doi.org/10.1177/0003122419884917

Kluegel, J. R. (1990). Trends in Whites' Explanations of the Black-White Gap in Socioeconomic Status, 1977-1989. *American Sociological Review*, *55*(4), 512–525. https://doi.org/10.2307/2095804

Kmec, J. A. (2008). The Process of Sex Segregation in a Gender-Typed Field: The Case of Male Nurses. *Sociological Perspectives*, *51*(2), 259–279. https://doi.org/10.1525/sop.2008.51.2.259

Kovács, B., & Sharkey, A. J. (2014). The Paradox of Publicity: How Awards Can Negatively Affect the Evaluation of Quality. *Administrative Science Quarterly*, *59*(1), 1–33. https://doi.org/10.1177/0001839214523602

Kulich, C., Trojanowski, G., Ryan, M. K., Haslam, S. A., & Renneboog, L. D. R. (2011). Who Gets the Carrot and Who Gets the Stick? Evidence of Gender Disparities in Executive Remuneration. *Strategic Management Journal*, *32*(3), 301–321.

Lamont, M. (2012). Toward a Comparative Sociology of Valuation and Evaluation. *Annual Review of Sociology*, *38*(Volume 38, 2012), 201–221. https://doi.org/10.1146/annurev-soc-070308-120022

Lane, J. N., Lakhani, K. R., & Fernandez, R. M. (2023). Setting Gendered Expectations? Recruiter Outreach Bias in Online Tech Training Programs. *Organization Science*. https://doi.org/10.1287/orsc.22.16499

Lee, M., & Huang, L. (2018). Gender Bias, Social Impact Framing, and Evaluation of Entrepreneurial Ventures. *Organization Science*, *29*(1), 1–16. https://doi.org/10.1287/orsc.2017.1172

Leppert, R. (2014, February 16). A Look at Black-Owned Businesses in the U.S. *Pew Research Center*. https://www.pewresearch.org/short-reads/2024/02/16/a-look-at-black-owned-businesses-in-the-us/

Lerner, M. J. (1980). The Belief in a Just World. In M. J. Lerner (Ed.), *The Belief in a Just World: A Fundamental Delusion* (pp. 9–30). Springer US. https://doi.org/10.1007/978-1-4899-0448-5_2

Leslie, L. M., Manchester, C. F., & Dahm, P. C. (2017). Why and When Does the Gender Gap Reverse? Diversity Goals and the Pay Premium for High Potential Women. *Academy of Management Journal*, *60*(2), 402–432. https://doi.org/10.5465/amj.2015.0195

Leung, M. D., & Koppman, S. (2018). Taking a Pass: How Proportional Prejudice and Decisions Not to Hire Reproduce Gender Segregation. *American Journal of Sociology*, *124*(3), 762–813. https://doi.org/10.1086/700677

Li, D. (2017). Expertise versus Bias in Evaluation: Evidence from the NIH. *American Economic Journal: Applied Economics*, *9*(2), 60–92. https://doi.org/10.1257/app.20150421

Luo, H., & Zhang, L. (2022). Scandal, Social Movement, and Change: Evidence from #MeToo in Hollywood. *Management Science*, *68*(2), 1278–1296. https://doi.org/10.1287/mnsc.2021.3982

Magee, J. C., & Galinsky, A. D. (2008). Social Hierarchy: The Self-Reinforcing Nature of Power and Status. *Academy of Management Annals*, *2*(1), 351–398. https://doi.org/10.5465/19416520802211628

Marquez-Velarde, G., Grashow, R., Glass, C., Blaschke, A. M., Gillette, G., Taylor, H. A., & Whittington, A. J. (2023). The Paradox of Integration: Racial Composition of NFL Positions from 1960 to 2020. *Sociology of Race and Ethnicity*, *9*(4), 451–469. https://doi.org/10.1177/23326492231182597

Martin, A. E., & North, M. S. (2022). Equality for (Almost) All: Egalitarian Advocacy Predicts Lower Endorsement of Sexism and Racism, but not Ageism. *Journal of Personality and Social Psychology*, *123*(2), 373–399. https://doi.org/10.1037/pspi0000262

Mazmanian, M., & Beckman, C. M. (2018). "Making" Your Numbers: Engendering Organizational Control Through a Ritual of Quantification. *Organization Science*, *29*(3), 357–379. https://doi.org/10.1287/orsc.2017.1185

McDonnell, M.-H., Odziemkowska, K., & Pontikes, E. (2021). Bad Company: Shifts in Social Activists' Tactics and Resources After Industry Crises. *Organization Science*, *32*(4), 1033–1055. https://doi.org/10.1287/orsc.2020.1410

Melamed, D., Munn, C. W., Barry, L., Montgomery, B., & Okuwobi, O. F. (2019). Status Characteristics, Implicit Bias, and the Production of Racial Inequality. *American Sociological Review*, *84*(6), 1013–1036.

Merluzzi, J., & Phillips, D. J. (2016). The Specialist Discount: Negative Returns for MBAs with Focused Profiles in Investment Banking. *Administrative Science Quarterly*, *61*(1), 87–124. https://doi.org/10.1177/0001839215610365

Mijs, J. J. B. (2021). The Paradox of Inequality: Income Inequality and Belief in Meritocracy go Hand in Hand. *Socio-Economic Review*, *19*(1), 7–35. https://doi.org/10.1093/ser/mwy051

Milliken, F. J., & Martins, L. L. (1996). Searching for Common Threads: Understanding the Multiple Effects of Diversity in Organizational Groups. *Academy of Management Review*, *21*(2), 402–433. https://doi.org/10.5465/amr.1996.9605060217

Mishel, E. (2020). Contextual Prejudice: How Occupational Context and Stereotypes Shape Bias against Gay and Lesbian Employees. *Social Currents*, *7*(4), 371–391. https://doi.org/10.1177/2329496520919912

Mobasseri, S. (2019). Race, Place, and Crime: How Violent Crime Events Affect Employment Discrimination. *American Journal of Sociology*, *125*(1), 63–104. https://doi.org/10.1086/703883

Moss, P., & Tilly, C. (2001). *Stories Employers Tell: Race, Skill, and Hiring in America*. Russell Sage Foundation. https://www.jstor.org/stable/10.7758/9781610444101

Moss-Racusin, C. A., & Rudman, L. A. (2010). Disruptions in Women's Self-Promotion: The Backlash Avoidance Model. *Psychology of Women Quarterly*, *34*(2), 186–202. https://doi.org/10.1111/j.1471-6402.2010.01561.x

Mun, E., & Kodama, N. (2022). Meritocracy at Work?: Merit-Based Reward Systems and Gender Wage Inequality. *Social Forces*, *100*(4), 1561–1591. https://doi.org/10.1093/sf/soab083

National Center for Education Statistics. (2023). *The NCES Fast Facts Tool provides quick answers to many education questions (National Center for Education Statistics).* https://nces.ed.gov/fastfacts/display.asp?id=61

Neumark, D., Bank, R. J., & Van Nort, K. D. (1996). Sex Discrimination in Restaurant Hiring: An Audit Study. *The Quarterly Journal of Economics*, *111*(3), 915–941. https://doi.org/10.2307/2946676

Niessen-Ruenzi, A., & Ruenzi, S. (2019). Sex Matters: Gender Bias in the Mutual Fund Industry. *Management Science*, *65*(7), 3001–3025. https://doi.org/10.1287/mnsc.2017.2939

Niven, D. (2020). Stifling Workplace Activism: The Consequences of Anthem Protests for NFL Players. *Social Science Quarterly*, *101*(2), 641–655. https://doi.org/10.1111/ssqu.12756

O'Reilly, C. A., & Chatman, J. A. (1996). Culture as Social Control: Corporations, Cults, and Commitment. In *Research in Organizational Behavior: An Annual Series of Analytical Essays and Critical Reviews* (Vol. 18, pp. 157–200). Elsevier Science/JAI Press.

Pager, D., Bonikowski, B., & Western, B. (2009). Discrimination in a Low-Wage Labor Market: A Field Experiment. *American Sociological Review*, *74*(5), 777–799. https://doi.org/10.1177/000312240907400505

Pager, D., & Karafin, D. (2009). Bayesian Bigot? Statistical Discrimination, Stereotypes, and Employer Decision Making. *The ANNALS of the American Academy of Political and Social Science*, *621*(1), 70–93. https://doi.org/10.1177/0002716208324628

Pager, D., & Pedulla, D. S. (2015). Race, Self-Selection, and the Job Search Process. *American Journal of Sociology*, *120*(4), 1005–1054. https://doi.org/10.1086/681072

Pedulla, D. S. (2016). Penalized or Protected? Gender and the Consequences of Nonstandard and Mismatched Employment Histories. *American Sociological Review*, *81*(2), 262–289.

Pedulla, D. S., Allen, S., & Baer-Bositis, L. (2023). Can Customers Affect Racial Discrimination in Hiring? *Social Psychology Quarterly*, *86*(1), 30–52. https://doi.org/10.1177/01902725221109533

Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. *The American Economic Review*, *62*(4), 659–661.

Podolny, J. M. (2005). *Status Signals: A Sociological Study of Market Competition*. Princeton University Press. https://www.jstor.org/stable/j.ctt7sbsh

Rahman, H. A. (2021). The Invisible Cage: Workers' Reactivity to Opaque Algorithmic Evaluations. *Administrative Science Quarterly*, *66*(4), 945–988. https://doi.org/10.1177/00018392211010118

Ray, V. (2019). A Theory of Racialized Organizations. *American Sociological Review*, *84*(1), 26–53. https://doi.org/10.1177/0003122418822335

Reskin, B. F. (2000). The Proximate Causes of Employment Discrimination. *Contemporary Sociology*, *29*(2), 319–328. https://doi.org/10.2307/2654387

Ridgeway, C. L. (1991). The Social Construction of Status Value: Gender and Other Nominal Characteristics. *Social Forces*, *70*(2), 367–386. https://doi.org/10.2307/2580244

Ridgeway, C. L. (1997). Interaction and the Conservation of Gender Inequality: Considering Employment. *American Sociological Review*, *62*(2), 218–235. https://doi.org/10.2307/2657301

Ridgeway, C. L. (2009). Framed Before We Know It: How Gender Shapes Social Relations. *Gender & Society*, *23*(2), 145–160. https://doi.org/10.1177/0891243208330313

Ridgeway, C. L. (2011). *Framed by Gender: How Gender Inequality Persists in the Modern World*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199755776.001.0001

Ridgeway, C. L., Backor, K., Li, Y. E., Tinkler, J. E., & Erickson, K. G. (2009). How Easily Does a Social Difference Become a Status Distinction? Gender Matters. *American Sociological Review*, *74*(1), 44–62. https://doi.org/10.1177/000312240907400103

Ridgeway, C. L., Boyle, E. H., Kuipers, K. J., & Robinson, D. T. (1998). How Do Status Beliefs Develop? The Role of Resources and Interactional Experience. *American Sociological Review*, *63*(3), 331–350. https://doi.org/10.2307/2657553

Ridgeway, C. L., & Correll, S. J. (2006). Consensus and the Creation of Status Beliefs. *Social Forces*, *85*(1), 431–453.

Ridgeway, C. L., & Erickson, K. G. (2000). Creating and Spreading Status Beliefs. *American Journal of Sociology*, *106*(3), 579–615. https://doi.org/10.1086/318966

Ridgeway, C. L., & Nakagawa, S. (2017). Is Deference the Price of Being Seen as Reasonable? How Status Hierarchies Incentivize Acceptance of Low Status. *Social Psychology Quarterly*, *80*(2), 132–152. https://doi.org/10.1177/0190272517695213

Rivera, L. A. (2012). Hiring as Cultural Matching: The Case of Elite Professional Service Firms. *American Sociological Review*, *77*(6), 999–1022. https://doi.org/10.1177/0003122412463213

Rivera, L. A. (2020). Employer Decision Making. *Annual Review of Sociology*, *46*(1), 215–232. https://doi.org/10.1146/annurev-soc-121919-054633

Rivera, L. A., & Tilcsik, A. (2016). Class Advantage, Commitment Penalty: The Gendered Effect of Social Class Signals in an Elite Labor Market. *American Sociological Review*, *81*(6), 1097–1131. https://doi.org/10.1177/0003122416668154

Rivera, L. A., & Tilcsik, A. (2019). Scaling Down Inequality: Rating Scales, Gender Bias, and the Architecture of Evaluation. *American Sociological Review*, *84*(2), 248–274. https://doi.org/10.1177/0003122419833601

Rivera, L. A., & Tilcsik, A. (2023). Not in My Schoolyard: Disability Discrimination in Educational Access. *American Sociological Review*, *88*(2), 284–321. https://doi.org/10.1177/00031224221150433

Robbins, T. L., & DeNisi, A. S. (1993). Moderators of Sex Bias in the Performance Appraisal Process: A Cognitive Analysis. *Journal of Management*, *19*(1), 113–126. https://doi.org/10.1177/014920639301900108

Rosette, A. S., Leonardelli, G. J., & Phillips, K. W. (2008). The White Standard: Racial Bias in Leader Categorization. *Journal of Applied Psychology*, *93*(4), 758–777. https://doi.org/10.1037/0021-9010.93.4.758

Rudman, L. A. (1998). Self-promotion as a Risk Factor for Women: The Costs and Benefits of Counterstereotypical Impression Management. *Journal of Personality and Social Psychology*, *74*(3), 629–645. https://doi.org/10.1037/0022-3514.74.3.629

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status Incongruity and Backlash Effects: Defending the Gender Hierarchy Motivates Prejudice Against Female Leaders. *Journal of Experimental Social Psychology*, *48*(1), 165–179. https://doi.org/10.1016/j.jesp.2011.10.008

Rudman, L. A., & Phelan, J. E. (2008). Backlash Effects for Disconfirming Gender Stereotypes in Organizations. *Research in Organizational Behavior*, *28*, 61–79. https://doi.org/10.1016/j.riob.2008.04.003

Rudman, L. A., & Phelan, J. E. (2010). The Effect of Priming Gender Roles on Women's Implicit Gender Beliefs and Career Aspirations. *Social Psychology*, *41*(3), 192–202. https://doi.org/10.1027/1864-9335/a000027

Rust, J., & Golombok, S. (2009). *Modern Psychometrics: The Science of Psychological Assessment* (pp. xiii, 258). Routledge/Taylor & Francis Group.

Salganik, M. J., & Watts, D. J. (2008). Leading the Herd Astray: An Experimental Study of Self-fulfilling Prophecies in an Artificial Cultural Market. *Social Psychology Quarterly*, *71*(4), 338–355. https://doi.org/10.1177/019027250807100404

Shen, X. (Subrina), Kim, H., & Li, J. (2022). Funding Ventures Similar to One of Us: How Status Dynamics within Heterogeneous Groups Affect Venture Evaluation. *Strategic Management Journal*, *43*(10), 2135–2155. https://doi.org/10.1002/smj.3400

Sherf, E. N., Venkataramani, V., & Gajendran, R. S. (2019). Too Busy to Be Fair? The Effect of Workload and Rewards on Managers' Justice Rule Adherence. *Academy of Management Journal*, *62*(2), 469–502. https://doi.org/10.5465/amj.2016.1061

Skrentny, J. D. (2014). *After Civil Rights: Racial Realism in the New American Workplace*. Princeton University Press. https://www.jstor.org/stable/j.ctt4cgb4p

Smith, S. S. (2005). "Don't put my name on it": Social Capital Activation and Job-Finding Assistance among the Black Urban Poor. *American Journal of Sociology*, *111*(1), 1–57. https://doi.org/10.1086/428814

Smith-Doerr, L., Alegria, S., Husbands Fealing, K., Fitzpatrick, D., & Tomaskovic-Devey, D. (2019). Gender Pay Gaps in U.S. Federal Science Agencies: An Organizational Approach. *American Journal of Sociology*, *125*(2), 534–576. https://doi.org/10.1086/705514

Snellman, K., & Solal, I. (2023). Does Investor Gender Matter for the Success of Female Entrepreneurs? Gender Homophily and the Stigma of Incompetence in Entrepreneurial Finance. *Organization Science*, *34*(2), 680–699. https://doi.org/10.1287/orsc.2022.1594

Srivastava, S. B., & Sherman, E. L. (2015). Agents of Change or Cogs in the Machine? Reexamining the Influence of Female Managers on the Gender Wage Gap. *American Journal of Sociology*, *120*(6), 1778–1808. https://doi.org/10.1086/681960

Stainback, K., Kleiner, S., & Skaggs, S. (2016). Women in Power: Undoing or Redoing the Gendered Organization? *Gender & Society*, *30*(1), 109–135. https://doi.org/10.1177/0891243215602906

Stainback, K., & Tomaskovic-Devey, D. (2012). *Documenting Desegregation: Racial and Gender Segregation in Private Sector Employment Since the Civil Rights Act*. Russell Sage Foundation.

Sterling, A. D., & Fernandez, R. M. (2018). Once in the Door: Gender, Tryouts, and the Initial Salaries of Managers. *Management Science*, *64*(11), 5444–5460. https://doi.org/10.1287/mnsc.2017.2880

Stuart, T. E., Hoang, H., & Hybels, R. C. (1999). Interorganizational Endorsements and the Performance of Entrepreneurial Ventures. *Administrative Science Quarterly*, *44*(2), 315–349. https://doi.org/10.2307/2666998

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

Tajfel, H., & Turner, J. (1979). An Integrative Theory of Intergroup Conflict. In *The Social Psychology of Intergroup Relations* (pp. 33–47). Wadsworth.

Tetlock, P. E. (1983). Accountability and Complexity of Thought. *Journal of Personality and Social Psychology*, *45*(1), 74–83. https://doi.org/10.1037/0022-3514.45.1.74

Tilcsik, A. (2021). Statistical Discrimination and the Rationalization of Stereotypes. *American Sociological Review*, *86*(1), 93–122. https://doi.org/10.1177/0003122420969399

Turco, C. J. (2010). Cultural Foundations of Tokenism: Evidence from the Leveraged Buyout Industry. *American Sociological Review*, *75*(6), 894–913. https://doi.org/10.1177/0003122410388491

U.S. Bureau of Labor Statistics. (2004). *Employed persons by detailed occupation, sex, race, and Hispanic or Latino ethnicity*. https://www.bls.gov/cps/cpsaat11.htm

Valentino, L., & Vaisey, S. (2022). Culture and Durable Inequality. *Annual Review of Sociology*, *48*(Volume 48, 2022), 109–129. https://doi.org/10.1146/annurev-soc-030320-102739

Wald, K., Abraham, M., Pike, B., & Galinsky, A. D. (2024). Gender Differences in Climbing up the Ladder: Why Experience Closes the Ambition Gender Gap. *Psychological Science*.

Walker, H. A., Rogers, L., & Zelditch, M. (1988). Legitimacy and Collective Action: A Research Note. *Social Forces*, *67*(1), 216–228. https://doi.org/10.2307/2579108

Wayne, S. J., Sun, J., Kluemper, D. H., Cheung, G. W., & Ubaka, A. (2023). The Cost of Managing Impressions for Black Employees: An Expectancy Violation Theory Perspective. *Journal of Applied Psychology*, *108*(2), 208–224. https://doi.org/10.1037/apl0001030

Williams, C. L. (1992). The Glass Escalator: Hidden Advantages for Men in the "Female" Professions. *Social Problems*, *39*(3), 253–267. https://doi.org/10.2307/3096961

Wingfield, A. H. (2009). Racializing the Glass Escalator: Reconsidering Men's Experiences with Women's Work. *Gender & Society*, *23*(1), 5–26. https://doi.org/10.1177/0891243208323054

Wingfield, A. H., & Chavez, K. (2020). Getting In, Getting Hired, Getting Sideways Looks: Organizational Hierarchy and Perceptions of Racial Discrimination. *American Sociological Review*, *85*(1), 31–57. https://doi.org/10.1177/0003122419894335

Women in Management. (2022). *Catalyst*. https://www.catalyst.org/research/women-in-management/

Women in Management (Quick Take). (2021). *Catalyst*. https://www.catalyst.org/research/women-in-management/

Younkin, P., & Kuppuswamy, V. (2018). The Colorblind Crowd? Founder Race and Performance in Crowdfunding. *Management Science*, *64*(7), 3269–3287. https://doi.org/10.1287/mnsc.2017.2774

Zarate, M. A., & Smith, E. R. (1990). Person Categorization and Stereotyping. *Social Cognition*, *8*(2), 161–185. https://doi.org/10.1521/soco.1990.8.2.161

Zou, L. X., & Cheryan, S. (2017). Two Axes of Subordination: A New Model of Racial Position. *Journal of Personality and Social Psychology*, *112*(5), 696–717. https://doi.org/10.1037/pspa0000080