University of Pennsylvania
ScholarlyCommons

Management Papers

Wharton Faculty Research

12-2008

The Challenge of Debiasing Personnel Decisions: Avoiding Both Under- and Overcorrection

Philip E. Tetlock
University of Pennsylvania

Gregory Mitchell

Terry L. Murray

Follow this and additional works at: http://repository.upenn.edu/mgmt_papers

Part of the Business Administration, Management, and Operations Commons

Recommended Citation

The Challenge of Debiasing Personnel Decisions: Avoiding Both Under-
and Overcorrection

Disciplines
Business Administration, Management, and Operations

The Challenge of Debiasing Personnel Decisions:
Avoiding Both Under- and Over-Correction

Philip E. Tetlock,[*] **Gregory Mitchell,**[**] **and Terry L. Murray**[***]

Introduction

This commentary advances two interrelated scientific arguments. First, we endorse Landy's (2008) concerns about the insufficient emphasis placed on individuating information by scholars eager to import social-cognition work on stereotyping into employment law. Building on Landy's analysis, we emphasize that greater attention needs to be given to the power of accountability and teamwork incentives to motivate personnel decision-makers to seek and utilize individuating information that is predictive of job-relevant behavior. Second, we note how easy it is for exchanges between proponents and skeptics of unconscious stereotyping to lead to ideological stalemates. We therefore stress the need for proponents of different viewpoints to engage in adversarial collaborations in which the goal is to design studies that each camp agrees have the potential to change minds on the conditions under which, and the extent to which, unconscious stereotyping is likely to bias personnel decision-making. Without aggressive testing of organizational boundary conditions on stereotyping hypotheses, judges, regulators and managers will be compelled to rely on their best guesses about what exactly organizations need to do–by way of structuring personnel decision processes–to guarantee equality of opportunity.

Access-to-Individuating-Information *X* Motivation-to-Use-It Interactions

Landy rightly notes that personnel decision-makers almost always have more "individuating" information about the specific employees they are judging than lab subjects possess about the hypothetical paper people they are judging. Indeed, in the most extreme case involving the Implicit Association Test, subjects are provided with only brief glimpses of faces of different racial, ethnic, or sexual make-up and the subjects must pair these faces with positive or negative stereotype trait terms as quickly as possible; millisecond differences in reaction time determine whether a subject supposedly holds an unconscious positive or negative stereotype of the groups represented by the faces. We challenge social psychologists to come up with a test setting in which subjects are provided with less information about the target of a judgment than this. These are not "stranger-to-stranger" interactions, to use Landy's term borrowed from Copus (2005); these are not interactions at all. This information-free zone may be appropriate for social-cognition research into basic-level psychological processes, but it does not provide a sound foundation for making inferences about stereotyping in the workplace.

---

[*] Lorraine Tyson Mitchell Chair of Organizational Behavior, Haas School of Business, University of California, Berkeley, 545 Student Services Building #1900, Berkeley, CA 94720-1900 510-642-2571, Tetlock@haas.berkeley.edu.
[**] Professor of Law & E. James Kelly, Jr.-Class of 1965 Research Professor, University of Virginia School of Law, 580 Massie Road, Charlottesville, VA 22903-1738, 434-243-4088, greg_mitchell@virginia.edu.
[***] LASSC, LLC, 2954 Beaumont Farm Road, Charlottesville, VA 22901, 434-964-1914, lassconsulting@lassconsulting.com. All three authors are affiliated with LASSC, LLC, which provides expert consulting services on social science issues in corporate and litigation settings, including issues surrounding the application of stereotyping research to personnel decisions.

Here we would add that personnel decision-makers are often much more motivated to seek out and make the best possible use of the available information than are the sophomore conscripts in social-cognition experiments. Specifically, we see strong empirical and theoretical grounds for the following propositions:

(a) Relative to laboratory subjects who rarely expect to interact with the mostly "paper people" they are judging, personnel decision-makers have strong outcome-interdependence incentives to assemble the most productive possible teams and therefore to be on sharp lookout for individuating information that confers predictive advantage in spotting those who possess key job-relevant skills and who have demonstrated good work ethic. Indeed, Neuberg and Fiske (1987; see also Fiske & Neuberg, 1990) show that even relatively weak outcome interdependence incentives are sufficient to shift lab subjects to rely less on stereotypic category labels and more on individuating information.

(b) Relative to laboratory subjects who usually expect their judgments of paper people to be anonymous and devoid of consequences, personnel decision makers often feel accountable to high-status others in the organization who expect them to make decisions that simultaneously advance the economic efficiency and profitability of the firm and demonstrate the organization's commitment to EEO norms. There are good theoretical and empirical reasons for believing that these types of accountability pressures will motivate many decision makers to ground their judgments in diagnostic job-relevant individuating information--and to monitor their thought processes to ensure that inappropriate category-based stereotypes do not influence their judgments (see, e.g., Ford et al., 2004; Ruscher & Duval, 1998). Indeed, Tetlock (1992) and Lerner and Tetlock (1999) have reviewed large bodies of research literature that demonstrate that even relatively weak lab forms of accountability are sufficient—under the right conditions—to check a host of judgmental biases often linked to stereotyping, including insensitivity to dissonant evidence (Tetlock, 1983a), over-confidence (Tetlock and Kim, 1987), belief perseverance (Tetlock, 1983b), and the fundamental attribution error (Tetlock, 1985).

Navigating Between Under- and Over-Correction

We therefore share Landy's view that findings from social-psychological studies on stereotypes hold little direct relevance for work settings, primarily because these studies typically ignore the potentially large interaction effects of individuating information and motivation to use such information on the likelihood of reaching accurate and fair conclusions about employees (for the much longer version of our argument, see Mitchell & Tetlock, 2006). Not everyone shares our skepticism, however. Some are confident that laboratory findings on stereotyping will hold up in real-world settings (e.g., Bielby, 2003; Kang & Banaji, 2006) and have used this research to criticize the personnel processes of organizations that have been sued for employment discrimination (e.g., Bielby, 2005; Borgida & Kim, 2007).

If we skeptics are correct, then embracing the prescriptions of the generalizers is likely to lead to undue interference with personnel processes and over-correction of bias with real consequences (most notably backlash effects from more aggressive diversity efforts or unnecessarily straining interactions between unprejudiced Whites and minorities as Whites engage in formal behaviors

that minorities interpret as hostility rather than as efforts at bias-controlling impartiality; see Goff et al., 2008; Kidder eta al., 2004; Norton et al., 2006; Vorauer & Turpie, 2004). If the generalizers are correct, then considerable work needs to be done to determine how to prevent the possibly pernicious effects of unconscious biases in the workplace (see Kalev et al., 2006). Accordingly, we believe that high priority should be placed on determining who is correct (or rather, on determining the conditions under which the findings from laboratory studies do and do not generalize to real world work settings).

Unfortunately, it is all too easy for the two camps to speak past each other. Lab researchers tend to see themselves as exploring fundamental perceptual-cognitive processes and see little theoretical value in exploring the possible interactive effects of social-context manipulations, such as accountability, with their hypothesized processes. If anything, their goals are explicitly reductionist and their attention is directed in the opposite level-of-analysis direction, toward ever more fine-grained and ultimately neurological characterizations of process mechanisms (e.g., Eberhardt, 2005). The reductionists' attitude toward finding that accountability checks bias might well be: "that is just demand characteristics at work." The social-contextualists might then reply: "organizations are loaded with normative rules that are the functional equivalent of demand characteristics." And the reductionists might then counter: "you can temporarily suppress unconscious stereotyping with heavy handed accountability manipulations—but the biasing effects will simply resurface in other aspects of organizational life." To which the social-contextualists might rejoin: "your framework becomes non-falsifiable if you don't specify the precise boundary conditions for the effects."

Concerns over avoiding such protracted level-of-analysis stalemates led Tetlock and Mitchell (in press) to urge both generalizers and skeptics to engage in a process of scientific dispute resolution known as adversarial collaboration (Mellers, Hertwig & Kahneman, 2000), in which each side works together to jointly design empirical studies that each agrees, ex ante, have the potential to yield data that could induce them to modify their views.

Imagine the following stylized debate. When pressed to be precise about effect sizes in a low-accountability/high-anonymity control condition, proponents of unconscious stereotyping effects estimate that personnel decision-makers see equivalently qualified black promotion candidates for job x as 1.2 standard deviations less capable than white candidate. By contrast, skeptics estimate the effect to be much smaller on average (say, 0.2 of a standard deviation) and, moreover, insist that substantial fractions of the population. Those who score on scales of motivation to control prejudice (e.g., Glaser & Knowles, 2008) will either show no bias or even a slight-to-moderate pro-black bias.

Let's also posit that, when pressed to be precise about effect sizes in a setting with accountability measures in place designed to check prejudice (measures that mandate procedures to check cognitive biases but fall short of mandating equality of result), proponents of the unconscious-stereotyping position see such measures as likely to reduce the anti-black bias from 1.2 standard deviations to 0.8 standard deviations whereas skeptics see such measures as likely to transform a slight anti-black bias into a slight pro-black bias.

Speaking for ourselves, if unconscious stereotyping effects held up under reasonably realistic accountability-process protections against discrimination, we would indeed feel an obligation to scale down our earlier expressed skepticism. And we would hope that if the opposite results materialized, generalizers might be more inclined to change their minds. The great advantage is that, when each camp signs off on the methodological details before the data are collected, it becomes more difficult to criticize the study for methodological defects which can always be found ex post.

Closing comments

The policy stakes in debates over unconscious stereotyping are unusually high. Stereotyping research, including the new research into implicit stereotypes, has become a centerpiece of plaintiffs' cases in employment discrimination class actions (see Monahan, Walker & Mitchell, in press). Prominent plaintiffs' experts have argued that unconscious stereotyping effects are so pervasive and potent that it is reasonable to assume that virtually all managers in a company are predisposed to discriminate against particular protected categories of employees whenever there is "excessive" subjectivity or discretion in personal decision-making and "inadequate" accountability for how personnel decisions are made (e.g., Bielby, 2005). For these experts, the solution is to move to greater objectivity in decision-making (despite evidence that this may not serve minorities well; see Roth et al., 2003) and strict monitoring of employment outcomes for possible disparities among groups (an approach that is likely to lead to implicit, if not explicit, quotas; see Fryer, in press).

In our view, these applications of stereotype research go far beyond what the existing science can support and risk causing perverse effects within organizations. For all of the reasons that Landy sets out, we simply have no empirical basis for believing that personnel decisions studied in artificial lab settings approximate personnel decisions in real world settings. We believe that the fastest way to bridge the gap between the lab and the real world, and to do so in a way that convinces both skeptics and generlizers, is for researchers from each camp to engage in adversarial collaborations that require making reasonably precise, ex ante, predictions in agreed-upon research designs. Researchers who are sincerely interested in promoting effective debiasing in organizations, and in avoiding the negative effects of under- and over-correction of bias, should be willing to engage in this adversarial-collaboration process.

References

Bielby, W.T. (2003). Can I get a witness? Challenges of using expert testimony on cognitive bias in employment discrimination litigation, *Employee Rights & Employment Policy Journal*, 7, 377.

Bielby, W.T. (2005). Applying social research on stereotyping and cognitive bias to employment discrimination litigation: The case of allegations of systematic gender bias at Wal-Mart stores. In R. L. Nelson & L. B. Nielsen (eds.), *Handbook of employment discrimination research: Rights and realities* (pp. 395-407). New York: Springer.

Borgida, E., & Kim, A. (2007). Reflections on being an expert witness in class action sex discrimination litigation. In F.J. Crosby, M.S. Stockdale, & S.A. Ropp (eds.), *Sex discrimination in the workplace* (pp. 117-130). Oxford: Blackwell Publishing.

Copus, D. (2005). A lawyer's view: Avoiding junk science. In F.J. Landy (Ed.), *Employment discrimination litigation: Behavioral, quantitative, and legal perspectives* (pp. 450-502). San Francisco: Jossey-Bass.

Eberhardt, J. L. (2005). Imaging race. *American Psychologist*, *60*, 181-190.

Fiske, S.T., & Neuberg, S.L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M.P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1-74). New York: Academic Press.

Ford, T.E., Gambino, F., Lee, H., Mayo, E., & Ferguson, M.A. (2004). The role of accountability in suppressing managers' preinterview bias against African-American sales job applicants. *Journal of Personal Selling & Sales Management*, *24*, 113-124.

Fryer, R.G. (in press). Implicit quotas. *Journal of Legal Studies*.

Glaser, J., & Knowles, E.D. (2008). Implicit motivation to control prejudice. *Journal of Experimental Social Psychology*, *44*, 164–172.

Goff, P. A., Steele, C. M., & Davies, P. G. (2008). The space between us: Stereotype threat and distance in interracial contexts. *Journal of Personality and Social Psychology*, *94*, 91-107.

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*, 589-617.

Kang, J., & Banaji, M.R. (2006). Fair measures: A behavioral realist revision of "affirmative action." *California Law Review*, *94*, 1063-1118.

Kidder, D.L., Lankau, M.J., & Chrobot-Mason, D. (2004). Backlash toward diversity initiatives: Examining the impact of diversity program justification, personal and group outcomes. *International Journal of Conflict Management*, *15*, 77-102.

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

Lerner, J., & Tetlock, P.E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*, 255-275.

Mellers, B., Hertwig, R., and Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, *12*, 269-275.

Mitchell, G., & Tetlock, P.E. (2006). Antidiscrimination law and the perils of mindreading. *Ohio State Law Journal*, *67*(5), 1023-1121.

Monahan, J., Walker, L., & Mitchell, G. (In press). Contextual evidence of gender discrimination: The ascendance of "social frameworks." *Virginia Law Review*, *94*.

Neuberg, S.L., & Fiske, S.T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology, 53*, 431-444.

Norton, M. I., Sommers, S. R., Apfelbaum, E. P., Pura, N., & Ariely, D. (2006). Colorblindness and interracial interaction: Playing the "Political Correctness Game." *Psychological Science*, *17*, 949-953.

Roth, P.L., Huffcutt, A.I., & Bobko, P. (2003). Ethnic group differences in measures of job performance: A new meta-analysis. *Journal of Applied Psychology, 88*, 694-706.

Ruscher, J. B., & Duval, L. L. (1998). Multiple communicators with unique target information transmit less stereotypical impressions. *Journal of Personality and Social Psychology*, *74*, 329-344.

Tetlock, P.E. (1983a). Accountability and complexity of thought. *Journal of Personality and Social Psychology*, *45*, 74-83.

Tetlock, P.E. (1983b). Accountability and perseverance of first impressions. *Social Psychology Quarterly*, *46*, 285-292.

Tetlock, P.E. (1985). Accountability: A social check on the fundamental attribution error. *Social Psychology Quarterly*, *48*, 227-236.

Tetlock, P.E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 331-376). New York: Academic Press.

Tetlock, P.E., & Kim, J.L. (1987). Accountability and judgment processes in a personality prediction task. *Journal of Personality and Social Psychology*, *52*, 700-709.

Tetlock, P.E., & Mitchell, G. (in press). Unconscious prejudice and accountability systems: What must organizations do to prevent discrimination? *Research in Organizational Behavior*, *30*.

Vorauer, J. D., & Turpie, C. A. (2004). Disruptive effects of vigilance on dominant group

members' treatment of outgroup members: Choking versus shining under pressure. *Journal of Personality and Social Psychology*, *87*, 384-399.