*Industrial and Organizational Psychology*, **1** (2008), 393–398.


COMMENTARIES

# Gender Stereotypes Are Alive, Well, and Busy Producing Workplace Discrimination

MADELINE E. HEILMAN
*New York University*

ALICE H. EAGLY
*Northwestern University*

A critical examination of research on the relationship between stereotyping and workplace discrimination must meet three requirements. The first requirement is an understanding of the theory that guides this research. The second requirement is an unbiased review of relevant research. The third requirement is comprehension of the ways that different types of research are informative about behavior in organizations. Landy (2008) meets none of these requirements. He misstates the consensual social scientific theory about the relation between stereotyping and discrimination, presents only a selective portion of the relevant research, and misconstrues the basis for generalizing research findings to organizations. As a result, Landy misrepresents the evidence for stereotype-based workplace discrimination.

For brevity, we consider only sex discrimination. Also, consistent with Landy's emphasis, we address the consequences of stereotypes that describe women and men as opposed to stereotypes that prescribe normatively acceptable behavior for them and thus sanction behavior deviating from gender norms (see Eagly & Karau, 2002; Heilman, 2001).

## The Prevailing Theory of Discrimination

Landy asserts that psychological theory views stereotypes' negativity as the basis of workplace discrimination. This assertion is an inaccurate simplification of contemporary theory concerning discrimination in general and sex discrimination in particular (e.g., Glick & Fiske, 2007). Theory has moved beyond the idea of discrimination as simply an outcome of antipathy (see Dovidio, Glick, & Rudman, 2005). In fact, women, who are regarded as the nicer, kinder sex, have a cultural stereotype that is in general more positive than that of men (Eagly & Mladinic, 1994), yet women often are victims of prejudice. Theorists resolved this paradox by recognizing that it is not the negativity of gender stereotypes but their mismatch with desirable work roles that underlies biased workplace evaluations. Heilman (1983, 2001) articulated this "lack of fit" model (see also Dipboye, 1985), and

Correspondence concerning this article should be addressed to Madeline E. Heilman. E-mail: madeline.heilman@nyu.edu

Address: Department of Psychology, New York University, 6 Washington Pl Room 576, New York, NY 10003

Madeline E. Heilman, Department of Psychology, New York University; Alice H. Eagly, Department of Psychology, Northwestern University.

Eagly and Karau (2002) promoted this idea in their role incongruity theory of prejudice toward female leaders.

Consistent with prevailing theory, biased personnel decisions about women can flow from positive or negative attributes ascribed to women. Female stereotypical attributes, including the positive communal qualities of warmth and niceness, tend to be inconsistent with the attributes believed to be required for success in many key organizational positions. People associate such roles with more masculine than feminine attributes (e.g., Schein, 2001). What results from this mismatch between a group stereotype (e.g., women) and a job role (e.g., manager) is negative performance expectations, which in turn produce biased evaluations. And because individuals are commonly assimilated to group stereotypes, this evaluative penalty is exacted even from women who do not possess the qualities that are stereotypical of their group but do possess those regarded as necessary for success in the job role. Discrimination is a behavioral expression of this evaluative penalty.

## Differing Research Paradigms Relevant to Stereotype-Based Sex Discrimination

Landy errs in identifying relevant research. He buttresses his claim that the research has relied on an "artificial work" paradigm by singling out such studies and excluding other types of studies. Among his glaring omissions is a body of research in which persons responsible for personnel selection in organizations receive job applications or resumes (or sometimes phone calls or in-person applicants) through ordinary channels. These research participants do not realize that they are providing data for a research project or that certain applications have been constructed to differ only in sex or other attributes to meet the requirements of an experimental design. This omission of these natural setting experiments, which are known as "audit studies," is puzzling, given their popularity (see Pager, 2007) and their direct implications for understanding labor markets. Although most of these field experiments have addressed racial and ethnic discrimination, some have examined sex discrimination (see review by Riach & Rich, 2002). Their findings conformed to the lack-of-fit formulation: There was a high incidence of sex discrimination against women in the more senior jobs that yield higher status and wages and against both sexes when they applied for jobs dominated by the other sex.

Other researchers have exploited natural variation in conditions that deter discrimination. For example, Goldin and Rouse (2000) conducted a quasi-experimental study of major symphony orchestras' hiring of musicians, positions traditionally held almost exclusively by men. This project compared the success of women and men when the orchestras did (vs. did not) use a screen to hide the identity of the musicians who were auditioning. The screen increased women's success by approximately 50% in the initial round of auditions and made them 1.6 times more likely to win an orchestral position.

Field studies focusing on performance evaluation also support the lack-of-fit principle, including the one such study that Landy cites (Lyness & Heilman, 2006). Using archival organizational data from 489 upper- and senior-level managers from a large financial services organization, this study compared women and men after controlling for age, education, organizational level, and organizational tenure (which Landy incorrectly implied was not controlled). The results showed that women in line jobs (male gender–typed positions) received lower performance ratings than women in staff jobs (female gender–typed positions) or men in either line or staff jobs. Moreover, a meta-analysis of 96 studies of the effectiveness of leaders, as assessed mainly by performance evaluations in organizational field studies, found that men fared better than women in male-dominated leadership roles or culturally masculine settings but that women surpassed men in settings that were less male dominated or less culturally masculine (Eagly, Karau, & Makhijani, 1995). Thus, the lack-of-fit principle prevails in natural settings, whether field experiments, quasi-experiments, or

correlational field studies are the method of investigation and whether selection or performance evaluation is the outcome of interest.

This natural setting research is consistent with laboratory experiments, which also show that evaluative bias follows from lack of fit to job roles. With respect to performance evaluations of leaders, a meta-analysis of experiments that held leaders' behavior constant while varying only their sex showed that men were evaluated more favorably than women in male-dominated leadership roles but equally in other leadership roles (Eagly, Makhijani, & Klonsky, 1992). Also, taking 49 experiments in simulated employment contexts into account, Davison and Burke (2000) classified the research according to stereotype-driven mismatch between job candidates and roles. In male sex–typed jobs (e.g., auto salesperson, life insurance agent), men were preferred over women. However, in female sex–typed jobs (e.g., secretary, director of a day care center) women were preferred over men.

If natural setting research and laboratory experimental research show basically the same pattern of findings, what do these experiments add? Laboratory experiments have the advantage of the more perfect control that follows from randomly assigning participants to conditions that differ only in the sex of the target job applicant or employee. This control reduces the chances that extraneous variables are correlated with sex and thus eliminates alternative interpretations of bias effects. Given the inevitable trade-offs between the control of laboratory experiments and the realism of studies conducted in natural settings, the best support for lack-of-fit principles emerges from consistent findings across these different design types. Also, as we elaborate in the next sections, another advantage of experiments is that they can allow the precise identification of conditions that favor (vs. deter) stereotype-based discrimination.

Both laboratory and field findings are highly supportive of the idea that sex bias in evaluations of women arise not from their sex but from the perceived mismatch between their inferred female attributes and the requirements of male gender–typed jobs. Nonetheless, we agree with Landy's point that a meta-analysis that integrates research from all the relevant research paradigms, treating paradigm as one of many moderator variables, would speak to the issue of the relative magnitude of sex effects in different paradigms. However, such a project is not required to recognize the converging laboratory and field support for the lack-of-fit hypothesis. Also, the lack-of-fit principle and supportive findings provide a plausible explanation for the absence of pro-male bias in studies in which the focal jobs are not male in gender type or in which the gender type of positions is not taken into account. In fact, sex discrimination can penalize men but only when the position is female gender typed —positions men seldom seek.

## Generalization of Evaluative Bias Findings to Organizational Settings

Landy's failure to recognize the findings that have emerged across varied research designs is compounded by his lack of comprehension of how laboratory experiments inform understanding of behavior in work settings. Instead, he offers a one-dimensional critique of laboratory research that rests on its presumed lack of generalizability to the "real world" because of descriptive differences between the laboratory and work settings and the greater complexity of typical work environments.

Landy's discussion of generalizability contains two erroneous assumptions. The first is that research from any one setting can be generalized to another setting only if the two settings "appear" to be similar. His second assumption is that work organizations are similar to one another but different from psychology laboratories. However, there are many differences among work organizations, for example, in the average time employees spend in a given position; human resources practices and procedures; and organizational history, climate, and culture. In fact, many of what Landy assumes

are "common characteristics" of work-related decisions in organizations are actually quite variable and may differ more from one organization to another than from one organization to a laboratory setting. If, for instance, evaluators in one organization feel accountable because they have to publicly justify their evaluations, their decision making is apt to be different from that of evaluators in another organization who do not feel accountable, no matter how similar the organizations appear to be. However, the decision making of accountable organizational evaluators may not differ much from that of laboratory study participants for whom accountability has been created experimentally. In short, the similarity of constraints on evaluative behavior across social settings is not accurately described by Landy's dichotomy between artificial and real-world settings.

The bridge between research and application is theory. Only theoretical principles supported by research reveal the essential similarity or dissimilarity of various types of settings. Theory guides the search for the conditions that promote or hinder discrimination. Once such conditions are identified, they expand the scope of theory. For example, studies have identified features of situations that make female stereotypes more salient (e.g., scarcity of women) and characteristics of women that highlight their femininity (e.g., physical attractiveness or motherhood). Such conditions increase evaluative bias because they exaggerate lack of fit with male sex–typed roles (Heilman & Parks-Stamm, 2007). Also, research has identified situational factors that can affect subjectivity in judgment and reliance on stereotypes. These factors include degree of accountability, anticipated interdependence, and ambiguity in the decision process (e.g., in information, measures, criteria, standards, or source of performance; see Heilman & Haynes, 2008). It is similarity or dissimilarity in such moderating conditions that appropriately guides generalizations from research to organizational settings. Research thereby provides a road map for assessing, predicting, and remedying discrimination.

Related to these issues of generalizability is Landy's assertion that dramatic changes in the nature of work render earlier findings irrelevant to contemporary work settings. Yet, recent findings as well as earlier findings produce the signature lack-of-fit pattern (e.g., Lyness & Heilman, 2006). Nonetheless, workplaces are always changing, and the implications of these changes for discrimination can be predicted from theory that explains their relevance. For example, the effects of virtual environments or work in teams can be anticipated by understanding the moderating effects of important theoretical constructs such as ambiguity, accountability, or the decision makers' stake in their decisions. Therefore, relevant theory and supporting research are essential to explaining the effects of work conditions—even rapidly changing ones—on the relation between gender stereotyping and workplace discrimination.

## The Relevance of Individuating Information to Stereotyping

Another of Landy's major points is that individuating information deters stereotyping in organizational settings. Landy points out that knowing someone in an organizational setting usually entails more than the limited information available to decision makers in laboratory studies and that access to this additional information nullifies the effects of stereotypes. Consistent with robust evidence of stereotype-based discrimination in the field research that we have cited, there are multiple reasons why stereotypes are often not quashed by additional information.

Landy would have us believe that overriding stereotypes is an uncomplicated process and a common event. But neither is the case. Although individuating information can indeed deter stereotyping, there are many circumstances under which it does not have this effect. As well established in research on social cognition, whether stereotypes are activated and applied depends on a complex of cognitive and motivational variables (see Kunda & Spencer, 2003), including whether

individuals have the cognitive resources to inhibit stereotyping and whether they are motivated to be accurate. The conditions that deter stereotyping are often absent in work settings because of decision makers' cognitive overload, motivation to maintain the status quo, and limited dependence on subordinates for valued outcomes. Thus, stereotype-based perceptions of lack of fit are likely to take hold despite the availability of additional information. Moreover, the expectations that these perceptions produce can be tenacious and self-sustaining, even in the face of disconfirming information. They can affect what information social perceivers attend to, how they interpret the information, what they remember, and how they engage in social interaction, often in a manner that reinforces their prior expectations. Therefore, the individuating information available to decision makers in organizational settings does not necessarily prevent discriminatory decision making.

Landy's misunderstanding about the power of individuating information leads him to argue that the simplicity and control of laboratory research make it easier to find bias in the laboratory than in the field. However, there are some features of experimental laboratory research that make bias less rather than more likely. Specifically, the participants tend to be younger and therefore more accepting of contemporary cultural norms that espouse gender equity than many people who populate organizations. Also, laboratory participants are not likely to be as pressed for time or cognitively busy or distracted as organizational decision makers. Consequently, laboratory participants ordinarily have the cognitive resources to be more deliberative and use individuating information in making inferences. Therefore, it is not at all clear that the deck is stacked in favor of discrimination in laboratory experiments.

## Bottom Line

In summary, Frank Landy falls short of making his case. He fails to accurately present the prevailing theory of how gender stereotypes produce sex discrimination. Moreover, by providing a highly selective review of the relevant research, he wrongly suggests that evidence for workplace sex discrimination derives almost exclusively from experiments using artificial work paradigms when in fact there is converging evidence from laboratory and field studies using a broad range of research paradigms. He also misinterprets research on the mitigating effects of individuating information on stereotyping and its consequences. And he shows no appreciation of the widely accepted methodological principle that the best bridge to applications derives from deploying theory. If Landy had thoroughly reviewed the sex discrimination literature and paid heed to the scientific issues involved in his claims, the result would no doubt have been a more productive exchange of views.

## References

Davison, H. K., & Burke, M. J. (2000). Sex discrimination in simulated employment contexts: A meta-analytic investigation. *Journal of Vocational Behavior, 56*, 225–248.

Dipboye, R. L. (1985). Some neglected variables in research on discrimination in appraisals. *Academy of Management Review, 10*, 116–127.

Dovidio, J. F., Glick, P., & Rudman, L. (Eds.). (2005). *On the nature of prejudice: Fifty years after Allport*. Malden, MA: Blackwell.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*, 573–598.

Eagly, A. H., Karau, S. J., & Makhijani, M. G. (1995). Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin, 117*, 125–145.

Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin, 111*, 3–22.

Eagly, A. H., & Mladinic, A. (1994). Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology* (Vol. 5, pp. 1–35). New York: Wiley.

Glick, P., & Fiske, S. T. (2007). Sex discrimination: The psychological approach. In F. J. Crosby, M. S. Stockdale, & S. A. Ropp (Eds.), *Sex discrimination in the workplace: Multidisciplinary perspectives* (pp. 155–187). Malden, MA: Blackwell.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review, 90*, 715–741.

Heilman, M. E. (1983). Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior, 5*, 269–298.

Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, *57*, 657–674.

Heilman, M. E., & Haynes, M. C. (2008). Subjectivity in the appraisal process: A facilitator of gender bias in work settings. In E. Borgida & S. T. Fiske (Eds.), *Beyond common sense: Psychological science in the courtroom* (pp. 127–155). Malden, MA: Blackwell.

Heilman, M. E., & Parks-Stamm, E. J. (2007). Gender stereotypes in the workplace: Obstacles to women's career progress. In S. J. Correll (Ed.), *Social psychology of gender: Advances in group processes* (Vol. 24, pp. 47–77). Oxford, UK: Elsevier.

Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, *129*, 522–544.

Landy, F. J. (2008). Stereotypes, bias, and personnel decisions: Strange and stranger. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *1*, 379–392.

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

Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *Annals of the American Academy of Political and Social Science*, *609*, 104–133.

Riach, P. A., & Rich, J. (2002). Field experiments of discrimination in the market place. *Economic Journal*, *112*, F480–F518.

Schein, V. E. (2001). A global look at psychological barriers to women's progress in management. *Journal of Social Issues*, *57*, 675–688.