VIRGINIA LAW REVIEW
IN BRIEF

VOLUME 95 FEBRUARY 28, 2010 PAGES 109-115

RESPONSE

GOOD SCHOLARLY INTENTIONS DO NOT GUARANTEE GOOD POLICY

Greg Mitchell[*]

PROFESSOR Bartlett has written a bold article pushing back against what might be called inchoate or half-hearted empiricism.[1] The half-hearted empiricist recognizes the value of empirical evidence to help solve a legal problem but, for whatever reason, fails to acknowledge the complexity and uncertainties of the evidence and as a result offers haphazard prescriptions. Professor Bartlett's article demonstrates what a whole-hearted commitment to empiricism looks like: it involves an engagement with primary sources rather than a reliance on secondary sources (or tertiary sources in the form of law review summaries of secondary sources), a review of research relevant to a problem rather than a review of a subset of research focused on one particular aspect of a problem, and a struggle to find usable prescriptive lessons in a literature that ranges from basic-level research with little obvious real-world application to applied research that can be so situation-specific that its generalizability can be questioned. Whole-hearted empiricism is hard, messy work, as well as a frustrating undertaking. Unlike theory- or model-based approaches to policy that self-consciously simplify the world and avoid many empirical complications, an evidence-based approach must confront empirical complexities and try to find meaning in the midst of ever-changing empirical evidence, imperfect studies, and

[*] Daniel Caplin Professor of Law, University of Virginia, greg_mitchell@virginia.edu.
[1] Katharine T. Bartlett, Making Good on Good Intentions, 95 Va. L. Rev. 1893 (2009).

contradictory findings located among the fragmented social sciences. Consequently, rarely will consensus exist on the prescriptive meaning of the assembled evidence. It should thus not be surprising that, after saying a few words on the dangers of half-hearted empiricism, I raise some questions about the empirical conclusions and recommendations offered by Professor Bartlett and suggest that her attempt to foster good intentions might more usefully be focused on firms rather than managers.

A LITTLE EMPIRICISM CAN BE A DANGEROUS THING

The present concern among employment law scholars about implicit bias as a cause of workplace discrimination no doubt arises from good intentions: the desire to inform legal policy using the latest empirical evidence on why minorities and women may continue to lag behind white males on various aggregate indicators of workplace outcomes. When prominent psychologists, sociologists, and science writers proclaim unconscious biases against minorities and women to be a likely cause of these workplace disparities, and when these empiricists partner with prominent legal scholars to provide helpful overviews of the relevant empirical research in top law reviews,[2] it is understandable why Professor Bartlett and other legal scholars start from the assumption that much workplace discrimination today involves subtle behaviors motivated by implicit bias. Where Professor Bartlett departs from some others working in this area is in going beyond the empirical research on implicit bias to ask how this work fits with empirical research on motivation and the debiasing of workplaces. From a review of this additional research, Bartlett concludes that some of the prescriptions offered by legal scholars are likely to be ineffective against implicit bias and, more distressingly, counterproductive by undercutting the intrinsic motivation of managers to act without bias. If Bartlett is right, then these well-intentioned scholars are advocating policies that not only may fail but may make matters worse.

The potential negative consequences of inchoate empiricism become more serious when such empiricism influences legal opinions. For instance, the dissenters in *Ricci v. DeStefano* embraced the use of assessment centers as a selection tool that they considered likely to have less adverse effects than written tests of the kind used by the New Haven

[2] See, e.g., Anthony G. Greenwald & Linda Hamilton Krieger, Implicit Bias: Scientific Foundations, 94 Cal. L. Rev. 945 (2006).

Fire Department.[3] What these justices failed to acknowledge, however—perhaps because they limited their research review to the incomplete information contained in the appellate record—is that there is little systematic evidence that assessment centers reliably reduce adverse impacts on African American candidates. A recent meta-analysis (or systematic quantitative review) of twenty seven studies examining subgroup differences in assessment center evaluations found "that Black-White mean differences for assessment centers are not as small as has been suggested by a number of researchers" and concluded that much more research is needed on this important topic.[4] Thus, if municipalities turn to assessment centers in hopes of avoiding the problems experienced by the City of New Haven, they may be sorely disappointed.

Half-hearted empiricism may be a function of deference to empirical experts pushing promising theories, the effort to say something fresh in the crowded legal scholarship market, or the opportunistic use of studies that support one's preferred causal or remedial theories. Whichever is the case, the result is likely to be the same: important qualifications to the empirical research will be neglected and inapt prescriptions offered. A sincere commitment to the formulation of evidence-based antidiscrimination policy requires a hard look at a wide range of empirical research on both the causes of discrimination *and* its cures, and it requires that scholars understand the limits of this research to avoid oversimplifying complex problems that may have context-dependent or yet-to-be-found solutions.

A LOT MORE EMPIRICISM MAY STILL NOT BE ENOUGH

Of course, whole-hearted empiricism is no guarantee of good policy either, given the contested and incomplete nature of much of the empirical evidence. Consider the key, interrelated empirical claims in Professor Bartlett's discussion of how to regulate workplaces to reduce the im-

[3] Ricci v. DeStefano, 129 S. Ct. 2658, 2704–05 (2009) (Ginsburg, J., dissenting).

[4] Michelle A. Dean, Philip L. Roth & Philip Bobko, Ethnic and Gender Subgroup Differences in Assessment Center Ratings: A Meta-Analysis, 93 J. Applied Psychol. 685, 690 (2008). The researchers did find a smaller Black-White difference in ratings of incumbents relative to applicants (i.e., for promotion as opposed to hiring), but differences remained and the incumbent sample sizes were not as large, meaning that the result is less reliable. See id. at 688. The larger point brought out by this quantitative review was the sparseness of data addressing subgroup differences in assessment center outcomes and the need for more research to reach more reliable conclusions. See id. at 689.

pact of subtle, even unconscious, biases on human resource decisions: (1) stronger legal prohibitions against discrimination are likely to crowd out intrinsic motivations to treat all workers equally, leading to greater overall discrimination because of the difficulty of proving discrimination in a world of subtle discriminatory behaviors motivated by stereotypes and prejudicial attitudes operating beneath or at the fringes of consciousness;[5] (2) recent research points to effective means to reduce subtle discrimination that will not crowd out positive intrinsic motivations.[6] A closer look at the underlying research raises questions about whether Professor Bartlett's proposed workplace policies will work as advertised.

Drawing on the literature on norm internalization and intrinsic motivation, Bartlett argues that strong legal sanctions against subtle, uncon-

[5] This claim actually depends on the sub-claim that individuals are often intrinsically motivated to avoid acting in discriminatory ways. I endorse a related view and discuss relevant evidence elsewhere. See Gregory Mitchell, Second Thoughts, 40 McGeorge L. Rev. 687, 702–09 (2009).

[6] Professor Bartlett also assumes, as do several other scholars recently, that much workplace discrimination is caused by subtle behaviors motivated by implicit bias. I will not address that assumption here, since I and my co-authors have discussed it at length in other places. See Hart Blanton, James Jaccard, Jonathan Klick, Barbara A. Mellers, Gregory Mitchell & Philip E. Tetlock, Strong Claims & Weak Evidence: Reassessing the Predictive Validity of the Race IAT, 94 J. Applied Psychol. 567, 567–68 (2009); Mitchell, supra note 5, at 687–89; Gregory Mitchell & Philip E. Tetlock, Antidiscrimination Law and the Perils of Mindreading, 67 Ohio St. L.J. 1023, 1023–34 (2006); Philip E. Tetlock & Gregory Mitchell, Implicit Prejudice and Accountability Systems: What Must Organizations Do to Prevent Discrimination?, 29 Research in Org. Behav. 3–38 (Barry Staw & Arthur Brief eds., 2009); Philip E. Tetlock & Gregory Mitchell, Calibrating Prejudice in Milliseconds, 71 Social Psychol. Q. 12, 12 (2008); Philip E. Tetlock, Gregory Mitchell & Terry L. Murray, The Challenge of Debiasing Personnel Decisions: Avoiding Both Under- and Overcorrection, 1 Indus. and Organizational Psychol. 439 (2008). It is my position that implicit bias research is very important research that does not presently explain actual personnel decisions made in complex, uncontrolled environments where there are many potential influences on managerial judgments and decisions. Further, there exists little empirical evidence to support the supposed cures being offered for discrimination supposedly motivated by implicit bias. Were the focus on implicit bias harmless and the imposition of possibly ineffective cures for it costless, then my concerns about legal uses of the implicit bias research would vanish. But in a world of limited resources for workplace regulation and of over-reaching opinions by experts in class action litigation, this focus on implicit bias is not costless or harmless. See also Ralph Richard Banks & Richard Thompson Ford, (How) Does Unconscious Bias Matter? Law, Politics, and Racial Inequality, 58 Emory L.J. 1053, 1122 (2009) ("Because it misdescribes the problem of racial injustice, unconscious bias theory inspires misguided reform efforts. It fuels fruitless attempts to ferret out individual bias and places too much emphasis on individual acts of discrimination. Ultimately, the unconscious bias approach may obscure, or even undermine, the substantive goals of racial justice.").

sciously-motivated discrimination may be seen as unfair and intrusive, may compromise managers' feelings of self-determination with respect to fair treatment, and may thus "crowd out" intrinsic motivations to act without bias.[7] Bartlett proposes that workplace interventions be designed to "crowd in" positive intrinsic motivations by creating situations that facilitate positive intergroup interactions, encourage the adoption of egalitarian norms in these interactions, and permit managers and coworkers to feel that they acted fairly because it was the right thing to do rather than to avoid punishment. Of course, the devil is in the details for the organization or policymaker looking for guidance on how to make this broad proposal a reality. And when we turn to the research that Bartlett discusses to guide the formulation of policy, it is not clear that it provides sufficient or consistent guidance.

Much of the research on the reduction of implicit bias that Bartlett discusses has not yet been shown to be effective in workplaces, as Bartlett acknowledges when noting the potential value rather than the proven track record of this work. Reliance on largely experimental work is necessary because surprisingly little research has examined what actually works inside organizations: "Whereas there has been a great deal of research on the sources of inequality, there has been little on the efficacy of different programs for countering it. . . . We know a lot about the disease of workplace inequality, but not much about the cure."[8] Ideas that look good in the laboratory when tested on college students making low-stakes or hypothetical decisions in limited interactions may run into considerable complications when moved to real workplaces. A perennial problem in human resource management is convincing practitioners to embrace practices that have not been field tested and shown to be suc-

[7] As Bartlett notes, there is some debate about the power of the crowding-out effect and the conditions under which it occurs. Space constraints do not permit an in-depth discussion of this research. For entry into this literature, see Samuel Bowles, Policies Designed for Self-Interested Citizens May Undermine "The Moral Sentiments": Evidence from Economic Experiments, 320 Science 1605 (2008); Bruno S. Frey & Reto Jegen, Motivation Crowding Theory, 15 J. Econ. Surveys 589, (2001); Marylène Gagné & Edward L. Deci, Self-determination Theory and Work Motivation, 26 J. Organizational Behav. 331 (2005).

[8] Alexandra Kalev, Frank Dobbin & Erin Kelly, Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies, 71 Am. Soc. Rev. 589, 590 (2006). See also Elizabeth Levy Paluck & Donald P. Green, Prejudice Reduction: What Works? A Review and Assessment of Research and Practice, 60 Ann. Rev. of Psychol. 339, 357–58 (2009) ("We currently do not know whether a wide range of [prejudice-reduction] programs and policies tend to work on average, and we are quite far from having an empirically grounded understanding of the conditions under which these programs work best.").

cessful within their industry.[9] To the extent that many of the practices Bartlett recommends would need to be adopted voluntarily rather than imposed on a company, the evidence that Bartlett presents is unlikely to be persuasive to many HR managers.

Furthermore, it is not clear that all of the research that Bartlett touts fits with her crowding-in thesis. In particular, the organizational research by Kalev and Dobbin finds that the centralization of authority and responsibility for diversity planning and the imposition of affirmative action plans are among the most effective means to improving the number of women and African Americans in private sector management.[10] These practices involve external pressure and oversight of managers by others inside the firm. It is certainly possible that successful diversity officers and affirmative action plan administrators impose their directives in ways that do not interfere with ordinary managers' feelings of autonomy and competence or intrude on managers' intrinsic motives to enhance diversity, but that remains an open question.

REFOCUSING ON THE FIRM

Professor Bartlett's article offers good reasons to be cautious about legal proposals that employ the simple economic logic that increased external sanctions can compensate for the low probability of detection and punishment of unconsciously motivated discriminatory behavior. Such sanctions may cause some managers to increase their supply of unbiased behavior, but they may backfire in some cases and cause managers who would otherwise act fairly to reduce their supply of unbiased behavior. Unfortunately, the net effects of increased external sanctions are difficult to estimate from the existing evidence.

Professor Bartlett focuses her inquiry on how legal sanctions may affect managerial norm internalization and intrinsic motivation, but her argument can usefully be scaled up to the level of the firm. As I see it, the

[9] See, e.g., David E. Guest, Don't Shoot the Messenger: A Wake-up Call for Academics, 50 Acad. Mgmt. J. 1020, 1023 (2007) ("Another major concern expressed by these managers was the desire for evidence of successful application from other organizations in their own sectors."); Edward E. Lawler III, Why HR Practices Are Not Evidence-Based, 50 Acad. Mgmt. J. 1033, 1035 (2007) ("Gathering the type of data that will influence practice usually requires field research that uses an action or evaluation research approach." (citation omitted)).

[10] See Alexandra Kalev, Frank Dobbin & Erin Kelly, Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Programs, 71 Am. Soc. Rev. 589, 602–07, 611 (2006).

central crowding-out problem in antidiscrimination law today is that even those companies that may want to engage in the sorts of self-critical review and program evaluation necessary to determine whether their antidiscrimination and diversity-enhancement policies are working as desired have strong externally imposed incentives not to do so. Evidence generated by in-house audits and organizational culture reviews may be subject to discovery, adopting evidence-based policies proven to reduce workplace disparities will not immunize companies from attacks by plaintiff experts who can always find some aspect of a personnel process to criticize, whether merited or not, and there are no formal legal benefits from engaging in self-critical review.[11] Companies are in the best position to detect and correct workforce disparities. We need to create a set of legal rules and policies that reward internal monitoring and self-correction and that penalize deliberate ignorance about the discriminatory impact of a company's personnel policies.[12] In short, we need to consider how best to crowd in positive intrinsic motivations and crowd out negative intrinsic motivations at the firm level, perhaps more so than at the managerial level.

[11] Of course, there may be factual advantages depending on how the company reacts to the products of its self-monitoring. For instance, a company that acts proactively in response to negative findings will surely reduce an already limited risk of punitive damages being awarded against it. For a fuller discussion of the hurdles companies face to engaging in self-monitoring and self-correction, see Susan Sturm, Second Generation Employment Discrimination: A Structural Approach, 101 Colum. L. Rev. 458, 543–46 (2001).

[12] Professor Sturm has undertaken important work in this regard:

> [T]he practice that has emerged in forward-looking workplaces offers fertile ground for beginning to elaborate principles of effective, accountable, and legitimate workplace processes, along with the regulatory framework that would encourage their development. Indeed, the challenge of my project is to enable the courts, along with administrative agencies and nongovernmental organizations, to encourage the evolution of accountable and legitimate internal problem-solving processes. Increasing evidence exists of employers who develop effective systems for identifying, addressing, and minimizing particular manifestations of workplace bias.

Id. at 491 (2001) (footnotes omitted). Continuing Sturm's project and trying to figure out how to encourage more companies to self-regulate effectively, without losing ground for women and minorities in those companies that will not, should be a priority for empirical researchers and employment law scholars.