Triangulation Across Methodologies: All Signs Point to Persistent Stereotyping and Discrimination in Organizations

LISA M. LESLIE
University of Minnesota

EDEN B. KING
George Mason University

JILL C. BRADLEY
California State University

MICHELLE R. HEBL
Rice University

Landy (2008) claims that (a) field-based evidence of discrimination is scarce and (b) laboratory-based research is "largely useless" (p. 379) for understanding real-world organizational contexts. He therefore concludes that we lack sufficient evidence that stereotyping and discrimination influence organizational decision making.

We strongly disagree with Landy's conclusion and show that a large, multimethod, multidisciplinary body of research documents persistent discrimination in organizations. We counter Landy's first claim by providing several examples of field-based research that document discrimination in real-world organizational contexts but were overlooked entirely by Landy. We then counter Landy's assertion that laboratory

findings do not generalize by highlighting the imperfect assumption upon which Landy bases this claim. We conclude that triangulated evidence—spanning both laboratory and field settings—documents the unfortunate reality that discrimination is alive and well in organizations.

Evidence of Discrimination in Real-World Organizational Contexts

Audit studies and work outcome studies, both of which are field-based research, provide evidence that discrimination influences personnel decision making. To be consistent with Landy, we interpret evidence of discrimination as a manifestation of stereotyping when reviewing these studies. At the same time, we acknowledge that stereotypes (i.e., beliefs about a group) are only one potential explanation for discrimination (i.e., behaviors toward that group). Although laboratory research shows that stereotypes drive discrimination (e.g., Correll, Benard, & Paik, 2007), field-based evidence of the stereotype–discrimination linkage is rare, given that stereotype

Correspondence concerning this article should be addressed to Lisa M. Leslie. E-mail: lmleslie@umn.edu

Address: Carlson School of Management, Department of Psychology, University of Minnesota, 321, 19th Avenue South, Suite 3-300, Minneapolis, MN 55455

Lisa M. Leslie, Carlson School of Management, Department of Psychology, University of Minnesota; Eden B. King, Department of Psychology, George Mason University; Jill C. Bradley, Craig School of Business, California State University; Michelle R. Hebl, Rice University.

assessment is not a naturally occurring organizational phenomenon.

Audit studies: real situations and evaluators, fictitious targets. Compelling field-based evidence of discrimination comes from audit studies in which researchers apply to real job postings by sending fictitious resumes or having confederates pose as job applicants. Researchers vary applicant demographics experimentally and hold all other information (e.g., qualifications, ability) constant, allowing them to draw inferences regarding whether demographic characteristics affect personnel decisions. Unlike the laboratory studies that Landy criticizes, audit studies are conducted in the context of actual personnel decisions (i.e., real situations) made by job incumbents (i.e., real evaluators). The targets of discrimination are fictitious in that they are not actual job applicants, yet they are real in the eyes of the evaluators.

Two exemplary audit studies provide evidence that applicants who are either Black or mothers are discriminated against in resume-screening decisions. Bertrand and Mullainathan (2004) manipulated applicant race by varying the names that appeared on resumes sent to job postings (e.g., White names: Emily, Greg; Black names: Lakisha, Jamal). A 50% Black–White gap emerged, such that Black applicants received one callback for every 15 resumes submitted, whereas White applicants received one callback for every 10 resumes submitted. Correll et al. (2007) conducted a similar audit study in which they manipulated gender by varying applicants' names (e.g., Allison or Brad) and parental status by varying applicants' volunteer activities (i.e., parent–teacher or alumni association). Childless women received twice as many callbacks as mothers, but there was no difference in callbacks between childless men and fathers. Taken together, these studies provide evidence that mothers and Black applicants are discriminated against in real-world hiring decisions.

Audit studies are not only limited to resume screening but also have been conducted in face-to-face situations. A set of studies in which confederates interviewed for various positions demonstrate that Black applicants are less likely than White applicants to be called back and offered a job (Fix & Turner, 1998). Similarly, a study in which male and female confederates were sent to restaurants to apply for waiter positions revealed hiring discrimination against women (Neumark, 1996). One limitation of these studies is difficulty matching all characteristics across confederates. Hebl, King, Glick, Singletary, and Kazama (2007) circumvented this issue by examining the experiences of job applicant confederates who did or did not wear a pregnancy prosthesis. Reports by the confederates, observers, and naïve coders of the audiotaped interactions indicated that the same applicant was treated with greater hostility when she appeared to be pregnant than when she did not. Similarly, when confederates applied for retail jobs in Texas wearing hats reading "gay and proud" or "Texan and proud," the same individuals experienced more hostility when wearing gay and proud hats (Hebl, Foster, Mannix, & Dovidio, 2002). Although pregnant and nonpregnant women and gay and nongay individuals did not differ in callbacks, hostile treatment provides evidence of stereotyping and may negatively affect personnel decisions (Shapiro, King, & Quiñones, 2007).

In all, audit studies conducted in economics (e.g., Bertrand & Mullainathan, 2004), sociology (e.g., Correll et al., 2007), and psychology (e.g., Hebl et al., 2002) provide causal evidence that members of low-status groups experience discrimination. They document stereotyping in *real situations*, implying that evaluators make biased decisions even when they are accountable for decision outcomes and use instruments designed by the organization (see Landy, p. 383, point 5 and p. 383, point 8). Similarly, these studies document stereotyping using *real evaluators*, which implies that rater training and experience making personnel decisions do not prevent bias (see Landy, p. 383, point 2 and p. 383, point 9).

Although compelling, audit studies are largely constrained to the context of hiring

decisions that involve fictitious applicants because monitoring the performance evaluations or pay of confederates posing as real employees is not feasible, practically or ethically. As a result, evaluators in these studies have limited information about targets. Landy argues that the absence of such personal knowledge (i.e., individuating information) limits the generalizability of discrimination research (pp. 389–390 points 1, 3, 4, 5, and 7). The next group of studies we discuss, however, uses real targets, as well as real situations and evaluators, and therefore suggests that discrimination persists even when individuating information is available.

Work outcome studies: real situations, evaluators, and targets. Databases containing information on the work outcomes (e.g., pay, performance) of job incumbents provide another source of evidence regarding discrimination in organizations. This research examines differences in work outcomes based on group membership, controlling for human capital (e.g., education, experience) and other relevant factors (e.g., occupation, industry, job level). Differences not accounted for by these controls lend support to the conclusion that stereotypes and discrimination continue to affect work outcomes under the presumption that all relevant predictors are included and reliably measured.

Research on group-based differences in pay, which is common in economics, industrial relations, and sociology, documents persistent pay gaps based on gender, motherhood, and ethnicity. In the case of gender, women's salaries in the United States are 80% of men's. Education, experience, occupation, industry, and other qualifications explain 40% of this gap, yet after controlling for these factors, women still earn only 91% of what men earn (Blau & Kahn, 2007). Moreover, the size of the gender wage gap is larger for mothers than for single women. On average, mothers earn approximately 5% less per child than single women do, after accounting for the same human capital and occupational controls (Correll et al.,

2007). An ethnic pay gap also exists in the United States such that Black men earn 69–76% of what White men earn, and 12–15% of the gap remains after controlling for human capital (Fugazza, 2003). Moreover, studies that control for job performance find that performance differences do not explain the Black–White wage gap (Kahn & Sherer, 1988) and in some cases even increase it (Coleman, 2003).

Wage gap research suggests that gender, motherhood, and ethnicity correspond to lower pay after controlling for numerous alternatives to the explanation of discrimination. These findings, however, do not provide unequivocal evidence of stereotyping and discrimination, as uncontrolled factors and measurement unreliability may explain part of the pay gap. Yet, a number of studies that have used varied methodologies to circumvent these limitations similarly conclude that discrimination is the likely explanation (see Blau & Kahn, 2007, for a review).

In addition to pay, work outcome studies also find gender and ethnic differences in performance ratings, although as Landy noted, the differences are small. Smaller effect sizes for performance appraisals than for pay, however, may not be surprising given that performance appraisals reflect a single evaluation at a single point in time. When performance appraisal differences are considered in a summative manner, as in the accrual of pay through raises, small differences can accumulate and create meaningful disparities between advantaged and disadvantaged groups (Martell, Lane, & Emrich, 1996).

In all, work outcome studies provide evidence consistent with continued stereotyping and discrimination in organizations. Like audit studies, this research suggests that managers (i.e., real evaluators) discriminate when making actual personnel decisions (i.e., real situations). In addition, work outcomes studies use *real targets* (i.e., job incumbents) and therefore find evidence of discrimination when evaluators have individuating information about targets.

In summary, field-based discrimination research converges across a range of

methodologies (audit studies and work outcomes studies) and documents persistent discrimination in real-world organizational contexts. Landy claims that field evidence of discrimination is limited, yet bases his conclusion primarily on psychological research on performance ratings. Bias in performance appraisals is only one of the many ways in which stereotypes may affect personnel decisions. Taking a broader perspective, one that spans a variety of employment outcomes (e.g., hiring, pay) and disciplines (e.g., psychology, economics, sociology), field-based research paints a consistent picture regarding continued discrimination in organizations.

Generalizing Laboratory-Based Discrimination Research

In addition to claiming that field-based evidence of discrimination is scarce, Landy also asserts that laboratory studies do not generalize to real-world organizational contexts. Reminiscent of a long-standing debate (e.g., Black, 1955), Landy argues that because laboratory studies use fictitious evaluators, situations, and targets (pp. 382), they are too far removed from the real world to provide meaningful evidence of discrimination in organizations. Laboratory findings, however, largely converge with the field studies reviewed above. As Landy himself states, "it is common to find some evidence … of stereotypes in laboratory studies" (p. 381). Similar findings across laboratory and field settings challenge Landy's criticisms of laboratory research. As we will describe in more detail, the likely explanation for the discrepancy between Landy's claim that laboratory findings do not generalize and the reality of laboratory-field convergence is that Landy's criticism of laboratory research is based on a faulty assumption.

Landy assumes individuating information uniformly reduces stereotyping and discrimination. This assumption forms the basis for his most emphatic criticism of laboratory studies—that they use fictitious targets, about whom evaluators have little individuating information (pp. 379–392). Landy,

however, discusses stereotypes as a purely *descriptive* heuristic of how a group "is," which individuals use to deal with the cognitive overload they encounter in their daily lives. In fact, stereotypes are often *prescriptive* and therefore describe how a group "ought to be" (Heilman, 2001). Unlike descriptive stereotypes, prescriptive stereotypes often serve a motivational purpose, such as enhancing self-esteem, granting access to limited resources, or maintaining the status quo (e.g., Allport, 1954). Because prescriptive stereotypes are motivational in nature, they are more immune to the effects of individuating information than descriptive stereotypes.

For example, someone might have individuating information about a particular woman (e.g., Hillary Clinton is a strong leader, assertive, and effective) and realize that this information is inconsistent with stereotypes about women; however, this does not mean that he or she will no longer stereotype Hillary Clinton or women more generally. Individuating information simply does not address the prescriptive aspect of stereotyping, which suggests that Hillary Clinton, as a woman, *should* act in a certain way (e.g., nurturing, emotional). In fact, women who violate stereotypes may threaten the status quo and traditional hierarchy; hence, stereotype-inconsistent individuating information may reinforce prescriptive stereotypes to prevent other women from similarly violating traditional gender roles.

Although individuating information surely reduces stereotypes and discrimination in some cases, many empirical findings highlight the flaw in Landy's logic. First, people are not always motivated to seek out or use individuating information (Cameron & Trope, 2004). Indeed, Sherman, Stroessner, Conrey, and Azam (2005) found that prejudiced individuals are less likely than unprejudiced individuals to form individuated impressions in the first place. Second, individuating information does not necessarily eliminate stereotyping and discrimination in organizational decision making. For example, Glick, Zion, and Nelson (1988) examined the role of individuating

information in business professionals' evaluations of experimentally manipulated resumes. Individuating information reduced gender-typed personality inferences about male and female applicants but did not eliminate gender discrimination. More recently, Singletary and Hebl (in press) investigated the impact of individuating information on sexuality-based discrimination. They similarly found that having gay applicants disclose individuating information in an interview neither reduced interpersonal discrimination (measured from the applicant's as well as an independent observer's perspective) nor increased the likelihood of gay applicants receiving an interview callback.

Individuating information is not only often ineffective at reducing discrimination but may even *increase* discrimination. For example, King, Shapiro, Hebl, Singletary, and Turner (2006) conducted a field study on customer service interactions and examined the role of individuating information in weight-based discrimination. Confederate customers who appeared to be obese disclosed either stereotype-consistent (i.e., were not dieting, had not competed in an athletic event) or stereotype-inconsistent (i.e., were dieting, had competed in an athletic event) individuating information. Stereotype-inconsistent individuating information had no effect on interpersonal treatment. Stereotype-consistent individuating information, however, increased the degree of interpersonal discrimination experienced by overweight targets. Thus, in contrast to Landy's claims, individuating information is not a panacea for preventing prescriptive stereotypes and, in fact, enhances discrimination under certain conditions.

In all, Landy's assumption that individuating information always reduces stereotyping and discrimination is flawed, making his disregard for laboratory studies on the grounds that evaluators have little individuating information about fictitious targets difficult to justify. Rather, the convergence of laboratory findings with a variety of field methodologies, including those in which individuating information is present, suggests that labora-tory research provides meaningful evidence regarding discrimination in organizational contexts.

Conclusions

A large body of evidence—one that spans multiple disciplines and includes both laboratory and field methodologies—documents the unfortunate conclusion that stereotyping and discrimination persist in organizations. This conclusion directly counters Landy's assertions that field-based evidence of discrimination is scarce and that laboratory findings provide evidence of discrimination but do not generalize to real-world organizational contexts. Of course, no methodology is without limitations and, as Landy has done with laboratory research, each methodology discussed here could be critiqued. Nevertheless, triangulated findings across a wide range of methodologies provide strong evidence that continued discrimination is a robust and generalizable phenomenon (cf. Lykken, 1968).

References

Allport, G. W. (1954). *The nature of prejudice*. Oxford, UK: Addison Wesley.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review, 94*, 991–1013.

Black, V. (1955). Laboratory versus field research in psychology and the social sciences. *British Journal for the Philosophy of Science, 5*, 319–330.

Blau, F. D., & Kahn, L. M. (2007). The gender pay gap: Have women gone as far as they can? *Academy of Management Perspectives, 21*, 7–23.

Cameron, J. A., & Trope, Y. (2004). Stereotype-biased search and processing of information about group members. *Social Cognition, 22*, 650–672.

Coleman, M. G. (2003). Job skill and Black male wage discrimination. *Social Science Quarterly, 84*, 892–906.

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology, 112*, 1297–1338.

Fix, M., & Turner, M. A. (1998). *A national report card on discrimination in America: The role of testing*. Washington, DC: Urban Institute Press.

Fugazza, M. (2003). Racial discrimination: Theories, facts and policy. *International Labour Review, 142*, 507–541.

Glick, P., Zion, C., & Nelson, C. (1988). What mediates sex discrimination in hiring decisions? *Journal of Personality and Social Psychology, 55*, 178–186.

Hebl, M., Foster, J. M., Mannix, L. M., & Dovidio, J. F. (2002). Formal and interpersonal discrimination: A field study examination of applicant bias. *Personality and Social Psychological Bulletin, 28*, 815–825.

Hebl, M. R., King, E. B., Glick, P., Singletary, S., & Kazama, S. (2007). Hostile and benevolent reactions to pregnant women: Complementary interpersonal punishments and rewards that maintain traditional gender roles. *Journal of Applied Psychology, 92*, 1499–1511.

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.

Kahn, L. M., & Sherer, P. D. (1988). Racial differences in professional basketball players' compensation. *Journal of Labor Economics, 6*, 40–61.

King, E., Shapiro, J. L., Hebl, M., Singletary, S., & Turner, S. (2006). The stigma of obesity in customer service: A mechanism for remediation and bottom-line consequences of interpersonal discrimination. *Journal of Applied Psychology, 91*, 579–593.

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

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin, 70*, 151–159.

Martell, R. F., Lane, D. M., & Emrich, C. (1996). Male-female differences: A computer simulation. *American Psychologist, 56*, 157–158.

Neumark, D. M. (1996). Sex discrimination in restaurant hiring: An audit study. *Quarterly Journal of Economics, 111*, 915–941.

Shapiro, J., King, E. B., & Quiñones, M. (2007). Expectations of obese trainees: How stigmatized trainee characteristics influence training effectiveness. *Journal of Applied Psychology, 92*, 239–249.

Sherman, J. W., Stroessner, S. J., Conrey, F. R., & Azam, O. A. (2005). Prejudice and stereotype maintenance processes: Attention, attribution, and individuation. *Journal of Personality and Social Psychology, 89*, 607–622.

Singletary, S. L. B., & Hebl, M. (in press). Compensatory strategies for reducing interpersonal discrimination: The effectiveness of acknowledgment, increased positivity, and individuating information. *Journal of Applied Psychology.*