

**Metrics, Diversity, and Law  
Papers and Proceedings of a Conference of  
the Research Group on Legal Diversity  
May 5 – 6, 2016  
Chicago, IL  
American Bar Foundation**



**Editors  
Atinuke Adediran and Robert L. Nelson  
Online Publication August 30, 2018  
[http://www.americanbarfoundation.org/research/  
project/1119](http://www.americanbarfoundation.org/research/project/1119)**

Editors' Note: The papers in this volume are conference drafts, may contain manuscript errors, and are subject to change. Authors retain copyright.

## Table of Contents

Acknowledgements .....	3
Chapter 1: Introduction .....	5
<b><u>PART I: Legal Dimensions of Standards and Measurement</u></b>	
Chapter 2: Diversity Structures as Symbolic Metrics in the Federal Courts / <i>Brent K. Nakamura and Lauren B. Edelman</i> .....	12
Chapter 3: Litigation, Diversity, and Metrics / <i>Richard Tonowski</i> .....	63
Chapter 4: Marking to Benchmark: Using Peer Comparison Standards to Measure Diversity and EEO Performance / <i>Pamela Coukos</i> .....	102
<b><u>PART II: Innovations in Metrics and Organizational Performance</u></b>	
Chapter 5: How effective are organizational diversity initiatives? An illustrative evaluation of the National Football League’s Rooney Rule / <i>Christopher Rider, James Wade, Anand Swaminathan and Andreas Schwab</i> .....	133
Chapter 6: Solving the Legal Diversity Problem / <i>William D. Henderson</i> .....	175
Chapter 7: Measuring Employee Health: Balancing Costs, Control, and Controversy / <i>Jennifer Bennett Shinall*</i> .....	201
<b><u>PART III: Standards and Measurements in Business, Organizations, and Law Firms</u></b>	
Chapter 8: Excerpt from the IILP Review (2017): <i>An Innovative Approach to Hiring Lawyers: One Firm’s New Program Reflects Its Firm Values and Eliminates Implicit Bias</i> / <i>Lisa A. Brown</i> .....	253
Chapter 9: Excerpt from <i>Diversity in Action: A Manual for Diversity Professionals in Law – The Measure of Success or What Matters Is Measured and What Is Measured Matters</i> / <i>Theresa D. Cropper and Anna L. Brown</i> .....	261
Chapter 10: Diversity and Inclusion Standard and Practice in the 21st Century / <i>Effenus Henderson</i> .....	275

**Chapter 11: Diversity & Inclusion - A Service Academy Approach to Assessments / Patricia L. Williams ..... 298**

**Appendix of Conference Materials..... 318**

**Acknowledgements**

We gratefully acknowledge the support of many co-sponsors and supporters for this conference. Several entities at Northwestern University contributed to the effort: Pritzker School of Law, Kellogg School of Management, the Office of the Provost, The Graduate School, and Weinberg College of Arts and Sciences. Microsoft Corporation provided lead funding for the conference. Allstate Insurance and Kirkland & Ellis LLP gave significant donations. The conference would not have been possible without this support.

This conference benefitted from a major diversity research initiative of the American Bar Foundation, which culminated in the creation of the William H. Neukom Fellows Research Chair in Diversity and Law. The inaugural holder of the chair, Professor Rachel F. Moran of UCLA Law School, participated in the conference. We gratefully acknowledge all the contributors to the Neukom Chair.

Several individuals made notable contributions to the planning and execution of this event. Spencer Headworth, the Graduate Coordinator for the Research Group on Legal Diversity, was instrumental in all aspects of the conference—both logistical and intellectual. Lucinda Underwood, ABF Director of Development and Communications, led the critical outreach and media relations work on the conference. Among many ABF board members who supported the conference were ABF President David Collins, board member Jimmy Goodman, and board member Kay Hodge. Then ABA President Paulette Brown delivered opening remarks, as did ABF Executive Director Ajay Mehrotra. Ronit Dinovitzer and David Wilkins,

co-directors of the Research Group on Legal Diversity, gave intellectual encouragement to the effort.

A full list of conference presenters and attendees is shown at the American Bar Foundation website [http://www.americanbarfoundation.org/uploads/cms/documents/contact\\_information\\_list.pdf](http://www.americanbarfoundation.org/uploads/cms/documents/contact_information_list.pdf).

We want to single out a few individuals from business, government, the media, the legal profession, and the academy: Rima Alaily of Microsoft, Bruce Goldberg of Allstate, Deneese Walia Levin of Kirkland & Ellis, Kathryn Simpson of Northrop Grumman, Alan Bryan of Walmart, Walter Sutton of Walmart, Peggy Davis of the Chicago Community Trust, Leslie Richards-Yellen of Hinshaw and Culbertson, Paulette Brown-President of the American Bar Association, Robert Morse of U.S. News & World Report, Aric Press of Bernero & Press, Richard Tonowski of the Equal Employment Opportunity Commission, Lee Webster of the University of Texas Medical Branch, Rina Alvarez of Kirkland & Ellis, Sandra Yamate of the Institute for Inclusion in the Legal Profession, Sally Olson of Sidley & Austin, Jon Geier from Paul Hastings, Brian Levine of Mercer Consulting, Swethaa Ballakrishnen of NYU Abu Dhabi, Carole Silver of Northwestern University, William Bielby of the University of Illinois-Chicago, Lauren Edelman of the University of California-Berkeley, Wendy Espeland of Northwestern University, Michael Sauder of the University of Iowa, and many others. The convening of these individuals from diverse backgrounds, with diverse interests, and with dedication to advancing our understanding of diversity metrics made for a stimulating conference.

## Chapter 1: Introduction

On May 5<sup>th</sup> and 6<sup>th</sup>, 2016, the Research Group on Legal Diversity of the American Bar Foundation convened a conference on Metrics, Diversity, and Law. The modern world is replete with schemes of measurement and assessment. In the law and other professions, metrics play key roles in the decision-making processes of gatekeepers at critical junctures throughout careers, and in drawing conclusions about the successes and failures in efforts to advance diversity and inclusion. The conference brought together the social scientists and legal academics who study how metrics are employed in law and organizations; diversity professionals who seek to use metrics in their work; metrics experts and standards writers who seek to develop a set of agreed upon best practices for the use of metrics; attorneys who advise clients on the promise and perils of using metrics concerning matters of diversity and inclusion; policymakers who seek to employ metrics to advance goals of equal opportunity; journalists who are an important source of and publisher of metrics; and leaders of bar associations, law firms, and corporate law departments who are interested in using data to advance diversity within the legal profession.

The conference produced a rich set of presentations and discussions that highlighted important debates concerning the use of metrics relating to diversity and law. A full reporting of those discussions is available in links in the Appendix to this collection and on the website of the Research Group on the Legal Diversity on the ABF website.

<http://www.americanbarfoundation.org/research/project/1119>

The purpose of this online volume is to present in one place the papers submitted by some authors after the conference. We have organized the papers in three parts. Part I deals with the legal dimensions of standards and metrics. Brent Nakamura and Lauren Edelman argue

that organizations equate diversity structures, that is, symbolic metrics of diversity, including diversity statements, diversity trainings and policies as compliance with civil rights laws and affirmative action. Judges and courts also equate those structures as compliance with civil rights laws. The authors draw on two representative samples of federal civil rights opinions to determine how judges evaluate diversity structures within organizations. They find that judges fail to evaluate the efficacy of diversity structures and simply take them as an indication of compliance with civil rights laws.

Richard Tonowski, Chief Psychologist for the Equal Employment Opportunity Commission, argues that society has moved from a world of overt discrimination to more subtle forms of marginalization. Therefore, it is important that organizations move beyond mere diversity (metrics) to inclusion (how people are treated). He believes that more cooperative measures may get us towards inclusion than EEO litigation, which is adversarial. He advocates for structural changes within organizations to eliminate root causes of unlawful discrimination, intergroup hostility, and negative stereotyping. He maintains that the focus should be on how people are treated within organizations, rather than just the number of people of a certain gender or race employed within the organization.

Pamela Coukos advocates for constructing an Equal Employment Opportunity or diversity metric in the form of free or low cost national peer organization benchmarks for workplace pay and representation for U.S. based employers, rather than generalized labor market data. With a public benchmarking tool, individual organizations could voluntarily test their data against the benchmarks to set their internal diversity goals. Establishing a free national database of peer benchmarks has the potential of conferring social benefits, such as making it easier for a range of organizations to implement diversity metrics in the workplace, and providing external

stakeholders with useful information about companies, universities, nonprofits and other organizations on their diversity performance. By identifying the highest performing and highest risk levels among a set of comparable employers, investors, regulators, consumer and worker advocates could use the information to highlight best practices and direct resources, pressure or technical assistance accordingly. Job seekers could also incorporate this information into their decision making. Organizations hiring outside counsel could include this factor in their calculus.

The papers in part II examine innovations in metrics and organizational performance. Christopher Rider, James Wade, Anand Swaminathan, and Andreas Schwab analyze the Rooney Rule, established in 2003 by the National Football League, as an example of an innovation intended to spur greater diversity. The authors highlight statistical challenges for measuring the effectiveness of diversity initiatives. The Rooney Rule requires teams to interview at least one racial minority for every open head coaching position. While it was intended to increase minority representation among head coaches, it is unclear whether the Rooney Rule achieved its aim. Before the Rooney Rule came into existence, the authors observed a diversifying candidate pool of lower level coaches that may have influenced the diversification of head coach positions even without the Rooney Rule. This is especially true because one cannot know what would have happened had the Rooney Rule not been implemented. The authors then suggest guidelines that are more likely to produce definitive evidence of the effectiveness of diversity initiatives: clear objectives, a large sample, randomized controlled trials, and time series data.

William Henderson examines why the lack of workplace diversity is a problem that plagues law firms in the United States, especially at the partnership level. Henderson argues that the creation of high performing partners is influenced by five factors and explains how each of these factors can be modified or utilized to drive a diverse pool of candidates: (1) aptitude, also



known as cognitive ability – reducing a law firm’s reliance on academic proxies would increase the number of diverse candidates; (2) motivation, which is primarily a function of values alignment between the lawyer and the substance of his or her work – applying a job-relevant criteria during the interview process in a uniform, structured way to a diverse array of candidates; (3) the type and quality of work experience that a lawyer receives during his or her early career – ensuring that racial minorities have access to quality assignments; (4) the quality, quantity, and timeliness of training and feedback; and (5) the presence and quality of a mentorship or sponsorship relationship.

Jennifer Shinall examines how employers have been discouraged from gathering data on medical records acquired from wellness programs since at least 2000, because of the EEOC’s guidance on the Americans with Disabilities Act (ADA) rule that requires medical records acquired as part of wellness programs to be kept separate from personnel records. The ban was intended to prevent employers from discovering potentially disabling health conditions and using the information to discriminate against employees in the work place. Shinall argues that this ban on matching records prevents researchers and third-party wellness program administrators, in addition to employers, from compiling, analyzing and using these data to determine the efficacy of these programs. In the absence of available data, employers can subscribe to popular assumptions and regard individuals with visible health conditions as less productive than employees without visible conditions. Thus, disability advocates are unable to evaluate wellness programs and their effects on workplace productivity. The new ADA regulation that went into effect in 2016 – and the EEOC directives accompanying it – permitting the measurement and collection of health-related data is a step in the right direction for disability law advocates and scholars alike.

The papers in part III were authored by professionals who have led efforts to create diversity standards and measurement approaches in business, public organizations, and law firms. Lisa Brown discusses Schiff Hardin's innovative interviewing process created to identify talented and diverse new associates and prevent implicit bias in hiring. The process breaks from traditional large law firm hiring procedures of on campus interviewing at select schools to also include interviewing at a wider range of law schools, including one HBCU. In addition, the interview process includes a panel with 3 to 4 partners who ask behavioral questions that explore candidates' experiences solving real-world problems, working with and leading teams, learning new skills, and building successful relationships. Candidates also engage in writing exercises, which measure how they read and digest a legal issue. The results of this innovative program indicate that women and racial minorities perform well in both the panel and writing exercises and are more likely to have performed well after one year as associates. Associates report that they believe that the hiring process at the firm is fairer than those of other law firms.

Theresa Cropper and Anna Brown suggest that metrics are a critical part of diversity efforts in law firms. Metrics can call attention to key processes that determine the success of diverse groups within the firm, including not only hiring and retention, but also the staffing of Requests for Proposals or RFPs and important committees. The authors offer practical suggestions about what data to collect, how often to collect them, how to analyze them, how to ensure they are of high quality, and how to disseminate results within organizations. They advise diversity professionals to have a clear rationale for the data they collect and to carefully analyze diversity data in a way that educates the law firm about the diversity challenges it faces. They point out the potential downsides of mindless data collection or lack of sensitivity in releasing confidential data within the organization.

Effanus Henderson advocates for the creation of Diversity and Inclusion or D&I standards within organizations as diversity benchmarks. Without consistent standards and metrics, organizations' business leaders struggle with their diversity programs and rely only on workforce demographics about race, gender and other groups. Henderson talks about his work as part of the Society for Human Resource Management (SHRM), which convened a group of diversity thought leaders in 2009 to propose standards that define minimally effective diversity and inclusion practices for organizations across the country. Henderson was the co-chair of SHRM and this paper describes part of the work that he and others did until 2015. The paper also discusses growing interests in standards created by other organizations, including, the International Organization for Standardization (ISO).

Captain Patricia Williams highlights the mission of the United States Naval Academy Office of Diversity, Inclusion, & Equal Opportunity. The Office of Diversity is part of the Equity Study and Assessment Committee and the Proportional Outcomes study that assesses, among other things, the effectiveness of diversity and inclusion programs within the Academy. The Office also collects and analyzes data to identify causal factors for attrition rates among women and minorities. In addition, the Office conducts disaggregated data analysis for each race and ethnic category to better understand attrition patterns and make recommendations to college leadership. The Diversity Office also works in conjunction with the Office of Admissions to conduct numerous campus tours and events throughout the year to reach a diverse range of potential students.

These papers and the larger conference from which they came call attention to the continuing importance of metrics in diversity and law. Metrics can be a powerful impetus to change. As such they are inherently political in character. Scholars, policymakers, metrics

experts, and consumers of metrics need to be sensitive to the potential and pitfalls of quantification. We hope these papers and the conference proceedings advance this important discussion.

## **Part I: Legal Dimensions of Standards and Measurement**

**Chapter 2: Diversity Structures as Symbolic Metrics in the Federal Courts /  
Brent K. Nakamura and Lauren B. Edelman**

Jurisprudence & Social Policy Program  
University of California, Berkeley

Prepared for Diversity, Metrics, and Law  
American Bar Foundation

December 2016

The primary metric of diversity today is symbolic. Workplace policies that symbolize diversity – such as diversity statements, diversity training programs, and diversity policies – have become widely accepted indicia of compliance with civil rights laws, irrespective of their effectiveness. When we see company brochures that highlight their diverse workforces or university websites that emphasize their commitment to equity and inclusion, we tend to think of those organizations as fair and nondiscriminatory even though we know little about whether men and women of color and white women have equal access to management and professional positions or are subject to harassment that makes it difficult for them to succeed. Yet too often, corporations and courts alike measure diversity through symbolic metrics – that is, workplace structures that symbolize diversity – rather than through more substantive metrics of diversity such as the workforce representation of women and people of color or the real work opportunities afforded to groups that have traditionally been disadvantaged.

Symbolic metrics of diversity, which we call ‘diversity structures’ include corporations’ visible commitments to diversity, equal employment opportunity (EEO), or fair governance such as: diversity training programs, diversity mission statements, antidiscrimination or antiharassment policies, complaint procedures, and more generally, formalized organizational structures that are associated with fair governance such as progressive discipline policies, formal evaluation procedures, and multi-person decision making panels.

Diversity structures became increasingly common over the twentieth century. Edelman (2016) has shown that employers responded to the ambiguity of Title VII and other civil rights legislation of the 1960s with a variety of symbolic structures designed to symbolize attention to

civil rights ideals. Structures such as policies banning discrimination and later sexual harassment, grievance procedures, affirmative action offices and officers, and affirmative action recruitment and training programs quickly diffused throughout organizational fields. Later, as the term ‘diversity’ came to replace attention to ‘equal employment opportunity,’ (Edelman, Fuller, and Mara-Drita 2001), many organizations created diversity training programs, diversity offices, and diversity policies.

Virtually every organization now has a public commitment to diversity and inclusion as well as policies that purport to prohibit discrimination and harassment and complaint procedures that allow employees to complain about instances of discrimination or harassment. Many organizations have created antiharassment or diversity training programs and indeed some states mandate the use of these programs. Compliance professionals, professional organizations such as the Society for Human Resource Management (SHRM), management consulting firms, and insurance companies that sell employment practices liability insurance all strongly urge firms to have standard and standardized diversity structures in place. In the twenty-first century, diversity commitments and policies are standard and firms that lack such structures look suspect. In some cases, these symbolic metrics have helped organizations achieve substantive diversity by offering more and better employment opportunities to those who were disadvantaged by earlier discrimination, by reducing discrimination, or by encouraging more inclusive workplaces. But in other cases, workplace diversity policies are empty symbols that coexist with informal organizational practices that continue to place women, people of color, and other minorities at a disadvantage relative to white men. In these organizations, diversity structures serve to mask, rather than to substantively combat, discrimination and harassment.



Despite the substantial variation in the extent to which workplace diversity structures actually constrain discrimination and inequality, courts tend to view the mere presence of these structures as evidence of good faith efforts to achieve racial and gender diversity (Edelman 2016). When courts measure diversity by the mere presence of organizational policies and practices without serious scrutiny of the effectiveness of these structures, they unwittingly perpetuate a lack of substantive diversity.

In this article, we build on Edelman's earlier work on judicial deference to symbolic structures (Edelman et al. 2011; Edelman 2016) by focusing on the changes that have occurred in judicial deference to diversity structures since 2000 and by using case examples to show how and why those changes have occurred. In particular, we show how the assumptions that judges make about standards of proof and the drawing of inferences make judges more likely to defer to diversity structures and, therefore, place employment discrimination plaintiffs at a significant and unwarranted disadvantage.

### **Why Diversity Structures are not Evidence of a Diverse Workforce or a Nondiscriminatory Workplace**

Diversity structures are symbolic in that they are imbued with meaning; they invoke a sense of compliance and of the ideals underlying civil rights law. The creation of diversity is only the beginning of the process through which organizations define the meaning of compliance. Once in place, diversity structures become the sites in which the requirements and meaning of law are confronted and negotiated in the context of everyday organizational events. Where legal ideals conflict with business goals, compliance professionals tend to interpret the meaning of legal requirements in ways that render law closer to business values and managerial prerogatives. Over

time, the meaning of law tends to be understood in ways that incorporate managerial logic, values, and ways of understanding the world (Edelman 2016).

In some cases, compliance professionals are enthusiastic proponents of legal ideals, sometimes becoming activists who often confront organizational officials or employees who appear to be violating the law. Where compliance professionals are granted authority and autonomy, and where these structures are designed to achieve specific goals, diversity structures may be catalysts that engender the institutionalization of legal values within organizations (Kalev, Dobbin, & Kelly 2006). In some such cases, organizational efforts at compliance may even exceed what was envisioned by proponents of the law (Edelman and Petterson 1999; Kalev, Dobbin & Kelly 2006). In many other cases, however, compliance professionals who are steeped in the logic of organizational fields are likely to resolve conflicts between legal and organizational logics in ways that subtly introduce business logic into the meaning of law. Edelman (2016) argues that as this occurs, law becomes managerialized within organizations and diversity structures move further from substance and closer to pure symbolism. The transformation is gradual and subtle, and rarely involves conscious decisions to circumvent the law.

Edelman (2016) defines managerialization as the gradual infusion of managerial or business ideals into understandings of law. Managerialization can result from intentional efforts to circumvent legal requirements but it is more often the unintentional result of addressing everyday problems in ways that subtly infuse law with managerial values and objectives. As compliance professionals use professional networks to fill in the details that law has left

ambiguous, “law” within organizations acquires a managerial flavor that may differ in important ways from law within the public legal order. Law, in other words, becomes managerialized or infused with managerial values and interests.

As managerialization occurs within organizations, diversity structures may become less effective and may coexist with informal practices that diverge from organizations’ diversity commitments and formal antidiscrimination policies. Edelman (2016) specifies four ways in which managerialization can weaken diversity structures: (1) internalizing dispute resolution; (2) legislating, contracting, or managing away legal risk; (3) decoupling legal rules from organizational activities; and (4) rhetorically reframing legal ideals. These four forms of managerialization may coexist within particular organizations, and certainly coexist within organizational fields.

One of the most common diversity structures today is the complaint procedure. In theory, complaint procedures provide notice to employers that potential legal problems exist and allow employers to respond to those problems swiftly and effectively. When employers take their complaint procedures seriously, they can often resolve complaints quickly and avoid litigation. Yet there are numerous instances where employer threats or organizational culture lead employees to fear retaliation and therefore to avoid complaining. Although it now violates EEOC guidelines, some employers’ complaint procedures require employees to file a complaint first with their immediate supervisor, who is often the perpetrator of harassment or discrimination. In other cases, complaints – even when filed correctly – are ignored or evaluated in a cursory or biased manner.

Internal complaint procedures tend to managerialize the law as complaint handlers subtly reframe law as consistent with good management and complaints of rights violations as instances of poor management. Thus, even when corporations take complaints seriously, they rarely recognize employees' legal rights or take action to avoid future rights violations (Edelman 2016).

Complaint handlers, for example, frequently treat allegations of sexual harassment as instances of poor management or complaints of gender discrimination as evidence of personality problems. They take action to resolve the problem but do so under the rubric of remedying personality conflicts or providing counseling rather than of eliminating discrimination or harassment from the workplace. This orientation tends to deemphasize law and rights, and often leads even employees to view their problems as workplace problems rather than as violations of their legal rights (Quinn 2000; Marshall 2003, 2005).

Another form of managerialization occurs as compliance professionals navigate around legal assumptions or standards, in a manner somewhat akin to taking advantage of a tax loophole. Organizations may revise their rules or employment contracts to navigate around legal risk through pre-dispute mandatory arbitration clauses in their employment contracts or employment handbooks, which require employees to waive their right to sue for certain types of violations (Van Wezel Stone 1996; Roma 2011; Resnik 2015). In some cases, employers navigate around legal risk by insuring against the risk of legal liability (Mootz 1997; French 2012; Talesh 2015). Strategies of this type tend to be adopted quickly by other organizations as compliance professionals come to see them as successful ways of circumventing liability.

Perhaps the most common form of managerialization is “decoupling,” or disconnecting organizational practices from formal organizational policies (Weick 1976). This occurs where organizations have formal commitments to diversity or formal policies that ban discrimination or harassment yet fail to implement those commitments or policies into the everyday operations of the organization. Managers may overtly ignore the policies on the books, or decoupling may be more subtle, occurring through subjective standards for hiring or promotion that tend to favor whites or males or those members of minority groups who most successfully assimilate to white society (Carbado and Gulati 2013; Edelman 2016).

The subtlest form of managerialization is through the rhetorical reframing of legal ideals. As law is imported into the organizational setting, managers may reframe legal constructs in ways that alter their meaning in important ways. Through rhetorical reframing, ambiguous or politically-charged legal constructs may be subtly reshaped in ways that render law less challenging to traditional managerial prerogatives or business practices (Edelman, Fuller, and Mara-Drita 2001). The very concept of diversity is a form of rhetorical reframing. From the passage of the 1964 Civil Rights Act until about the late 1980s, the language of managerial compliance efforts centered around the meaning of civil rights, equal employment opportunity, and affirmative action. In the late 1980s, however, talk among managerial professionals began to shift from equal employment opportunity and affirmative action to diversity. Importantly, moreover, diversity rhetoric altered the focus from race and gender equality to diversity rhetoric rhetorically reconstructed management focus in a way that substantially deemphasized race civil rights law and race and gender equality. Managerial articles on diversity deemphasize the focus on civil

rights law and on race and gender equality and instead began to define diversity along other, nonlegal, dimensions such as culture, geographic location, dress style, and lifestyle (Edelman et al. 2001; Berrey 2015).

All of these forms of managerialization tend to render organizational diversity structures less effective, that is to move them closer to becoming merely symbolic and further from substantively achieving legal ideals. There is of course substantial variation across organizations. In some organizations, leaders and managers work hard to reduce discrimination and bias and to provide real opportunity for men and women of color and white women as well as other disadvantaged groups. But importantly, a substantial body of research now shows that, in part due to managerialization, diversity structures are often ineffective and may undermine legal ideals (Edelman, Erlanger, & Lande 1993; Edelman & Petterson 1999; Edelman, Uggen & Erlanger 1999; Edelman, Fuller, & Mara-Drita 2001; Kalev, Dobbin & Kelly 2006; Edelman, Krieger, Eliason, Albiston & Mellema 2011; Edelman 2016).

### **How Judges are Influenced by Diversity Structures**

Judges – just like employers, employees, and compliance professionals – over time come to equate the diversity structures that organizations create in response to civil rights law with the achievement of civil rights in organizations. Even though judges are in theory selected because of their expertise at law and critical thinking, judges are not immune to institutionalized ideas about diversity structures. Research by social psychologists shows that judges employ two types of reasoning. The first type of reasoning is intuitive, quick, and based upon heuristics, whereas the second is deliberative, slower, and based upon rules (Guthrie, Rachlinski, and Wistrich 2007;

Rachlinski 2011). The first type facilitates fast decisions but is subject to error; the second type is more accurate but requires more time, effort, and motivation. Judges are predominantly intuitive decision makers, and because intuitive decisions tend to be quick, automatic, and heuristic based, these decisions are highly subject to error (Guthrie, Rachlinski, and Wistrich 2002).

Diversity structures provide a heuristic for judges, which facilitate the intuitive assessment that all is in order, and that there is no discriminatory behavior. In a series of studies, Cheryl Kaiser and Brenda Major have shown that the presence of diversity structures has a strong influence on whether people believe in organization to be fair. Importantly, their research shows that even when subjects are explicitly told that women or minorities are unfairly disadvantaged in an organization, the mere presence of diversity structures creates an illusion of fairness that causes most people to overlook evidence of unfair treatment. To show that their findings had implications outside of the laboratory, Kaiser and Major replicated their studies using a sample of organizational managers. These managers were asked to list either diversity structures or structures related to environmental sustainability at their organizations and then were asked to review claims of race discrimination filed by African-American employees. Managers who had been primed to think about diversity structures ranked the race discrimination claims as less legitimate than did managers who primed to think about environmental structures. These experiments show that the presence of diversity structures and organizations causes most people to view organizations as fair and to overlook evidence of injustice or discrimination (Kaiser and Major 2006; Kaiser et al. 2013; Dover et al. 2014; Brady et al. 2015).

Kaiser and Major's research suggests that the mere presence of diversity structures is likely to cause judges to overlook evidence of discrimination, including evidence that organizational practices deviate from antidiscrimination policies, that organizations have cultures that promote bias or harassment, or that internal dispute resolution processes are unfair. Participants in these studies were more likely to believe that organizations were fair even when they were presented with evidence of discrimination against women or minorities. Further, white male participants were especially likely to believe that diversity structures indicated fairness (Kaiser and Major 2006; Kaiser et al 2013; Dover et al 2014). Similar studies have not yet been conducted on judges, and it may be that judges are more able than those without judicial experience to avoid the "illusions of fairness" that arose from the presence of diversity structures. However, it also may be that since most judges are disproportionately high status individuals and most are white males, that they are especially susceptible to these illusions.

Both management lawyers and plaintiffs' lawyers contribute to the heuristic, making it more likely that judges will infer nondiscrimination from the mere presence of diversity structures rather than engaging in careful scrutiny of their effectiveness. Management lawyers help to create and to reinforce assumptions about the fairness and rationality of diversity structures when they point to their clients' diversity structures as evidence of good faith, and of an absence of intent to discriminate, and when they cite judicial precedent that legitimates those structures. The more management lawyers rely on diversity structures as evidence of compliance, the more judges are exposed to the idea that these structures advance civil rights values (Edelman 2016).



One might expect that plaintiffs' lawyers would challenge the effectiveness of diversity structures, which might lead judges to become skeptical about the effectiveness of these structures, or at least more aware these structures may not protect employees' civil rights in all organizations. However, plaintiffs' lawyers too rarely challenge the effectiveness of diversity structures, in part because they also tend to accept the symbolic value of those structures and in part because they view organizations with diversity structures in place as difficult targets of litigation precisely because they look fair. Thus plaintiffs' lawyers may, largely inadvertently, discourage employees from pursuing litigation even in cases where managerialization renders diversity structures ineffective.

Even where plaintiffs' lawyers do challenge ineffective diversity structures and even where judges are skeptical of these structures, moreover, managerialization may be difficult to detect in a courtroom. Both plaintiffs' lawyers and judges are generally unaware of the extent to which organizations decouple their practices from formal structures that appear to protect employees from discrimination. Decoupling can be very difficult to prove because discrimination that occurs through subjective decision-making or through on the ground practices of lower level supervisors is far less visible than are the organizations' formal policies prohibiting discrimination. Similarly, the subtle ways in which internal complaint handlers discourage complaints, favor supervisors, or handle complaints in ways that undermine employee's legal rights are far less visible to judges than the fact that the organization has a diversity mission statement, an antiharassment policy, or a complaint procedure in place. Diversity training programs and diversity officers make organizations look responsible, but judges are unlikely to be aware of the way in which diversity rhetoric subtly shifts the focus from race and gender

equality to differences across a wide variety of dimensions or of research showing that these programs are often ineffective. And, of course, judges do not see those disputes in which organizations manage to avoid litigation through mandatory arbitration or the use of contractual language that manages away legal rights.

### **Symbolic Metrics in Court**

To observe the extent to which diversity structures have become symbolic metrics in courts, we analyzed judicial decision making in federal civil rights opinions with particular attention to whether and how judges evaluate diversity structures in organizations. As noted above, we define ‘diversity structure’ broadly to include any organizational structure that courts might interpret as evidence of an organization’s commitment to fair treatment. Some of these structures are explicitly geared toward diversity, such as diversity or equal employment opportunity policies or complaint procedures. Others are more generic structures that are often taken as indicia of fair governance such as progressive discipline policies, evaluation procedures, and multi-person decision-making structures. Our reading of civil rights opinions suggests that judges frequently understand these more generic structures as indicia of organizational rationality and fair governance.

### **Sample and Data**

We draw on two samples of federal civil rights opinions in the district and circuit courts. The first sample, collected by Lauren Edelman and Linda Krieger, is a representative sample of opinions from 1965 through 1999.<sup>1</sup> We selected a 2 percent sample consisting of 1,024 opinions,

---

<sup>1</sup> The pre-2000 sample used the Westlaw search term: (“title vii”) (“age discrimination”/3 “employment act”) (“rehabilitation act”) (“equal pay act”) (american!/3 disabilit!/1 act) (famil!/3

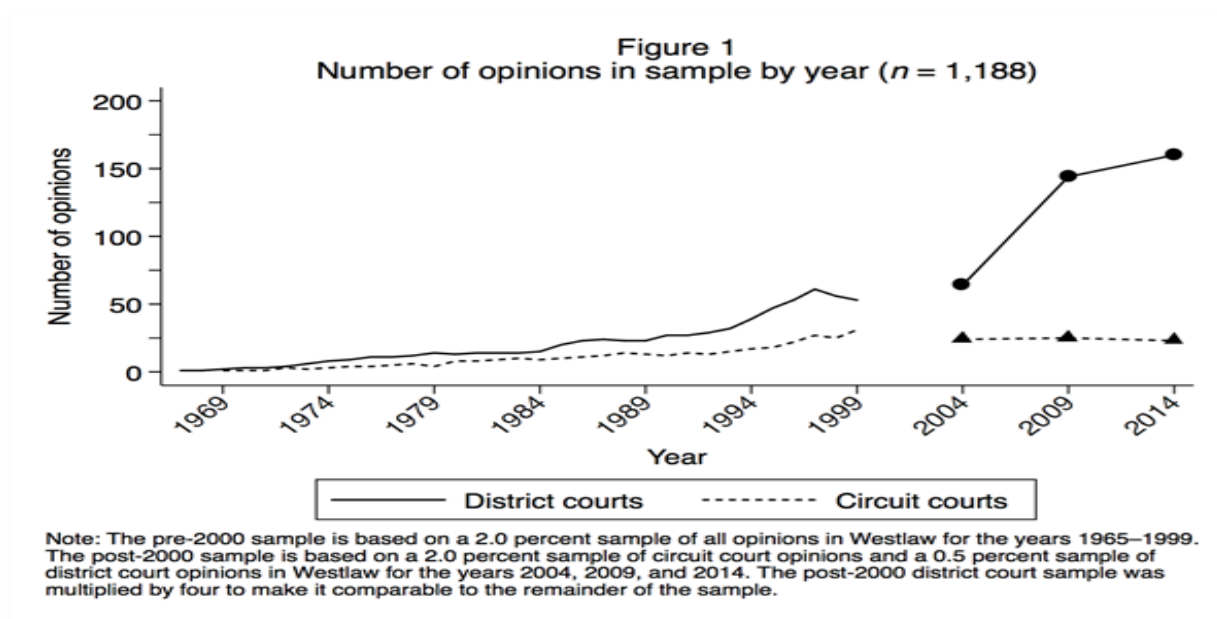
stratified by year and by district or circuit court. I refer to this as the pre-2000 sample. The second sample, which we collected in order to update the original data, is a representative sample of opinions in 2004, 2009, and 2014. We selected a 0.5 percent sample of district court cases and a 2 percent sample of circuit court cases for each of the three years, yielding a final post-2000 sample of 164 cases.<sup>2</sup> Both samples include opinions reported in Westlaw that result from cases brought under Title VII of the Civil Rights Act of 1964, the Age Discrimination in Employment Act of 1967 (ADEA), the Equal Pay Act of 1963 (EPA), and two post-Civil War civil rights statutes: 42 U.S.C. § 1981 and 42 U.S.C. § 1983. We did not include Supreme Court opinions both because of their relatively small number and so that we could treat important Supreme Court cases as independent variables that might influence judicial reasoning in the circuit and district courts. Figure 1<sup>3</sup> shows the number of opinions over time in our combined samples.

---

“medical leave act”) (fmla % (marin! lien))) & DA(aft 1-1-1965 & bef 12-31-1999). This search was performed separately in Westlaw’s Court of Appeals database and in its District Court database and was restricted to cases that were decided between January 1, 1965, and December 31, 1999. The search yielded 34,578 district court opinions and 16,604 circuit court opinions.

<sup>2</sup> For the post-2000 study, we used the same Westlaw search term to generate a universe of cases for the years 2004, 2009, and 2014. The search produced 18,305 district court cases and 3,588 circuit court cases for those three years. We selected a 0.5 percent sample of district court cases and a 2 percent sample of circuit court cases for each of the three years, yielding a final post-2000 sample of 92 district court cases and 72 circuit court cases (for a total of 164 cases).

<sup>3</sup> Figure 1 was previously published in Lauren B. Edelman, *Working Law: Courts, Corporations, and Symbolic Civil Rights* (Chicago: University of Chicago Press, 2016).

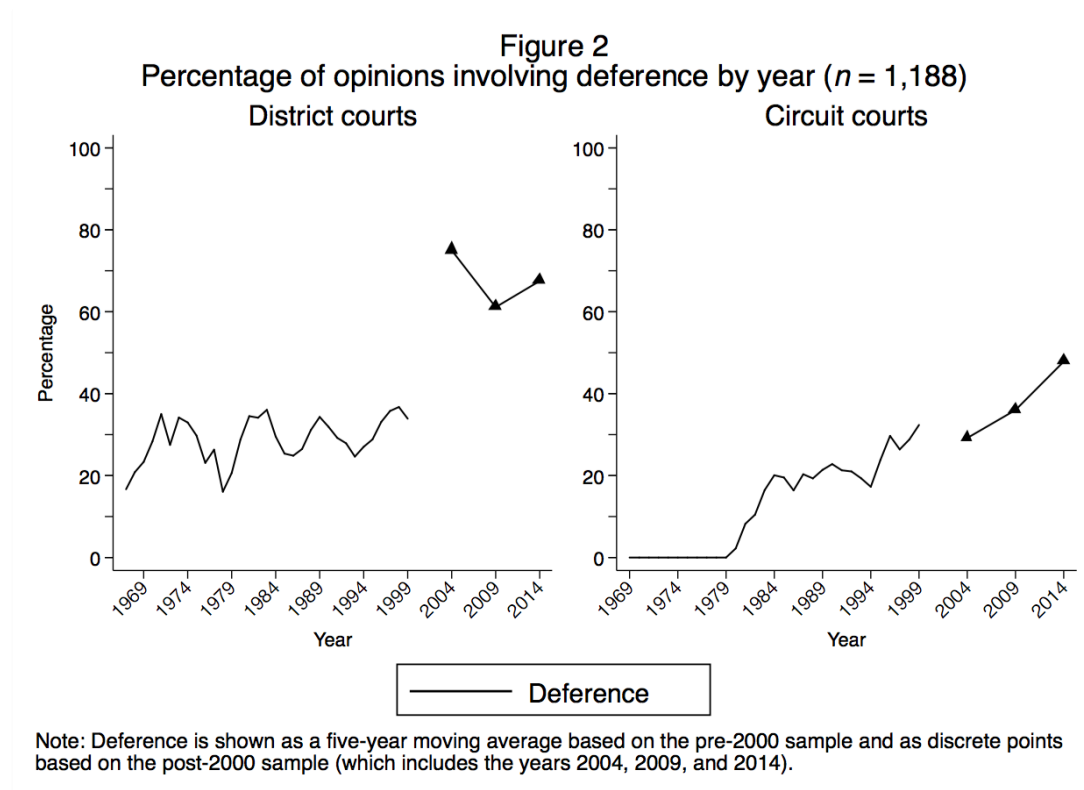


### Judicial Deference to Diversity Structures over Time

We define judicial deference to diversity structures as occurring where the opinion reflects that at least one diversity structure was considered relevant to the question of whether discrimination occurred and any one of the following three conditions existed: (1) the opinion reflects no attention to the adequacy or effectiveness of the diversity structure at all; (2) the opinion states that the diversity structure was ineffective but that the effectiveness of the structure was irrelevant to whether discrimination occurred; or (3) the opinion states that the diversity structure was adequate even though there is also substantial discussion of inadequacies of the structure. An example of the third condition would be if the opinion discusses the fact that a supervisor told an employee that she could not or should not use a complaint procedure and the court then penalizes the employee for failing to use that complaint procedure. If the opinion stated that the

diversity structure was adequate, that opinion would not be coded as involving judicial deference unless there was strong evidence presented that pointed to the inadequacy of the structure.

Figure 2 shows the percent of opinions for district and circuit courts that involved at least one instance of judicial deference to diversity structures without adequacy scrutiny. Figure 2 shows a gradual increase over time in the likelihood that judges would defer to diversity structures without adequate scrutiny through about 1999 and then a substantial increase after 2000. By 2014, judges defer without adequate scrutiny in about 75 percent of district court cases and 49 percent of circuit court cases. The dramatic rise increase in judicial deference without adequate scrutiny indicates that these structures have become, in and of themselves, symbols of organizations’ commitment to diversity. Diversity structures have acquired an aura of legality, irrespective of their effectiveness.



## **Why Judges Have Become More Likely to Accept Diversity Structures as Symbolic Metrics of Diversity?**

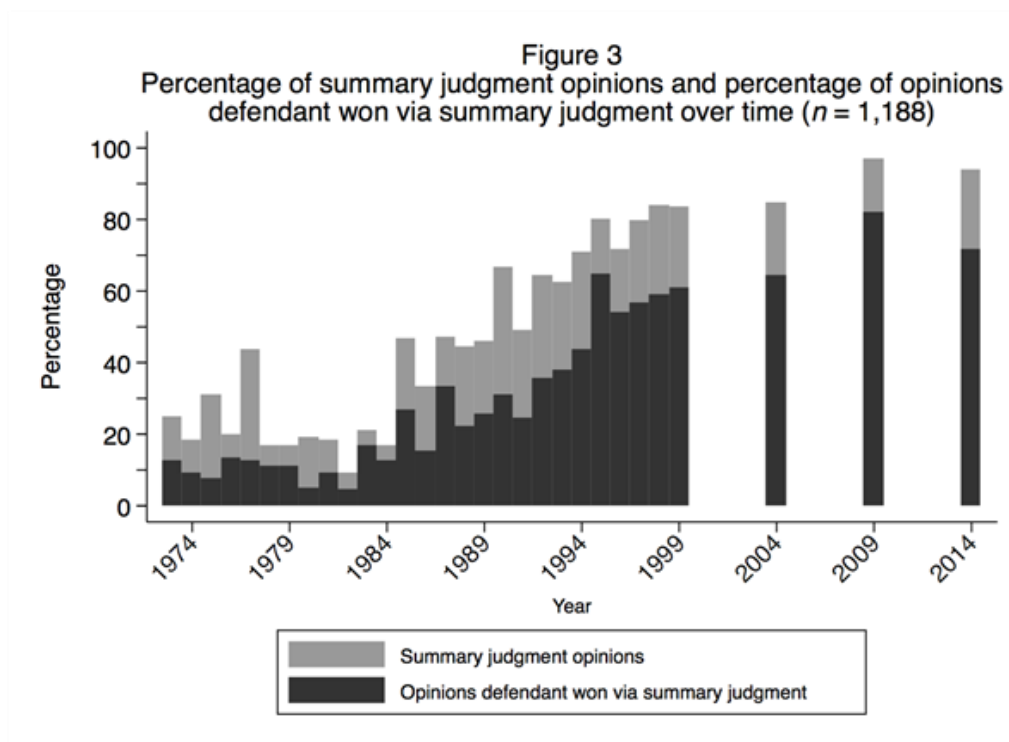
In this section, we offer two key reasons for the rise in judicial deference to diversity structures irrespective of their effectiveness since 2000. The first has to do with the rise of employers' grants of summary judgments in the federal district courts and a concomitant reticence to review those decisions in the federal circuit courts. The second is a result of two US Supreme Court decisions in 1998 that created an affirmative defense to hostile work environment sexual harassment: *Faragher v. City of Boca Raton* (524 U.S. 775, 1998) and *Burlington Industries, Inc. v. Ellerth* (524 U.S. 742, 1998).

### **The Rise in Grants of Summary Judgment to Employers**

The period since about 1990 has seen a significant increase in the proportion of civil rights cases terminated through grants of employers' motions for summary judgments and, in the circuit courts, denials of appeals of district court grants of summary judgment. Figure 3<sup>4</sup> shows that summary judgment cases have increased dramatically over time in both the district and circuit courts. Summary judgment is an increasingly important and frequent manner of case disposition in employment discrimination cases.

---

<sup>4</sup> Figure 3 was previously published in Lauren B. Edelman, *Working Law: Courts, Corporations, and Symbolic Civil Rights* (Chicago: University of Chicago Press, 2016).



Catherine Albiston (1999) argues that employers use summary judgment as a rulemaking opportunity. Pointing out that summary judgment permits piecemeal resolution of the case such as establishing liability without determining damages, she argues that employers use motions for summary judgment strategically to ensure that the early weight of authority addressing a new law will benefit employers. They can do so by settling those cases where employees might be most likely to win in order to avoid precedent that could potentially harm employers' future interests. In addition, as Albiston's argument suggests, because plaintiffs can only prevail on summary judgment if they are able to show that no disputed material facts remain and they are entitled to judgment as a matter of law on each and every element of their claims, plaintiffs' victories on summary judgment are necessarily rare and so are decisions laying down interpretations of the law favorable to plaintiffs. By contrast, in order to prevail on summary judgment, defendants need only show, as to a single element of plaintiffs' claims, that they are entitled to judgment as a matter of law and that no material disputed facts remain. This substantial asymmetry in the

burden of proof and persuasion between plaintiffs and defendants is significant in explaining the frequency of summary judgment motions by and victories for defendants relative to plaintiffs.

Retired federal judge Nancy Gertner offers further insights into the judicial tendency to grant summary judgments, particularly in employment discrimination cases. First, she argues that judges are encouraged to write detailed decisions when granting summary judgment but not to write decisions when denying it, which leads judges to see a skewed distribution of employment cases and hence to trivialize plaintiffs' claims. Second, building on Albiston's (1999) observation that employers settle those cases where plaintiffs have a strong argument, Gertner suggests that because judges rarely see the strongest claims by plaintiffs, they tend to see most employee claims as unjustified and to sympathize with employers (Gertner 2012). Third, echoing an argument originally made by Linda Krieger (1995), Gertner argues that judges tend to look for explicitly discriminatory policies or biased actors but that they fail to understand the many forms of structural and implicit bias that characterize today's workplace (Gertner 2012; Edelman 2016). Finally, while the standard of review for summary judgment orders at the appellate court is *de novo*, Gertner argues that appellate courts rarely reverse district court decisions (Gertner 2012).

### **Judicial Deference to Diversity Structures in Summary Judgment Cases**

Based on the legal standard, one would actually expect less judicial deference to diversity structures in summary judgment cases than in cases that go to trial on the merits. The legal standard for summary judgment requires the court to evaluate the facts and to draw all reasonable inferences in the light most favorable to the nonmoving party, which in civil rights



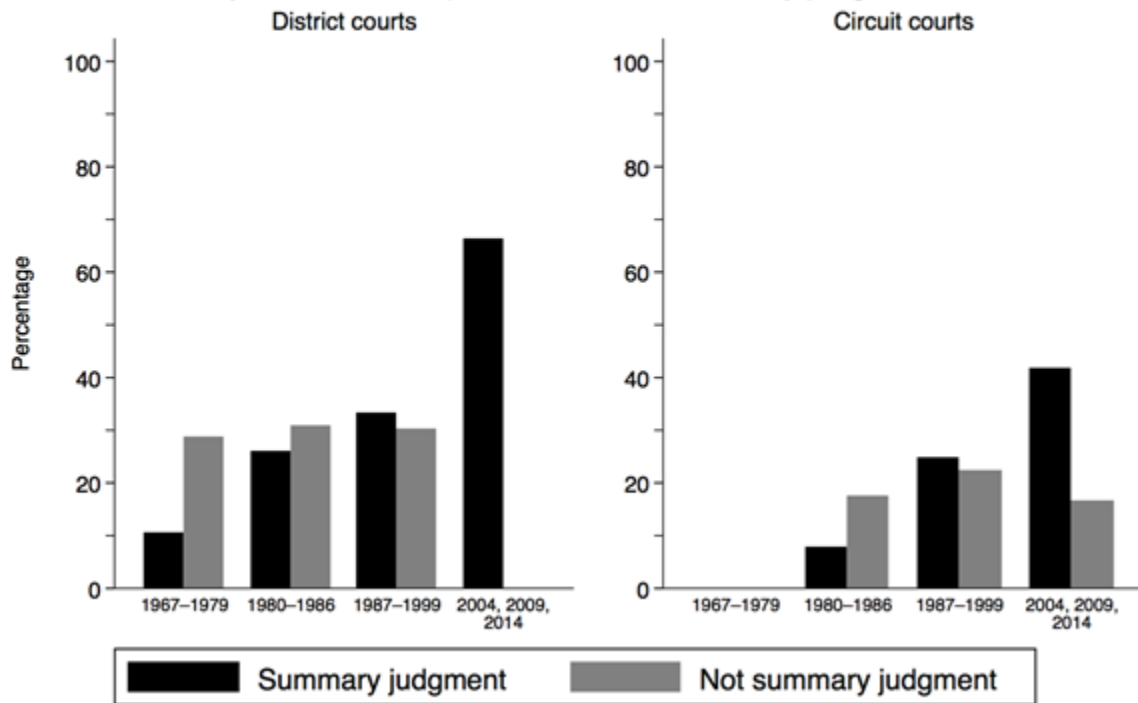
cases is almost always the employee. Because deference involves drawing an inference in favor of the employer on the basis of diversity structures, one would expect less deference in summary judgment actions than in cases resulting in a full trial on the merits. In fact, however, as deference to diversity structures has become more common, judges appear to defer to diversity structures more in summary judgment cases than in fully adjudicated cases.

Figure 4<sup>5</sup> shows the percent of opinions involving deference to diversity structures in summary judgment and non-summary judgment cases over time. After 1986, judges deferred to diversity structures more in opinions involving summary judgment than in opinions that did not involve summary judgment. In the circuit courts, there were relatively few summary judgment cases prior to 1986, when the Celotex trilogy made it easier for employers to prevail in a motion for summary judgment. Because there is less opportunity for judicial attention to the facts in summary judgment cases, judges are more likely to use the mere presence of diversity structures as a heuristic to infer that employers are rational and hence nondiscriminatory. After 2000, deference in summary judgment cases rose significantly, and notably, our random sample of federal district court civil rights opinions included only summary judgment opinions, which is why there is no column for non-summary judgment cases in the district courts after 2000. This finding also points to the difficult plaintiffs face in surviving motions for summary judgment in the district courts.

---

<sup>5</sup> Figure 4 was previously published in Lauren B. Edelman, *Working Law: Courts, Corporations, and Symbolic Civil Rights* (Chicago: University of Chicago Press, 2016).

**Figure 4**  
**Percentage of opinions with deference**  
**by whether the opinion involved summary judgment**



Why do judges defer so frequently to diversity structures in summary judgment cases? We think that judicial deference in summary judgment cases involves drawing inferences of fair treatment from diversity structures, or in other words, using diversity structures as a heuristic for proper governance. While likely inadvertent, such inference drawing amounts to a failure to take seriously the rule articulated by the Supreme Court in *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.* (475 U.S. 574, 587-88), which requires judges to draw all inferences in the light most favorable to the nonmoving party. Judges frequently fail to say anything about how they draw inferences, and they appear frequently to be drawing inferences in favor of the employer, who is in virtually every case the moving party. When judges defer to employers' diversity structures without adequate scrutiny of those structures, they are erroneously drawing inferences in favor of the moving party in violation of the rule articulated in *Matsushita*.

Consider, for example, *Serlin v. Alexander Dawson School* (No. 2:12–CV–1431 JCM (GWF), 2014 WL 1573535 (D. Nev. Apr. 17, 2014)), a case in which the judge misperceived the summary judgment standard and deferred to a company’s diversity structure without any discussion of its adequacy.

Cheri Serlin was a fifty-eight-year-old elementary school teacher. She had been a teacher for 18 years and had taught for four years at Alexander Dawson School, a private kindergarten through eighth-grade school in Las Vegas, Nevada. In 2016, the school noted that its average student to faculty ratio was 8:1.<sup>6</sup> The tuition for fifth-grade in the 2016-2017 school year was \$23,000.<sup>7</sup>

Serlin was diagnosed with breast cancer in 2009 and underwent 18 weeks of chemotherapy while she continued to work at Dawson. As a result of her breast cancer, Serlin was forced to undergo a bilateral double mastectomy, which caused interstitial cystitis. Serlin therefore had to use the bathroom approximately 10-20 times per day as a result of her mastectomy.

Serlin took some leave under the FMLA beginning in 2009, but apparently continued to work for part of that time. Serlin alleged that in 2010, she was harassed and bullied by another faculty member, Julie Tognoni. Serlin said that she was chastised by her co-workers, Tognoni and another co-worker in particular, for her frequent bathroom use, use of leave, and her inability to perform certain tasks that required some degree of physical exertion. Serlin alleged that Tonogi

---

<sup>6</sup> The Alexander Dawson School at Rainbow Mountain, Las Vegas, Nevada, About Dawson: Who We Are. Available at: <http://www.adsrm.org/page/about-dawson> (last accessed Nov. 25, 2016).

<sup>7</sup> The Alexander Dawson School at Rainbow Mountain, Las Vegas, Nevada, Tuition & Fees. Available at: <http://www.adsrm.org/page/admissions/tuition--fees/tuition--fees> (last accessed Nov. 25, 2016).

made derogatory comments about her religion as well as ethnic slurs, including repeated references to Serlin being from the “Bagel Belt,” which Serlin interpreted as an anti-Semitic remark. Tonogi allegedly knew Serlin was from Skokie, Illinois, where a large population of Jews and other Eastern Europeans lived.

Serlin filed a written complaint, contending that the term ‘Bagel Belt’ constituted an insulting ethnic slur. She then met with her supervisor, Russell Smith,<sup>8</sup> who told her that because he could not find “Bagel Belt” on the internet he did not believe it to be derogatory. According to Serlin, during the meeting, Smith “raised his voice sternly and intimidated Serlin by saying ‘do you really want to make something out of this?’” Serlin’s FMLA leave was renewed in February 2011. One month later, in March 2011, Serlin was told that her teaching contract would not be renewed. Serlin’s replacement was twenty-nine year-old ten-year veteran teacher, Angie Vetter. Serlin filed suit against Dawson and its affiliated entities in August 2012 alleging: interference with and unlawful retaliation under the Family and Medical Leave Act (FMLA), 28 U.S.C. § 2615(a)(1) and (2); violation of the Americans with Disabilities Act (ADA), 42 U.S.C. § 12111 et seq.; religious discrimination and hostile work environment based on religion in violation of Title VII; retaliation in violation of Title VII; age discrimination in violation of the Age Discrimination in Employment Act (ADEA), 26 U.S.C. § 621 et seq.; and two Nevada state law claims for unlawful blacklisting and wrongful termination in violation of Nevada public policy.<sup>9</sup> As our emphasis is on the effects of judicial deference to employment structures, we focus on

---

<sup>8</sup> Serlin Deposition 62:1-4.

<sup>9</sup> Serlin also made allegations under the analogous Nevada state statutes as to Serlin’s disability, religion, and age discrimination claims. The district court analyzed these claims “under the framework provided by the ADA, Title VII, and the ADEA as they operate under the same guiding principles.” 2014 WL 1573535, at \*2 n.1.

Title VII and ADEA claims. We do not discuss the ADA and FMLA claims in depth because the procedural and doctrinal strictures involved in those claims are somewhat more varied and idiosyncratic.

Dawson moved for summary judgment. In her opposition to the motion for summary judgment, Serlin's attorney submitted excerpts of Smith's deposition.<sup>10</sup> In his deposition, Smith admits that Serlin told him that the "Bagel Belt" comments upset her and that she "viewed it as a derogatory comment or an ethnic slur."<sup>11</sup> Smith also stated that Dawson School had an antiharassment policy and that the policy covered "derogatory comments and slurs."<sup>12</sup> Finally, Smith admitted that Serlin "explained that it was an ethnic slur towards Jews" and that, after Serlin's explanation, that he had "the impression that it was derogatory."<sup>13</sup>

Serlin's attorney also submitted three positive Annual Performance Reviews of Serlin by Dawson.<sup>14</sup> The 2008 review praised Serlin's performance. The 2008 evaluation, prepared by Smith, stated that he "enthusiastically recommend[ed] [Serlin] be offered a contract for the 2008-9 school year." The 2009 review, also prepared by Smith, contained only positive comments and remarked upon Serlin as "always professional in her dealings with students and parents." Finally, Smith's 2010 evaluation of Serlin contained no negative remarks on her professionalism or ability to interact and work with her colleagues and, by contrast, again noted that she "is

---

<sup>10</sup> Plaintiff's Opposition to Defendants' Motion for Summary Judgment, Exhibit 4, Deposition of Russell A. Smith ("Smith Deposition"), July 26, 2013, 12-cv-01431-JCM-NJK, ECF No. 48-23.

<sup>11</sup> *Id.* 38:16-20.

<sup>12</sup> *Id.* 38:24-39:2.

<sup>13</sup> *Id.* 40:1-12, 42:5-8

<sup>14</sup> Plaintiff's Opposition to Defendants' Motion for Summary Judgment, Exhibits 17-19, Annual Performance Reviews (2008, 2009, 2010), 12-cv-01431-JCM-NJK, ECF No. 48-4.

always professional in her dealings with students and parents.” Strikingly, Smith also wrote that Serlin and “Tonogi have developed bring real value to the students’ education, and the parents enjoy seeing their children excited to learn.”

In response to Serlin’s hostile work environment and retaliation claims, Dawson replied that it hired Vetter, and declined to renew Serlin’s contract, for two reasons. First, Dawson claimed that in 2011, it had considered a new requirement that all fifth-grade teachers be capable of teaching math and that Serlin was unwilling or unable to do so. Second, Dawson claimed that Serlin’s alleged inability to get along with Tognoni and another teacher was an additional factor in its decision.

Despite Smith’s admission that he had considered Tognoni’s behavior derogatory and a violation of the company’s antiharassment policy, the district court granted summary judgment to the employer. Notably, in its “Legal Standard” section, the district court failed to include a critical portion of the summary judgment standard. Nowhere in this section did the court indicate that, as the Supreme Court has required, “the inferences to be drawn from the underlying facts must be viewed in the light most favorable to the party opposing the motion.” *Matsushita*, 475 U.S. at 587-88 (citations and internal quotation marks and ellipses omitted) (emphasis added). That omission set the stage for the deferential analysis and cursory examination of the adequacy of the employer’s antidiscrimination structures that followed.

After making quick work of Serlin’s FMLA and ADA claims, the district court moved to Serlin’s Title VII claims. In three brief sentences regarding her Title VII religious discrimination claim,

the district court, without citing to any facts, concluded that Serlin “has provided no evidence that the named defendants discriminated against her because of her Jewish faith.” The court discounted Serlin’s version of events with respect to her meeting with Smith and concluded: “Candidly, the court is unconvinced that insinuating one likes bagels may constitute the type of harassment contemplated by Title VII and actionable under law.”

As to Serlin’s hostile work environment claim, the court found Serlin’s claim lacking because it concluded, again without reference any witness’s or party’s testimony, that Tognoni’s conduct was simply not severe or pervasive enough to create an objectively hostile or abusive work environment.” The court concluded: “While Tognoni’s comments implying that plaintiff enjoys bagels by virtue of her Jewish faith may have engendered offensive feelings, a reasonable person would not find that those comments created a hostile or abusive work environment.”

The court then also concluded that Serlin’s retaliation claim was without merit, explaining that Serlin had failed to establish a sufficient causal link between her complaint to Smith and failure to renew her contract. Again, the court did not cite to any facts in its analysis and failed to even begin to consider whether Dawson’s complaint procedure regarding written complaints was adequate or sensible. Interestingly, the court made no mention of Smith’s admissions regarding the existence of Dawson’s anti-harassment policy and the fact that Tognoni’s repeated “Bagel Belt” comments could easily have constituted the very “derogatory comments and slurs” that the policy forbade. Instead, the court accepted the validity and propriety of Dawson’s complaint procedures and, without discussion, assumed that it was valid on the way to granting Dawson’s summary judgment motion. The court privileged one diversity structure, Dawson’s complaint

procedure, over another, Dawson's anti-harassment policy. In doing so, the court made a substantive choice, one properly allocated to the jury, in choosing to find implicitly credible the diversity structure that bolstered Dawson's case for summary judgment.

Finally, in evaluating Serlin's ADEA claim, the district court overlooked significant testimonial evidence from Serlin and documentary evidence in the form of positive teaching evaluations written by Smith in accepting Dawson's claim that Serlin did not "get along" with the other fifth-grade teachers.

The district court appeared to overlook evidence of harassment, retaliation, and discrimination in granting summary judgment to the employer while paying attention to the employer's antiharassment policy and complaint procedure. The opinion makes no mention of the standard it used to review the facts and appears to have drawn inferences in a light favorable to the employer, in direct contrast to the Supreme Court's direction to do the opposite.

Unfortunately, nothing changed for Serlin on appeal. On July 28, 2016, the Ninth Circuit affirmed the district court's grant of summary judgment for Dawson. *Serlin v. Alexander Dawson School, LLC*, --- Fed. App'x ----, No. 14-15937, 2016 WL 4039713 (9th Cir. July 28, 2016). In a brief, unpublished, five-page memorandum disposition, the Ninth Circuit affirmed trial court's reasoning and disposition of the case.

The Ninth Circuit, like the district court, made no mention of the standard that requires courts to view the facts in the light most favorable to Serlin. This lack of mention of this key evidentiary



aspect of the summary judgment standard is strange when considering that, as the Ninth Circuit explicitly held as recently as 2009, in reviewing summary judgment grants: “[W]e are governed by the same principles as the district court: whether, with the evidence viewed in the light most favorable to the non-moving party, there are no genuine issues of material fact, so that the moving party is entitled to a judgment as a matter of law.” *San Diego Police Officers’ Ass’n v. San Diego City Emp. Ret. Sys.*, 568 F.3d 725, 733 (9th Cir. 2009).

As with the district court’s decision, the Ninth Circuit’s omission of a key evidentiary standard in its evaluation of the summary judgment motion led, quickly and inexorably, to the conclusion that the district court’s deference was proper and that its summary judgment decision was correct. The Ninth Circuit swiftly disposed of Serlin’s religious discrimination claim, concluding that even in the face of evidence that Dawson declined to renew other Jewish teachers’ contracts no triable issue of fact remained on this point. The Ninth Circuit further held that the alleged “Bagel Belt” comments failed to create a hostile work environment as the comments were, in the court’s opinion and without citation to any portion of the record, simply “not of a physically threatening or humiliating nature” and that they were therefore insufficiently severe, pervasive, or offensively objective to be actionable under Title VII.

In its decision, the Ninth Circuit deferred to several diversity structures, most implicitly but one explicitly. The court made no mention of the adequacy of Dawson’s complaint procedure, instead recasting and sanitizing Serlin’s written complaint as “the informal complaint she made regarding a coworker’s comments” and finding that she did not demonstrate that the complaint was a but-for cause of Dawson’s decision not to renew Serlin’s contract thus implicitly finding

that the structure must be adequate.<sup>15</sup> The court also made no mention of Dawson’s antiharassment policy and whether the policy would have categorized the “Bagel Belt” comments as harassment, thus implicitly deferring to the adequate application of that policy by Dawson through Smith’s inaction. Finally, the court explicitly found that Dawson’s hiring policies with respect to Serlin’s ADEA claim were adequate and proper. The court concluded: “[T]he decision makers’ statements that they desired creative and ‘dynamic’ teachers who have an ‘energized way of teaching’ and who will integrate technology into lessons are ‘at best weak circumstantial evidence of discriminatory animus’” based on age.”<sup>16</sup>

The Serlin case exemplifies courts’ tendencies to overlook the legal standard that requires deference to the nonmoving party and to defer to the mere presence of the employer’s diversity structures. While Serlin is a case that, on its facts, is less shocking than the City of Robertsdale case discussed below, it is a useful example of a more subtle way in which the interplay between the legal evidentiary standard on summary judgment and judicial deference can combine to the significant disadvantage of a plaintiff.

### **The Rise in Grants of Summary Judgment to Employers**

Sexual harassment was not expressively prohibited by the 1964 Civil Rights Act. In the early years of Title VII litigation, sexual harassment cases were brought as disparate treatment cases involving sex discrimination, but they were often dismissed, either because courts did not see sex discrimination as including sexual harassment or because courts were reluctant to hold

---

<sup>15</sup> 2016 WL 4039713, at \*2.

<sup>16</sup> 2016 WL 4039713, at \*2 (citation omitted).

employers liable for sexual harassment by individual supervisors or coworkers. In response to feminist activism, however, courts began in the 1970s and 1980s to hold that sexual harassment was a form of sex discrimination. The first sexual harassment case to reach the US Supreme Court was *Meritor Savings Bank v. Vinson* in 1986. Citing the EEOC guidelines, the Court established what is now known as hostile environment sexual harassment by holding that Title VII covers harassment that creates a hostile environment irrespective of whether there is a tangible economic loss.

The Court in *Meritor* also laid the foundation for employers to avoid liability by creating diversity structures, in particular antiharassment policies and grievance procedures. Until 1986 no Supreme Court cases had explicitly stated that organizational structures might protect employers from liability in the context of a Title VII case. In most areas of civil rights law, that is still the case. But when the Supreme Court defined hostile work environment sexual harassment in *Meritor*, it, for the first time, suggested that an effective antiharassment policy and a grievance procedure might protect an employer from liability when a supervisor harasses an employee. The *Meritor* decision led management consultants and human resource professionals to encourage employers to create these procedures and also dramatically increased the rate at which employers pointed to their diversity structures as evidence that they had taken reasonable measures to avoid sexual, and also racial, harassment (Edelman 2016). Over the next twelve years, judges became increasingly likely to defer to these structures (Edelman et al. 2011; Edelman 2016).

Then in 1998, *Faragher v. City of Boca Raton* and *Burlington Industries, Inc. v. Ellerth*, the Supreme Court formalized the defense that Meritor had hinted at which created the affirmative defense that the Meritor Court had hinted at. SHRM and other organizations representing employers' interests had submitted amicus briefs contending that the presence of antiharassment policies and complaint procedures for employees should always protect employers from liability. The Equal Employment Opportunity Commission (EEOC) support of this contention, albeit only in the context of hostile environment claims where there was no tangible economic loss and only where these procedures were shown to be effective. Other organizations representing employees' interests objected, contending that such a policy would allow employers to escape liability simply by creating diversity structures (Edelman 2016).

The Supreme Court followed the advice of the EEOC, establishing an affirmative defense in hostile work environment cases. The two-pronged affirmative defense, set forth in both *Faragher* and *Ellerth*, requires that employers prove:

- (a) that the employer exercised reasonable care to prevent and correct promptly any sexually harassing behavior, and (b) that the plaintiff employee unreasonably failed to take advantage of any preventive or corrective opportunities provided by the employer or to avoid harm otherwise (*Ellerth*, 524 U.S. at 765; *Faragher*, 524 U.S. at 807).

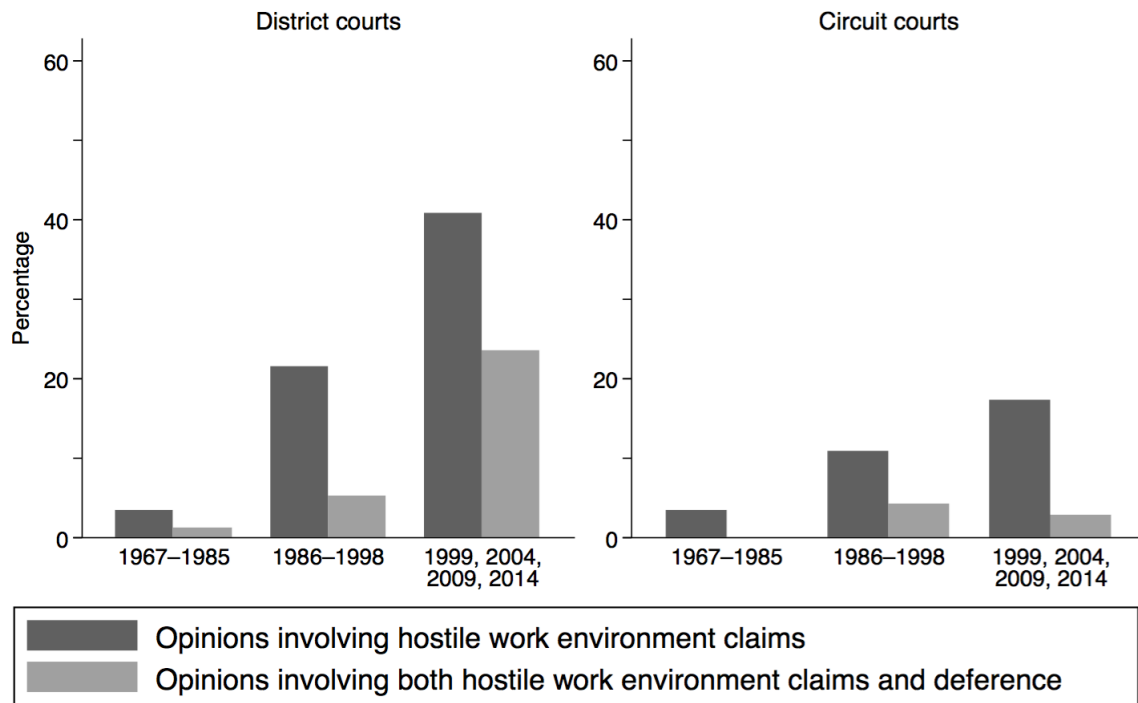
The opinions explicitly suggest, moreover, that antiharassment policies and complaint procedures would, in most cases, allow the employer to escape liability.

While proof that an employer had promulgated an anti-harassment policy with complaint procedure is not necessary in every instance as a matter of law, the need for a stated policy suitable to the employment circumstances may appropriately be addressed in any case when litigating the first element of the defense. And while proof that an employee failed to fulfill the corresponding obligation of reasonable care to avoid harm is not limited to showing any unreasonable failure to use any complaint procedure provided by the employer, a demonstration of such failure will normally suffice to satisfy the employer's burden under the second element of the defense. (Ellerth, 524 U.S. at 744; Faragher, 524 U.S. at 778).

### **Judicial Deference to Diversity Structures in Hostile Work Environment Harassment Cases**

Following the Faragher and Ellerth decisions, judicial deference to the presence of diversity structures increased dramatically in hostile work environment cases, especially in the district courts. Figure 5 shows both the increase in cases involving hostile work environment claims and the dramatic increase in judicial deference, first after the Meritor decision in 1986 and then after the Faragher and Ellerth decisions in 1998. In the district courts, during the twelve years prior to Faragher and Ellerth, deference occurred in only about 24 percent of opinions involving hostile work environment. After Faragher and Ellerth, deference occurred in about 58 percent of those opinions. In the latter period, nearly all hostile work environment harassment cases are decided via summary judgment.

**Figure 5**  
**Percentage of opinions involving hostile work environment claims and opinions involving both hostile work environment claims and deference**



Judicial deference in hostile work environment cases is not entirely due to the Supreme Court's Faragher and Ellerth decisions. As was shown in figure 2, the trend toward judicial deference to diversity structures began in the lower courts decades prior to these decisions. Edelman et al. (2011) show, moreover that courts found diversity structures more likely to be relevant in hostile work environment cases than in most other types of cases long before Faragher and Ellerth. But the Supreme Court's affirmation of the affirmative defense and specific mention of antiharassment policies and complaint procedures certainly spurred lower courts to defer to diversity structures.

Importantly, however, nothing in the Supreme Court decision should have been taken as a directive for lower courts to defer to diversity structures without considering the efficacy of these structures. Evidence that the policy is effective or that employers ignore their own policies would seem to violate the Supreme Court's presumption that a policy indicates that an employer took reasonable care to avoid harm. Similarly, efforts by an employer to preclude employees from using complaint procedures by threatening retaliation or evidence that complaint procedures are unfairly biased in favor of the employer would suggest that an employee's failure to make use of the complaint procedure was not necessarily unreasonable. Yet a reading of these cases indicates that judges regularly fail to evaluate the efficacy of these structures. Instead, judges defer to the mere existence of a harassment policy without any analysis of whether these policies are effectively implemented within organizations.

Consider, for example, *Howard v. City of Robertsdale*, Civil Action No. 03-0770-BH-C, 2004 WL 5551812 (S.D. Ala. Dec. 9, 2004), a case in which an Alabama federal district court considered whether to grant summary judgment for the City of Robertsdale on the plaintiff's Title VII hostile work environment sexual harassment claims and equal protection and due process Fourteenth Amendment claims pursuant to 42 U.S.C. § 1983.<sup>17</sup>

Elizabeth J. Howard was the plaintiff in *City of Robertsdale*. Howard began her career as the secretary and administrative assistant to the Chief of Police, Alan Lassiter, in May 1999. In this case, the defendant, City of Robertsdale actually conceded that shortly after Howard was hired, Lassiter engaged in frequent instances of sexual harassment of Howard. The opinion reveals

---

<sup>17</sup> Unless otherwise noted, the facts of the case are taken directly from the district court's order at 2004 WL 5551812, at \*1-\*2.

that: “He grabbed her breasts and pinched her buttocks on a weekly basis;” “On multiple occasions he poked her between the legs with the antenna of his police radio;” “He grabbed her and attempted to kiss her multiple times per week;” “He would pull open her blouse and look at her breasts multiple times per week;” and “He made countless sexual comments towards Howard regarding her body and his desire to have sex with her.”<sup>18</sup>

The City had an antiharassment policy and a complaint procedure in place, which stated:

All employees are responsible for helping to assure that we avoid harassment. If you feel you have experienced or witnessed harassment, you are to notify immediately (preferably within 24 hours) your immediate supervisor, personnel department, and/or the Mayor.<sup>19</sup>

Howard did not initially file a complaint, however, contending that she was “scared to death of [Lassiter]” and that he had prohibited his employees from going over his head to the Mayor.<sup>20</sup>

Beginning in December 20, 1999, Howard sought promotion to the position of Chief Dispatcher. On April 6, 2000, another woman, Katrina Griffin, was appointed to that position. Howard claimed that Lassiter told her that she was more qualified than Griffin but that “he would not promote her because he wished to keep her as his administrative assistant.”<sup>21</sup> In 2002, Lassiter

---

<sup>18</sup> 2004 WL 5551812, at \*1.

<sup>19</sup> *Id.* at \*2.

<sup>20</sup> *Id.* at \*1.

<sup>21</sup> *Id.*



suspended Howard without pay for half a day after she complained about his harassing behavior toward another employee.

In April 2002, Howard did approach Chief Dispatcher Katrina Griffin to complain about Lassiter's behavior but Griffin never reported the complaint. On May 27, 2002, Howard and her husband finally complained about the sexual harassment to the town's mayor, Charles H. Murphy. After Murphy received Howard's complaint, he contacted the City Attorney to investigate the situation. The City Attorney conducted an investigation and reported the results to Murphy on August 23, 2003. A disciplinary process was then initiated and Lassiter was terminated as Chief of Police on October 24, 2002 and was subsequently terminated by the City Council on November 4, 2002 for "lewd and immoral conduct" and "sexual harassment of a female subordinate."<sup>22</sup> As of December 9, 2004, Howard remained employed by the City as an administrative assistant for the City's police department.

In October, 2002, Howard filed a discrimination complaint with the EEOC and obtained a right to sue letter. She filed a complaint against the City of Robertsdale, alleging a hostile work environment in violation of Title VII and Section 1983. Robertsdale filed a motion for summary judgment. The district court articulated the proper evidentiary standard on summary judgment and recognized that it must "resolve all reasonable doubts about the facts in favor of the non-movant, and draw all justifiable inferences" in the non-moving party's favor. Nonetheless, the court ultimately deferred to the city's reporting policy and awarded summary judgment on all claims to the City.

---

<sup>22</sup> *Id.* at \*2.

The district court's decision turned on Howard's failure to follow the strictures of the City's reporting policy. The court viewed the facts as analogous to the Eleventh Circuit's decision in *Madray v. Publix Supermarkets, Inc.*, 208 F.3d 1290 (11th Cir.2000), in which the Eleventh Circuit held: "[O]nce an employer has promulgated an effective anti-harassment policy and disseminated that policy and associated procedures to its employees, then it is incumbent on the employees to utilize the procedural mechanisms established by the company specifically to address problems and grievances." The district court discounted Howard's complaint to Griffin by pointing to another holding in *Madray* that "informal complaints to individuals not designated to receive or process sexual harassment complaints" were insufficient to put the defendant on notice of sexual harassment.<sup>23</sup> And even though the City procedure's requirement that Howard complain to her direct supervisor – the perpetrator of the harassment – violated EEOC guidelines, the district court, without further examination, implicitly accepted the propriety of the city's reporting policy and its interpretation of its own definition of "immediate supervisor."

Further, the district court found that there were no triable issues of fact with respect to direct liability and actual knowledge on the part of anyone who could be construed to be Howard's immediate supervisor even though Howard had asserted that a number of supervisory personnel, including the Chief Dispatcher, a police lieutenant and sergeant, all had actual knowledge of Lassiter's verbal harassment. As to whether the city had constructive knowledge of Lassiter's repeated and severe sexual harassment, the court again deferred to the reasonableness and adequacy of the city's reporting policy.

---

<sup>23</sup> *Id.* at \*4.

Howard's hostile work environment claims likewise suffered defeat-by-deference. Using the Faragher-Ellerth affirmative defense, the city pointed to its antiharassment policy and grievance procedure as evidence that it had taken reasonable steps to prevent and correct sexual harassment. Although Howard presented substantial evidence that the antiharassment policy was ineffective and that the complaint procedure would have required her to complain directly to her perpetrator, and even though Howard presented testimony from Sergeant Middleton's deposition that "Police Chief Lassitter [sic] communicated a conflicting policy within his department that "he prohibited the employees in his department from going over his head and speaking with the Mayor,"<sup>24</sup> the court accepted the affirmative defense, pointing to the fact that Howard had "allowed the harassment to continue for approximately two and a half years without reporting it to any of the parties designated to handle such complaints under the policy."<sup>25</sup> In failing to consider evidence pointing to clear inadequacies in both the antiharassment policy and the complaint procedure, the district court appears to have drawn an inference in favor of the moving party (the employer).

Finally, the court granted summary judgment for the City on Howard's § 1983 claim in which it deferred to the formal language of the City's anti-harassment policy. The court simply found, as to formal policy: "The City of Robertsdale certainly does not have an official policy condoning or encouraging these inappropriate acts by Mr. Lassiter. Rather, the City has a comprehensive

---

<sup>24</sup> Initial Brief of Appellant, 168 Fed. App'x 883 (11th Cir. 2006) (No. 05-10023), 2005 WL 4843213.

<sup>25</sup> 2004 WL 5551812, at \*8.

anti-sexual harassment policy and reporting guidelines.”<sup>26</sup> It then concluded, without analysis, that “[t]here is also no evidence of an unofficial custom or practice permitting such actions by the City.”<sup>27</sup>

As in *Serlin*, Howard, as plaintiff and appellant, fared no better on appeal. In an unpublished decision, the Eleventh Circuit found Howard’s delay in reporting Lassiter’s behavior unreasonable and found that the City properly supported its Ellerth-Faragher defense.<sup>28</sup> The Eleventh Circuit, in failing to examine the adequacy of the City’s policy and instead using it to doom her case, placed Howard in a fatal Catch-22 – it notes that Howard “contends she never complained to the Mayor because Lassiter prohibited his employees from going to the Mayor” but that her “own actions . . . contradict this assertion” because she eventually did complain to Murphy.<sup>29</sup> Howard’s initial fear and hesitation of violating a policy laid down by her direct supervisor and department head were rendered per se unreasonable because she chose to eventually complain to another individual in violation of that policy.

In affirming the district court’s grant of summary judgment as to direct liability, the Eleventh Circuit first dismissed Howard’s direct liability complaint by finding that “Title VII is not a general civility code” and that Lassiter’s “frequent remarks about female employees’ bodies and sex lives” simply “do not rise to the level of discrimination under Title VII and cannot serve as the basis for constructive knowledge.” It also affirmed the grant of summary judgment as to

---

<sup>26</sup> *Id.* at \*9.

<sup>27</sup> *Id.*

<sup>28</sup> *Howard v. City of Robertsdale*, 168 Fed. App’x 883 (11th Cir. 2006).

<sup>29</sup> *Id.* at 887.

Howard’s section 1983 claims as it deferred to the adequacy of the City’s policy in finding that “[t]he record instead reveals a comprehensive sexual harassment policy, of which all employees were aware . . . .”<sup>30</sup>

In sum, City of Robertson is an example of deference by a trial and appellate court to the formal language of an antiharassment policy and reporting structure without consideration, as otherwise mandated by Supreme Court precedent in summary judgment cases, of the facts as presented by the plaintiff that tend to show the inadequacy and ineffectiveness of the reporting policy.

#### Symbolic Metrics, Judicial Politics, and Case Outcome

One might reasonably ask whether the rise in judicial deference to symbolic metrics of diversity might be attributable to conservative trends in the judiciary over the past half century. It is certainly true that the federal judiciary has become more conservative over time (Martin and Quinn 2002) and that conservative judges are generally more likely to rule in favor of employers (Krieger et al. 2015; Edelman 2016). It is not the case, however, that conservative judges are more likely than liberal judges to defer to symbolic metrics of diversity. In fact, Krieger et al. (2015) report that there is little difference in the likelihood of deference based on judicial politics in the district courts and that in the circuit courts, liberal judges are actually more likely than conservative judges to defer to diversity structures, perhaps because liberal judges are more impressed by the trappings of due process and rational governance.

---

<sup>30</sup> *Id.* at 889-90.

Judicial deference to diversity structures alone, moreover, does not guarantee that an employer will prevail because many other factors come into play in employment discrimination cases. Nonetheless, all else equal, judicial deference does make it much more likely that employers will win employment discrimination cases. When judges adequately scrutinize diversity structures, the outcome of cases depends largely on whether the judge rules that the structures are adequate (in which case the employer generally wins) or inadequate (in which case the employee generally wins). When judges defer to diversity structures without adequate scrutiny, however, employers win at nearly the same rate as is the case when judges ruled that the structures are adequate (Krieger et al. 2015).

### **The Logic of Judicial Deference**

We see judicial deference to diversity structures as part of a broader judicial reticence to review employers' personnel decisions. This logic is particularly evident in a phrase that we saw repeatedly in employment discrimination cases. With slight variations, the phrase suggested that courts should not act as "super-personnel departments." The phrase was typically used in a sentence like: "This court has repeatedly stated that it is not a super-personnel department that second-guesses employer policies that are facially legitimate." *Widmar v. Sun Chemical Corp.*, 772 F.3d 457, 464 (7th Cir. 2014).

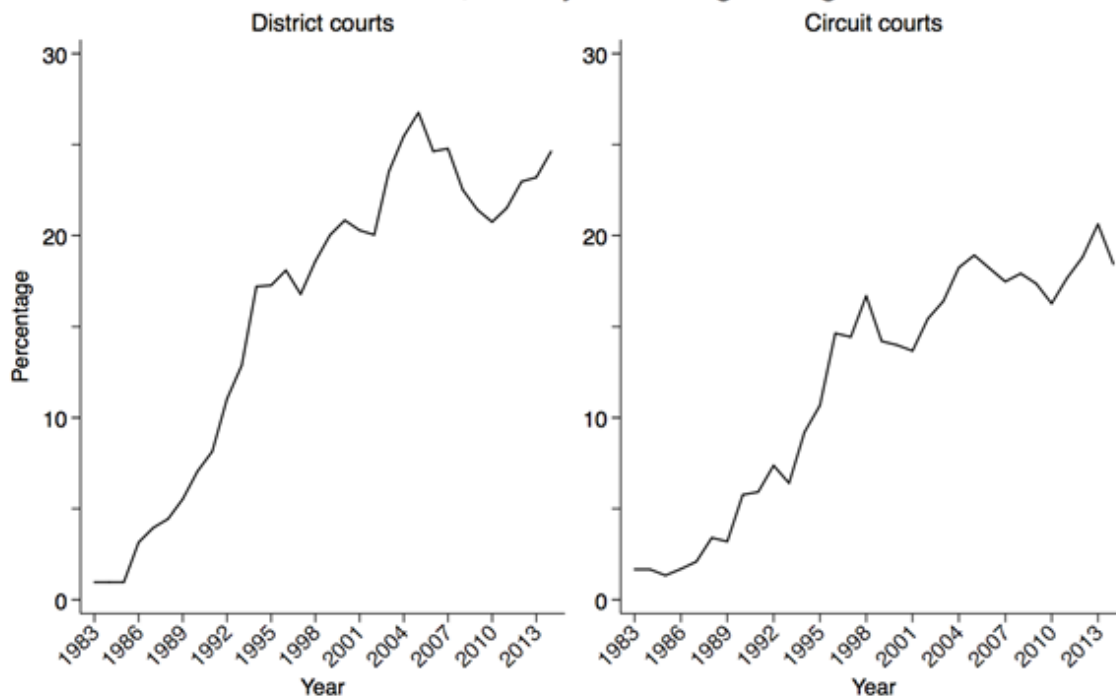
The increasing frequency with which the term 'super-personnel department' is used illustrates the resonance of the idea that courts prefer to stay out of business decisions. Figure 6 is based on a Westlaw search for the term super-personnel department in Title VII opinions involving grievance procedures, antiharassment policies, or diversity policies, shows the percentage of

opinions that use the term over time. The figure suggests that judicial reluctance to second-guess employers' business and personnel decisions was becoming more common at the same time that judicial deference to diversity structures was rising.<sup>31</sup> As of 2014, the term had been used in 498 circuit court opinions and in 2,855 district court opinions involving grievance procedures, antiharassment policies, or diversity policies. This term actually appears in only a small proportion of all EEO opinions in which judges defer to diversity structures, but it illustrates the general reticence of judges to second guess employers' personnel actions and helps to explain why judges so often give little weight to evidence that suggests that employers' diversity policies may be merely symbolic rather than both symbolic and substantive.

---

<sup>31</sup>We searched the Westlaw CTA and DCT databases, separately, for “super-personnel department” “super personnel department” ((super /3 person!) /3 department) within cases located through a Westlaw search for (“title vii”) & ((grievance appeals complaint /3 proc!) (“open door policy” ombud!)) ((anti! /3 polic!) (harass! /3 polic!) (divers! /3 polic!)). The percentage of cases in which the term “super-personnel department” appeared was calculated by dividing the number of such cases by the total number of cases found through the Westlaw search for Title VII cases involving grievance procedures, antiharassment policies, or diversity policies. Figure 8.6 starts in 1983, when the first instance of that term appears.

**Figure 6**  
**Percentage of opinions with the phrase *super-personnel department*, 1983 to 2014, three-year moving average**



## Conclusion

Diversity structures have become, to a great extent, symbolic metrics of diversity. If diversity structures were uniformly effective, this trend would not be problematic. But when judges defer to diversity structures without adequate attention to the adequacy of these structures, they undermine rather than protect civil rights in the workplace. We have presented quantitative evidence showing that judicial deference to diversity structures has become the norm since 2000, especially in cases involving motions for summary judgment and in cases alleging hostile work environment harassment. We have also presented two case examples, *Serlin* and *City of Robertson*, which illustrate how deference and the subtle use of different standards of proof can work together to the significant disadvantage of employment discrimination plaintiffs. The



graphs, based on quantitative analyses of a representative sample of federal trial and appellate cases from 1964 through 2014, suggest that the deference illustrated in the *Serlin* and *City of Robertson* cases is becoming more typical in civil rights adjudication generally.

Many diversity structures in today's corporations do promote greater equality and inclusion for women, people of color, and other protected groups. But it is critical that lawyers, judges, and policymakers recognize that the diversity structures that have become a commonplace feature of the American corporate workplace are not always evidence of nondiscrimination.

In hostile work environment cases, where employers invoke the *Faragher-Ellerth* affirmative defense, judges should recognize that Justice Kennedy's presumption in *Ellerth* and *Faragher* that anti-harassment policies and complaint procedures provide evidence that the employer took reasonable care to prevent and correct harassing behavior does not fit the reality of many workplaces. Social science evidence shows that everyday practices may discriminate against employees even where policies prohibiting such discrimination exist and that informal threats or organizational culture may lead employees reasonably to fear using employers' complaint procedures (Edelman 2016). Thus judges should take very seriously evidence suggesting that diversity policies are ineffective or inadequate in light of workplace culture.

Similarly, judges should recognize in summary judgment cases that any presumption that the presence of diversity structures automatically indicates an employers' good faith effort to comply with Title VII or other civil rights laws, especially in light of evidence of the inadequacies of those structures, may constitute drawing an inference in a light favorable to the moving party, in

contravention of the rule articulated by the Supreme Court in *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574, 587-88 (1986), which specifies that judges should draw all inferences in the light most favorable to the nonmoving party.

By illustrating the potential of diversity structures to serve as symbolic metrics of diversity irrespective of their substantive impact, we hope that this article will encourage plaintiffs' attorneys, judges, and policymakers to give greater attention to the potential inadequacies of diversity structures and to avoid deferring to symbolic metrics when they fail to protect legal rights.

## References

- Albiston, Catherine R. "The Rule of Law and the Litigation Process: The Paradox of Losing by Winning." *Law & Society Review* 33, no. 4 (1999): 869-910.
- Berrey, Ellen. *The Enigma of Diversity: The Language of Race and the Limits of Racial Justice*. Chicago: University of Chicago Press, 2015.
- Dover, Tessa L., Brenda Major, and Cheryl R. Kaiser. "Diversity Initiatives, Status, and System-Justifying Beliefs: When and How Diversity Efforts De-Legitimize Discrimination Claims." *Group Processes & Intergroup Relations* 17, no. 4 (2014): 485-93.
- Edelman, Lauren B. "Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law." *American Journal of Sociology* 97, no. 6 (1992): 1531-76.
- Edelman, Lauren B. *Working Law: Courts, Corporations, and Symbolic Civil Rights*. Chicago: University of Chicago Press, 2016.
- Edelman, Lauren B., Christopher Uggen, and Howard S. Erlanger. "The Endogeneity of Legal Regulation: Grievance Procedures as Rational Myth." *American Journal of Sociology* 105, no. 2 (1999): 406-54.
- Edelman, Lauren B., Howard S. Erlanger, and John Lande. "Internal Dispute Resolution: The Transformation of Civil Rights in the Workplace." *Law & Society Review* 27, no. 3 (1993): 497-534.
- Edelman, Lauren B., Linda H. Krieger, Scott R. Eliason, Catherine R. Albiston, and Virginia Mellema. "When Organizations Rule: Judicial Deference to Institutionalized Employment Structures." *American Journal of Sociology* 117, no. 3 (2011): 888-954.

- Edelman, Lauren B., Sally Riggs Fuller, and Iona Mara-Drita. "Diversity Rhetoric and the Managerialization of Law." *American Journal of Sociology* 106, no. 6 (2001): 1589-641.
- Edelman, Lauren B., and Stephen Petterson. "Symbols and Substance in Organizational Response to Civil Rights Law." *Research in Social Stratification and Mobility* 17 (1999): 107–35.
- French, Christopher C. "Debunking the Myth That Insurance Coverage Is Not Available or Allowed for Intentional Injuries or Damage." *Hastings Business Law Journal* 8, no. 1 (2012): 65-102.
- Gertner, Nancy. "Losers' Rules." *The Yale Law Journal Online* 122 (2012): 109-24.
- Guthrie, Chris, Jeffrey J. Rachlinski, and Andrew J. Wistrich. "Blinking on the Bench: How Judges Decide Cases." *Cornell Law Review* 93, no. 1 (2007): 1-44.
- Guthrie, Chris, Jeffrey J. Rachlinski, and Andrew J. Wistrich. "Judging by Heuristic-Cognitive Illusions in Judicial Decision Making." *Judicature* 86, no. 1 (2002): 44-50.
- Kaiser, Cheryl R., and Brenda Major. "A Social Psychological Perspective on Perceiving and Reporting Discrimination." *Law & Social Inquiry* 31, no. 4 (2006): 801–30.
- Kaiser, Cheryl R., Brenda Major, Ines Jurcevic, Tessa L. Dover, Laura M. Brady, and Jenessa R. Shapiro. "Presumed Fair: Ironic Effects of Organizational Diversity Structures." *Journal of Personality and Social Psychology* 104, no. 3 (2013): 504-19.
- Kaley, Alexandra, Frank Dobbin, and Erin Kelly. "Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies." *American Sociological Review* 71, no. 4 (2006): 589–617.
- Krieger, Linda Hamilton. "The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity." *Stanford Law Review* 47, no. 6

(1995): 1161-248.

Krieger, Linda Hamilton, Rachel Kahn Best, and Lauren B. Edelman. "When 'Best Practices' Win, Employees Lose: Symbolic Compliance and Judicial Inference in Federal Equal Employment Opportunity Cases." *Law & Social Inquiry* 40, no. 4 (2015).

Marshall, Anna-Maria. "Idle Rights: Employees' Rights Consciousness and the Construction of Sexual Harassment Policies." *Law & Society Review* 39, no. 1 (2005): 83-123.

Marshall, Anna-Maria. "Injustice Frames, Legality, and the Everyday Construction of Sexual Harassment." *Law & Social Inquiry* 28, no.3 (2003): 659-89.

Martin, Andrew D., and Kevin M. Quinn. "Dynamic Ideal Point Estimation Via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999." *Political Analysis* 10, no. 2 (2002): 134-53.

Mootz III, Francis J. "Insurance Coverage of Employment Discrimination Claims." *University of Miami Law Review* 52, no. 1 (1997): 1-78.

Quinn, Beth A. "The Paradox of Complaining: Law, Humor, and Harassment in the Everyday Work World." *Law & Social Inquiry* 25, no.4 (2000): 1151-85.

Rachlinski, Jeffrey J. "Processing Pleadings and the Psychology of Prejudgments." *DePaul Law Review* 60, no. 2 (2011): 413-30.

Reskin, Barbara F., and Debra Branch McBrier. "Why Not Ascription? Organizations' Employment of Male and Female Managers." *American Sociological Review* 65, no. 2 (2000): 210-33.

Roma, Elizabeth A. "Mandatory Arbitration Clauses in Employment Contracts and the Need for Meaningful Judicial Review." *American University Journal of Gender, Social Policy & Law* 12, no. 3 (2004): 519-44.

Stone, Katherine Van Wezel. "Mandatory Arbitration of Individual Employment Rights: The Yellow Dog Contract of the 1990s." *Denver University Law Review* 73, no. 4 (1996): 1017-50.

Talesh, Shauhin A. "Legal Intermediaries: How Insurance Companies Construct the Meaning of Compliance with Antidiscrimination Laws." *Law & Policy* 37, no. 2 (2015): 209–39.

Weick, Karl E. "Educational Organizations as Loosely Coupled Systems." *Administrative Science Quarterly* 21, no. 1 (1976): 1-19.

### **Cases Cited**

*Burlington Industries v. Ellerth*, 524 U.S. 742 (1998)

*Faragher v. City of Boca Raton*, 524 U.S. 775 (1998)

*Howard v. City of Robertsdale*, Civil Action No. 03–0770–BH–C, 2004 WL 5551812 (S.D. Ala. Dec. 9, 2004)

*Howard v. City of Robertsdale*, 168 Fed. App'x 883 (11th Cir. 2006)

*Madray v. Publix Supermarkets, Inc.*, 208 F.3d 1290 (11th Cir. 2000)

*Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574, 587-88 (1986)

*Meritor Savings Bank v. Vinson*, 477 U.S. 57 (1986)

*San Diego Police Officers' Ass'n v. San Diego City Emp. Ret. Sys.*, 568 F.3d 725, 733 (9th Cir. 2009).

*Serlin v. Alexander Dawson School*, No. 2:12–CV–1431 JCM (GWF), 2014 WL 1573535 (D. Nev. Apr. 17, 2014),

*Serlin v. Alexander Dawson School, LLC*, --- Fed. App'x ----, No. 14–15937, 2016 WL 4039713 (9th Cir. July 28, 2016)

Widmar v. Sun Chemical Corp., 772 F.3d 457 (7th Cir. 2014).

**Statutes, Rules, and Regulations Cited**

Age Discrimination in Employment Act of 1967, 29 U.S.C. § 621, et seq.

Americans with Disabilities Act (ADA), 42 U.S.C. § 12111 et seq.

Civil Rights Act of 1964, Title VII, 42 U.S.C. § 2000e et seq.

Civil Rights Act of 1991, 42 U.S.C. § 1981 et seq.

Family and Medical Leave Act (FMLA), 28 U.S.C. § 2601 et seq.

### **Chapter 3: Litigation, Diversity, and Metrics / *Richard Tonowski, Chief Psychologist, EEOC***

Equal employment opportunity (EEO) has reached a watershed moment. Blatant discrimination based on race, color, sex, national origin, religion, age, and disability have been unlawful for between a quarter and a half century. Such practice is even associated with going out of business (Pager, 2016). This is not to say that the war has been won<sup>32</sup>. But it does imply that the nature of the battles is changing. The problems are more nuanced: not just sex discrimination, but sexual orientation and gender identity; not just segregation, but implicit bias; not just employment, but equal pay and access to better jobs. Since the enforcement of Title VII overt discrimination has decreased, although perhaps to be replaced by more ambiguous manifestations of prejudice, sometimes unconscious as might occur with stereotyping people of a given demographic group. Perhaps the focus now should be with structural change in organizations to eliminate root causes of unlawful discrimination, intergroup hostility, and negative stereotyping. This is a matter of how people are regarded, more than just the number and demography of those who get on the payroll.

Diversity programs and litigation have become the mechanisms for driving EEO. If employers did not adopt the first voluntarily, there was the second to vindicate the rights of employees and employment applicants. Do they drive EEO now?

---

<sup>32</sup> For example, despite the changing scene, resegregation at industry level (Stainback & Tomaskovic-Devey, 2012) is a traditional-type issue of current concern. And some organizations and professions may still have a notable lack of demographic diversity. In one interesting twist, Rosen (2017) described how major corporations are pressuring law firms with which they do business to become more demographically diverse, at least for the teams that handle the business of the client corporations. The article quotes the head of an advocacy association as naming law as the least diverse of white collar professions.



This paper discusses the following, through an examination of the metrics often used for litigation and diversity:

- EEO lawsuits are in decline; while sometimes necessary, the sufficiency of litigation in promoting EEO in organizations is questionable.
- Diversity programs have always been a mixed bag of effectiveness, often implemented with little planning than a means of affirmative action.
- Current concerns for workplace inclusion are not well addressed by diversity programs or litigation.
- Still, litigation and diversity programs have a role in addressing those concerns, when used in conjunction with social science and human resources (HR) management principles.
- This implies a need to focus better on what objectives are to be sought and how progress toward those objectives is to be measured.

## **Perspectives**

The author is an industrial-organizational (I-O) psychologist. That branch of psychology deals with, among other things, the nature of work and the competencies necessary for its performance, and methods for the assessment of those competencies; workforce morale and related concepts, and their measurement; employee training and development; effective leadership and collaboration; and the mitigation of counterproductive workplace behavior. In addition, the author has been a human resources manager, and was a diversity manager for a brief time.

Much of this activity has been employer-side, dealing with personnel selection procedures.

More recently, I have been employed by the U.S. Equal Employment Opportunity Commission (EEOC)<sup>33</sup> in the role of an in-house expert for evaluation of personnel selection procedures and statistical analysis.

The author has been engaged in several social science studies in the field, including in “meta-analyses.” Meta-analysis is a generic term for statistical summaries of research across individual studies. The purpose is to aid in forming an overall conclusion of the research and to identify the types of situations that might amplify or diminish the strength of research findings.

### **Definitions and Metrics**

Several related terms are used here; their distinction follows from the Society for Human Resource Management (SHRM) documents, principally SHRM (2016) with some modification to avoid confusion between programs and program outcomes:

Equal Employment Opportunity (EEO) means freedom from discrimination on the basis of sex, color, religion, national origin, disability and age. This connotes compliance with civil rights law and regulation, usually enforceable through litigation. SHRM (2015) identifies “managing EEO” with legal compliance. But litigation per se is a record of disputes regarding EEO, not EEO itself. The measures of “freedom from discrimination” would seem to be those associated

---

<sup>33</sup> The views expressed here are the author’s and do not necessarily represent the views of the EEOC or any other government agency. Nothing in this article should be construed as legal advice.

with diversity efforts: numerical representation as previously-excluded groups increase their representation for desirable employment, and perceptions of being included rather than excluded.

Affirmative action defines an employer's standard for proactively recruiting, hiring and promoting women, minorities, disabled individuals and veterans. It is deemed a moral and social obligation to amend historical wrongs and eliminate the present effects of past discrimination. Affirmative action plans include numerical measures ("headcount") with the intent of increasing the representation of minorities. The usual metric is degree of goal attainment.

Diversity initiatives are devised to increase the acceptance of minorities by embracing cultural differences within the workplace. Diversity initiatives are twofold: valuing diversity and managing diversity. The value of diversity is achieved through awareness, education and positive recognition of cultural differences within the workplace. The management of diversity expounds upon this experience, and establishes the business case for diversity that is closely aligned with an employer's organizational goals. Note that the "value" of diversity does not give a definite metric. The "business case" also is vague; it seems to imply connecting diversity concepts with attainment of organizational goals. Presumably these goals have their own metrics that are not diversity metrics. This is discussed below. One ongoing concern, discussed next, is that the management of diversity amounts to managing headcount

Inclusion is the New Diversity (Bates, 2013). As a matter of general practice, "old" diversity is equated with the numerical representation of race, ethnic, and sex groups primarily to mitigate

the risk of EEO litigation. As such, it is akin to affirmative action<sup>34</sup>. Inclusion focuses on the “acceptance” aspect of diversity and goes beyond it, and its application to everyone in the workforce. Its goal is to have all members of the workforce committed to the work of the organization, generally by promoting communication, mutual respect, and participation; de-emphasis on hierarchy; and identification and removal of barriers to participation. This includes an interest in removing overt and unconscious discrimination, and may incorporate special emphasis groups along demographic lines. In summary, it deals with the perception and reality of EEO within an organization. Current metrics rely on measuring the sentiments of workforce members with surveys. Action planning generally follows the survey results, and initiatives that spring from that will have their own metrics.

It should be noted that there are no “right” metrics in the abstract. Whether a metric is right depends upon its intended use. It follows that a metric is wrong to the extent that it is unrelated to, or an inadequate proxy for, what was intended to be measured.

### **EEO Enforcement Metrics**

A half-century after the passage of the Civil Rights Act of 1964 and the establishment of the EEOC, employment discrimination persists. Roughly 90,000 new EEO charges filed each year<sup>35</sup>

---

<sup>34</sup> A criticism is that it is a watered-down version of affirmative action that, by making it applicable to everyone, either disguises the “unfair” benefiting of certain demographic groups or else dilutes the effectiveness of efforts to help historically under-represented groups.

<sup>35</sup> EEOC statistics are on the federal fiscal year (FY). The FY starts a quarter earlier than the calendar year; thus, FY 2017 started on October 1, 2016. EEOC statistics were retrieved from [www.eeoc.gov/statistics](http://www.eeoc.gov/statistics). Federal district and appellate court data (starting in 2000) were retrieved from Lexis CourtLink. Civilian labor force statistics were retrieved from the Bureau of Labor Statistics at <https://data.bls.gov/pdq/SurveyOutputServlet>.

testify to the extent to which people think they have been victimized; this number does not include charges filed exclusively with state and local civil rights agencies, nor those which are not filed out of fear or hopelessness, or lack of knowledge regarding legal rights.

But the picture of charges and subsequent litigation is more nuanced than charge filings alone. The following charts tell the story of those charges from 1997 to 2016<sup>36</sup>. Figure 1 shows charge intake at around 80,000 for each of the earlier years; a climb with the Great Recession cresting in 2011 at just under 100,000; and a slow decline thereafter. The figure also shows the percentage of charges determined<sup>37</sup> to have probable cause or not. In 2016 67% of the charges went “no cause” for lack of merit or of jurisdiction. Higher percentages accompanied the increase in charge volume in more recent years. The “cause” rate high point for the period was just under 10% in FY 2001, with a gradual decline thereafter to 3.2% in FY 2016.<sup>38</sup>

The “no cause” rate may be indicative of a difference in what people perceive to be discriminatory and what the law recognizes as such. Perceptions are important, whether measured by charge or opinion surveys, but they are a different type of data than legally actionable instances. If there is both over-reporting of claims that are not actionable and under-reporting of unlawful activity, then arriving at a “real” count of instances is difficult<sup>39</sup>. Add to

---

<sup>36</sup> See Nielson, Nelson, Lancaster, and Pedriana (2008) for a detailed analysis of 1987-2003 data. EEOC (2011) also has a summary of earlier data compiled from EEO-1 reports.

<sup>37</sup> Charge determination indicates reasonable inference of discrimination, but it is not full adjudication. Trials proceed de novo.

<sup>38</sup> Hartstein (2016) estimated that 40% of systemic charges result in a “cause” finding; this information is not on the EEOC website.

<sup>39</sup> See Nielson and Nelson (2005) for a discussion of studies regarding perceptions of discrimination and an analysis of the discrimination claiming process.

that a debate on what actions should be recognized by law as discriminatory and a reliable metric of discriminatory behavior becomes more difficult.

Obviously, the “cause” and “no cause” determinations do not account for all the charges. Many cannot be determined in a reasonable time frame, which can be influenced by the agency’s available resources. Regardless of determination, all charging parties receive a “right to sue” letter when the agency concludes its investigation. This indicates that the charging parties have exhausted possibilities for remediation through EEOC’s administrative process and are free to pursue their claims in court. Figure 2 illustrates the rise in charges corresponding to the number of actual and potential employees age 16 years and over in the civilian labor force (CLF). The linear trend lines, applied to data that has notable nonlinear variability, indicate that CLF has been increasing faster than charges. Although the number of charges is substantial, there is a relative decline in proportion to the CLF.

Figure 3 addresses EEOC’s litigation activity and total monetary recovery. “Merit” suits address allegedly unlawful employment practices, in contrast to ancillary activity such as subpoena enforcement. Monetary recovery includes court-mandated relief, voluntary settlements after the commencement of litigation, and pre-litigation conciliation agreements. In addition to direct financial relief for claimants, monetary relief includes the value of injunctive relief, such as diversity training.

There is a marked drop in the number of suits filed by the agency after FY 2001. This trend continued during the presidential administrations of George W. Bush (Republican) and Barack

H. Obama (Democrat). The Bush years saw roughly twice the number of suits as the Obama years.

Monetary recovery shows a different trend, with annual amounts generally rising. Every year of the Obama administration saw a higher total than the highest year's recovery for the Bush administration<sup>40</sup>.

Maatman, Janice, and Karasik (2016) see three metrics that underlie EEOC's systemic program as indicative of success for obtaining legal compliance: money recovered for plaintiffs or for injunctive relief, number of people benefited, and amount of structural change. Of these, only the first is featured in the agency's statistics on its website. The others find mention in reports on the agency's Strategic Enforcement Plan, although little is mentioned specifically on structural change. Schlanger and Kim (2014) note that EEOC's description of its systemic litigation program in 2006 emphasized merits over monetary value, which would imply tracking of a nonmonetary metric. Of course, the number of suits successfully concluded by settlement or judicial decision is a metric for litigation;<sup>41</sup> taking a financial bite out of organizations is a way to focus attention on systemic discrimination issues.

EEOC statistics show that the recent increases have more to do with settlements than successful suits. Often these settlements are occurring at the pre-suit conciliation stage.

---

<sup>40</sup> These figures have not been adjusted for inflation, which has generally been low in recent years.

<sup>41</sup> But as Kim (2015) noted, quantity is not a substitute for quality. High-impact systemic litigation can be considerably more resource-intensive than run-of-the-mill suits.

Figure 4 depicts federal EEO court filings. The trends for both the trial and appellate courts are for fewer suits, despite an increase in district court activity corresponding to the recession years. Relative scarcity and an uneven pattern makes it difficult to characterize class action suits. For 1997-2016 the average was 105 suits per year; this activity had its ups and downs, cresting at 142 in 2012, but declining after that to 100 in 2016. Because of the U.S. Supreme Court's decision in *Wal-Mart v. Dukes* (2011), several commentators (e.g., Kim, 2015) have expected a decline in the number of suits due to difficulty in meeting class certification requirements for commonality of interests<sup>42</sup>.

The outcome of those suits is not easy to determine. Clermont and Schwab (2004, 2008), using earlier data than the data here, found that plaintiffs tended to be unsuccessful both at trial and appeal. They noted that many cases derived from court data had missing outcomes, or missing identification to link trials with appeals<sup>43</sup>. Nielson and Nelson (2005) discuss a “pyramid model” wherein many grievances at the base of the claiming process ultimately result in few adjudications at the pinnacle. They also supply a detailed examination of claims from charge to disposition for 1990-2002.

Terpstra and Honorée (2016) sampled 401 federal court cases and found that plaintiff wins were 33% for the private sector, 32% for the federal sector, and 34% for the state/local sector.

Plaintiffs in race cases fared worse if they were in the federal sector (17%), while plaintiffs in sex cases fared best in federal court (67%).

---

<sup>42</sup> *Wal-Mart's* impact is discussed below.

<sup>43</sup> An attempt by the author to replicate with more recent data encountered similar difficulties.



For suits that are brought to successful conclusion by the class, Selmi (2003) had noted that outcomes had become more concerned with monetary transfer to the victors and their counsel rather than with organizational reform. If reform to promote EEO is the focus, then neither the number of suits nor monetary recovery seems an appropriate metric.

One key demographic is the movement of previously excluded groups into organizational management, and whether litigation has furthered that result. Kalev and Dobbin (2006) found little evidence that lawsuits were improving the status of women and African Americans.

Compliance reviews, in contrast, had a greater capacity for organizational change, but the effect depended on the regulatory environment. They concluded that litigation and compliance reviews had more than a ceremonial effect on advancing representation. In contrast, “In the current period, EEOC charges, OFCCP compliance reviews, and lawsuits seem to produce as much, or more, backlash as further equal opportunity progress” (Stainback & Tomaskovic-Devey, 2012).

Figures 5-10 are based on EEOC’s published EEO-1 statistics. The figures provide employment counts and relative representation for Asian, Blacks, Hispanics (collectively, non-Whites), and Whites<sup>44</sup>. Pacific Islanders and Native Americans are also tracked by EEOC but were not included here because they0 comprise a small portion of the national workforce. Starting in 2007

---

<sup>44</sup> Asian, Black, and White are considered as racial groups. Hispanic is an ethnic group whose members may be of various races. The groups are intended to be mutually exclusive; Hispanic gets priority over racial group. Group names here are those used in the EEOC tables.

the managerial category was divided into executives and senior managers (hereafter, “executives”) and mid- and first-level managers (hereafter, “low-mid”)<sup>45</sup>.

Note that the scale for the non-White groups (Asian, Black, Hispanic) are on the left of charts in this series. Whites are charted with the scale on the right. Whites are far more numerous than non-Whites; the units of the left and right scales generally differ. Because the non-White groups have the same scale, their relative numbers and representation are comparable. Caution should be used in comparisons involving the White group. Apparent change relative to the non-White groups may be exaggerated due to scale differences

Figure 5 indicates a drop in the number of executives in the 2007-2015 period, with most of the loss coming from Whites. There is a general decline noticeable around 2008, perhaps a product of the recession years. By 2015 all demographic groups were gaining numbers, but the trajectory of gains varied by group. Only Asians had increased their numbers in 2015 relative to 2007.

Figure 6 shows the decline in relative White representation among executives. Blacks were less represented at the end of the period than at the beginning, while Asians and Hispanics made gains.

Figures 7 and 8 provide a similar depiction for low-mid managers. Again, there is a dip in numbers around 2008 but increases thereafter. Numbers of Blacks dipped for most of the period but finished ahead by 2015. Whites, despite a growth in numbers, made up about 3% less of the

---

<sup>45</sup> There is nothing of intrinsic importance in starting with 2007. The intent was to provide about ten years of data; other studies cover previous years. Hawaiians were included, and the managerial category was split into higher and lower levels, starting in 2007.

managers by the end of the period. Blacks had a relatively small increase in representation; Asian representation increased throughout the period.

Figures 9 and 10 depict professionals. Despite the 2008 dip, professionals in all four demographic groups grew. White representation dropped by slightly more than 3%.

Kuang and Archer (2014) argued that gains for minorities in the managerial and professional ranks are a metric for the effectiveness of litigation since the Civil Rights Act of 1991's changes to Title VII, and more specifically to EEOC's enforcement efforts. Using the EEO-1 data described in this paper but for 1998-2012, they found that minority gains were either flat or were equivalent for those of Whites. They also noted that "minorities" obscured differences among constituent demographic groups. The analysis is different from what is reported here. Kuang and Archer (2014) cite to the EEOC (2011) statistics of workforce composition; there, as here, constituent demographic groups' representation adds to 100%. But their own analyses seem to be based on the percentage of groups moving into managerial or professional ranks. This seems to be more a measure of occupational mobility within demographic group, rather than a measure of demographic diversity within occupation.

Timing, if not everything, still makes a difference in the analyses. EEOC (2011) data indicating that minority representation in the Officials and Managers category increased from 1.8% to 8.2% from 1966 to 1983. That comes to a cumulative 6.4% over 18 years, an average of 0.35% per year. The data used in this paper for Executives, the "high end" of the older managerial

category, indicated a cumulative increase for minorities<sup>46</sup> of 1.43% from 2007 to 2015, and average increase of 0.16%. That might suggest a stall to progress. But if only the more current years 2012-2015 are examined, the cumulative increase is 2.21%, an annual average of 0.55%. Numbers tied to time frame and not causally linked to litigation practices perhaps are not a good metric for the relationship between litigation and diversity. If anything, the suggestion is that the marked decline in private and EEOC litigation has not returned the country to its occupational segregation of a half century ago.

### **Diversity Metrics**

The previous section contained some discussion of demographic diversity for managers and professionals.

For employers who wanted to avoid unlawful discrimination for ethical, business, and legal reasons, the range of voluntary action could extend from simply implementing policies and procedures that were nondiscriminatory, through various forms of training to avoid prejudice, to affirmative recruitment and preferential hiring.

As noted above, diversity can be divided into valuing diversity and managing diversity. But it can be said—and has—that diversity has stalled at legal compliance and demographic head-counting<sup>47</sup>. Legal compliance is important to the organization, and head-counting to ensure

---

<sup>46</sup> This includes Hawaiians, American Indians, and Two or More Races, as well as the three larger non-White demographic groups discussed previously.

<sup>47</sup> Recruiting a demographically diverse workforce can be said to fall under “managing diversity.” But an organization might embrace programs for “valuing diversity” to attract and retain demographically diverse talent. Thus, valuing diversity might also support headcount efforts.

representation is an understandable metric. The issue is what happens to the people after they are in the organization, a problem that has been described as shifting from a diverse workforce (numerical representation) to a diverse workplace (welcoming environment). “Diversity and inclusion” has become the more comprehensive term.

To some degree there seems to be consensus, backed by research, that diversity-as-headcount is fostered through organizational goal-setting and managerial accountability (Kurtulus, 2016; Dobbin and Kalev, 2007). These efforts have been directed at getting demographic groups into the workforce. They have not been as much concerned with what happens to people in historically under-represented groups once they get in. Presumably that should be covered by programs that encourage employees to “value diversity.” Whether they in fact prevent discriminatory behavior such as outright harassment and the erosion to morale by less overt hostility, biased performance evaluation, and exclusion from developmental opportunities necessary for advancement that militate against further equal employment opportunity is part of the current issue.

A fundamental problem is lack of consensus on what diversity initiatives should be accomplishing<sup>48</sup>. An available metric (e.g., headcount) might not have relevance for a particular intended purpose (e.g., stopping harassment). Accomplishing the purpose, rather than tying a program to some favorable indicator, needs to be the focus. As Gilrane, McCausland, King, and

---

<sup>48</sup> Heitner, Kahn, and Sherman (2013) used a panel of diversity experts to ascertain essential components of diversity initiatives’ success. They arrived at three key components: employee perceptions, organizational climate and culture, and measures based on employee lifecycle processes (attract, develop, and retain).

Jones (2012) noted, the important question is not whether diversity is associated with positive organizational outcomes, but what the organization is doing to encourage those outcomes.

Because organizations do not have the same expectations for their diversity programs, there is no one type of metric that applies universally. Table 1 gives a typology of metrics cited in the literature and provides some indication of those expectations and how they are measured; see Brenman (2013) for an expanded list.

Some metrics proposed for diversity have organizational scope, and it may be that this extent disqualifies these as specifically tied to diversity. Equitable compensation practices, for example, are applicable to all employees; to the extent that it is a diversity practice, it would seem to be only that good human resources management (HRM) sits well with a diversity of employees, i.e., all of them. Organizational indicators such as the firm's worth or financial performance have been claimed for diversity, but those results have been claimed for HRM without specific mention of diversity (e.g., Ulrich, 2016). Having diversity programs may itself be a product of a strong HRM presence<sup>49</sup>.

King, Gulick, and Avery (2010) indicated that by 2005 two-thirds of companies had some form of diversity training. The authors make a distinction between training (emphasis on skills and behavior for practical organizational needs) and education; the latter connotes knowledge, possibly without application. Both seem to share a lack of theoretical underpinning.

---

<sup>49</sup> Conversely, Pager (2016) mentioned that one possible reason that discriminating firms go out of business is that management with policies in this regard have other policies that are also detrimental to their organization.

TABLE 1: A Crude Taxonomy of Diversity Metrics		
METRIC	CONTENT	ACTIVITIES
Headcount	Demographic representation	Barrier analysis; recruitment and selection
Checklist	Policies and procedures building to a critical mass supporting EEO	Portfolio of programs
Throughput	Numbers exposed to diversity initiatives	Policies on who gets exposed, when
Learning evaluation	Outcomes from trainee reaction to assessing change in behavior	Outcome levels studies(Kirkpatrick, 1987)
HR programs with inclusiveness and headcount considerations (e.g., recruitment, tenure)	Sound HR management practices and its evaluation as applied broadly across organization, with some emphasis on subgroups; may tie to macro metrics, such as stock price.	Suite of HR programs
Employee surveys	Employee perceptions linked to process and outcome data.	Action planning based on survey results

A SHRM (2014) survey of diversity practices found that most respondent employers:

- Cover only ethnicity, gender, race, or age (50% to 61%, by category);
- Do not measure the impact of diversity practices (65%) or are not sure if they do (16%);
- Do not measure return on investment for diversity initiatives (93%); and
- Tend to house diversity activities within the HR function; few (16%) have dedicated diversity staff;

In addition, larger organization were more likely to have diversity programs than smaller ones<sup>50</sup>.

Kulik and Roberson (2008) identified 31 studies involving diversity training; only half addressed diversity-related behavior, and most of these were based on self-assessment.

“Valuing” has at least two aspects: value for the organization from diversity and having employees value the diversity of their co-workers. “The business case for diversity” seems essentially a headcount argument. By having a sufficiently diverse workforce, the organization is better able to recruit and retain a stellar workforce, gain access to diverse markets locally and globally, and better satisfy a diverse customer base. Insofar as diversity is linked to business interests, the metrics would be those of the business. This likely involves confounding the

---

<sup>50</sup> SHRM studies seem well designed, but participation rates tend to be limited. That leaves open what non-respondents are doing. Bendick, Egan, and Lofthjelm (2001) found that there is no good count on diversity programs and so a random sample was impossible. Their alternative approach was a stratified convenience sample with 73% response rate, yielding a completed sample of 108 for telephone interviews with a structured questionnaire. Still, convenience samples leave open questions of how representative were the respondents.



diversity effect with that of other aspects of business strategy. Diversity advocates who favor the business case concept would argue that diversity needs to be evaluated as any other business strategy. Advocates not so much in favor question whether a “business case” is necessary for a self-evident good, or if the business case concept implies that a commitment to diversity is subordinate to other business considerations. There is that “validity-diversity dilemma” that implies that the most able workforce may not be the most diverse<sup>51</sup>.

Arguably, exploiting demographic diversity can become exploitation as when employees are typecast for certain roles relating to their demographics but are not valued otherwise. There is also an issue of using diversity to satisfy discriminatory customer preference (Bendick, Egan, and Lanier, 2010)<sup>52</sup>. Increased demographic representation under those circumstances is likely not a good EEO measure.

A variation on the business case theme is the role diversity plays in team effectiveness. The issue, however, is the difference between diversity in characteristics contributing to successful

---

<sup>51</sup> See, for example, McKay and McDaniel’s (2006) cumulative investigation of black-white performance differences. As had previous reviews, this study found on average a small difference in favor of whites; the magnitude varied with the underlying competency and assessment method. Of course, the relevance of the difference and the assessment method are the stuff of lawsuits. Nielsen and Nelson (2005) noted that flipping the switch tomorrow to end all employment discrimination would not eliminate all labor market inequality. But the problems are not just historical disadvantage and hierarchical relationships among groups. Likely it is an interplay of nature and nurture factors long before a person’s first job application. None of this should deter efforts to flip that switch.

<sup>52</sup> *Knight* (1981) is an example of the conflict between the business case and potential exploitation. An African American police detective was assigned to community affairs and was good at his job—so good that he was not allowed to transfer to mainstream police work.

outcome and more surface characteristics. Demographic diversity may constitute more of an occasion for conflict than cooperation (King and Gilrane, 2015).

As to valuing diversity, discontent with diversity programs as effective resolutions of workplace inequity are longstanding. The matter was discussed by Kalev, Dobbin, and Kelly (2006; see Vedantam [2008] for a summary); a decade later, there was more of the same from Dobbin and Kalev (2016).

At issue is that while the presence of programs and the counts of employees annually provide measurable activity, producing results beyond signaling some willingness to address inclusion is lacking for many. Such programs may signal that the organization does not want discrimination; what is signaled regarding diversity and inclusion may be another matter.

It would be an overstatement that diversity training does nothing. Bezrukova, Spell, Perry, and Jehn (2016), in a study statistically summarizing over 40 years of research found a moderately strong overall effect for diversity training. Larger effects were for reactions to training and cognitive learning<sup>53</sup>; the latter persists over time. Smaller, more transient effects were found for behavioral and attitudinal/affective training. They concluded that “contrary to charges made by our predecessors over the years, diversity training research is no longer atheoretical, irrelevant, or dull” (p. 1246). King and Gilrane (2015) provided a short introduction for practitioners on evidence-based diversity initiatives. The issue is what kind of training, and for what result. For

---

<sup>53</sup> It was not clear to this writer what content cognitive learning included. One example (p. 1230) seems to involve factual information on Mexican culture. What the authors found “rarely encountered” (p. 1246) were studies that related training to discrimination and aspects of inclusion.

EEO, the issue is whether training just for mitigating liability for discrimination is sufficient. Bendick, Egan, and Lofthjelm (2001) presented the issue as whether the training was aimed at organizational development, i.e., changing the workplace; in their study, they estimated that about a quarter of programs have enough effort to be considered pursuing this aim.

### **The Interaction of Litigation and Diversity**

It may be that litigation and diversity need to be more focused on a common goal: encouraging workplace change. And their partners in this of necessity are HR management and social science—and organizational management as well.

This is not a simple task: Litigation without follow-up is unlikely to promote inclusion. Litigation is an adversarial process; organizational change needs to be cooperative<sup>54</sup>. Litigation without follow up is unlikely to promote inclusion. Selmi and Tsakos (2015) noted that employment discrimination cases characteristically involve some injunctive relief, but it tends to be a perfunctory matter. A “gladiatorial” contest (Schlanger & Kim, 2014) is marked by winners and losers, and the fighting can continue beyond the court’s decision. Where the employer thinks that the outcome was not fair, or there was a settlement accepted only to avoid the cost of protracted litigation, there may be little inclination to cooperate. The result is likely to be imposition of some specific, time-limited activities.

---

<sup>54</sup> At a panel discussion of consultants dealing with the issue, an analogy between plaintiffs’ victory and the American Civil War/Reconstruction Period was mentioned: fierce battles culminating in an imposed peace, sullen compliance from the defeated on civil rights issues, and backsliding when the enforcement pressure let up (Dickson et al., 2016). The implication was that this was sub-optimal.

In contrast, the “collaboration” model features cooperation with management and an “experimentalist” approach to solving problems of unequal treatment rather than a rigid follow-up to the litigation. It may also include court supervision to ensure that the terms of settlement are fulfilled as specified. Such a model requires mutual trust and commitment, and expertise to develop possible solutions and to refine them as necessary; it likely will require the investment of resources and time. “Collaboration,” of course, has both positive and negative connotations: cooperation for the common cause and selling out the cause.

At issue is not only whether there is follow up but whether it is effective. One complication is confusion regarding a HRM infrastructure and structural reform to end discrimination. The critique of the “managerialist” (Schlanger & Kim, 2014) response to discrimination is that it implements management practices widely accepted regardless of civil rights impact. Although Schlanger & Kim (2014) acknowledge the necessity of HRM, they specifically criticize EEOC’s injunctive practices in systemic cases as “pursuing standard, bureaucratic personnel practices.” The problem is not confined to just EEOC’s cases.

The managerialist critique has several elements that need to be examined here:

- The managerialist approach treats discrimination as a typical managerial problem with typical policy-and-procedure solutions. Adoption of, for example, a diversity program that is in line with what the organization does routinely (e.g., classroom training) can be quickly implemented and signals the organization’s intent to do something about discrimination. Insofar as the point of “doing something” is defined primarily as reducing legal liability, it may be successful.

- Reduction in liability alone does not necessarily deal with the underlying discrimination; it may serve only to mask it. Edelman and her colleagues (Edelman, Krieger, Eliason, Albiston, & Mellema, 2011) have noted, to their dismay, that courts are influenced by the mere presence of programs intended to alleviate workplace discrimination, regardless of their effectiveness.
- From the HR rather than legal perspective, “standard, bureaucratic personnel practices” enable the organization to function. These likely are not primarily directed at EEO. It would be the rare organization that functions primarily for the sake of providing EEO. However, the furtherance of EEO depends on how processes to select, retain, develop, and compensate employees are conducted. Whether these practices accomplish their intended purposes or are merely “cosmetic in nature,” “primarily designed to address public relations problems,” or “symbolic”<sup>55</sup> is a matter of how those purposes are defined and the metrics appropriate to the practice. For example, personnel selection procedures are expected to be valid for their purpose; there are professional standards and federal government guidance for demonstrating validity.
- Presumably it is in the general interest of management and employees alike that practices be effective, efficient, and fair. That does not preclude use of ineffective practices because of ignorance, lack of perceived need, cost considerations, or nefarious motives. Dobbin and Kalev (2016) noted that managers may not like to be told what selection procedures to use, and so they do not use them or else use them selectively. But that does not negate the value of good selection procedures as a means to further EEO.

---

<sup>55</sup> These are descriptive terms that various critics have applied, as recounted by Schlanger and Kim (2014, p. 1586).

- As Schlanger and Kim (2014) note, structural reform goes beyond specific legal issues to the structure of work that permits bias to operate within the organization. This implies an integrated solution to the problem, not isolated remedial actions. Rather than managerialist, what may be termed here as a fully “structuralist” orientation for the structural discrimination problem recognizes that an integrated system of practices to manage the workforce, both to dismantle barriers to EEO and to maintain a well-functioning organization. This system is the foundation for implementing and maintaining EEO. Obtaining this structure is an aspect of structural reform. It cannot be the totality of structural reform because the mechanisms of discrimination are not exclusively in the formal practices of the organization. But formal practices can be used to control those mechanisms.
- The new “business case” is that effective management practices and EEO outcomes are linked.

### **Litigation and Diversity**

Advancing EEO depends on organizations’ moving beyond minimizing litigation exposure. The matter is how to encourage this.

This starts with what has now become the often-repeated call to do more with injunctive relief than getting agreement that the employer will not do anything unlawful, and maybe put some employees through diversity training. Not so common are widely-disseminated descriptions of a new paradigm. I-O psychologists have a description of work done with Coca-Cola (Goldstein and Lundquist, 2010). Offermann and Basford (2014) have a set of cases on inclusion and HR

practices. Hegewisch, Deitch, and Murphy (2011) report on a study of more than 500 consent decrees. Schlanger and Kim (2014) discuss a “gladiatorial” case with consultative elements involving EEOC.

This literature is more descriptive than prescriptive; the final how-to book is still to be written. EEOC’s report on harassment (Feldblum and Lipnic, 2016) noted promising practices; this has joined other lists of practices and encouraged more.

Currently there is much discussion in HR and general management about employee engagement. Employees who do not feel included likely do not feel engaged. There is already a body of research and practice (although not necessarily definitive solutions) on which to build<sup>56</sup>.

This interest converges with research on “microaggressions,” a continuum of prejudiced behavior from the blatant to the unconscious.<sup>57</sup> These practices seem to be both a manifestation and cause of lack of inclusiveness. Microaggression has a long conceptual history but a shorter research one. A primary question is whether reaction to these practices is serious enough to gain the attention of employers. King et al. (2010) noted that they were serious enough for the recipients. A review of the research literature (Jones, Peddie, Gilrane, King, and Gray, 2013)

---

<sup>56</sup> The U.S. Office of Personnel Management (2015) has an introduction to engagement. Eisenberger, Malone, and Presson (2016) discuss managerial support for employee engagement.

<sup>57</sup> The term is synonymous with “everyday discrimination” and “micro-inequities.” See DeAngelis (2009) for a very brief introduction, and mention of controversy. The term originated in clinical work. Because it connotes a continuum, a specific instance might be a matter for legal redress, a structural issue for organizations, or an interpersonal dynamic to be resolved by individuals. Current EEO discussion focuses on the structural aspect.

concluded that the problem was at least as severe as with overt discrimination, affecting the well-being of individuals and the performance of organizations.

### **Litigation and Social Science**

The discussion above cites social science research regarding structural issues such as organizational culture and implicit bias as causes of discrimination. The suggestion for furthering EEO envisions a collaborative, rather than gladiatorial, approach. A possible alternative is for more litigation headed by EEOC, the governmental agency most critical to the furtherance of EEO. Some commentators have seen the agency's key role as stepping into the breach in class actions due to Wal-Mart; class certifications have become more difficult for private plaintiffs but EEOC is not subject to the same limitations. In the earlier years following the passage of Title VII, theories of structural reform "emphasized dramatic legal struggles to transform recalcitrant institutions" (Schlanger and Kim, 2014). The theory also included a willingness for long-term oversight of organizational change. Perhaps EEOC could revive this version of structural reform, although Kim (2015) notes several constraints: increased burden on the agency's resources, the need for complex litigation with expert witnesses, and procedural matters.<sup>58</sup>

Green (2017, p. 149) has criticized the theory of "organizational innocence" wherein "the organization's sole responsibility is to make complaints feasible for victims and respond adequately to any individual complaint." This criticism is in line with other writers, and with the

---

<sup>58</sup> Since the article was written, one constraint, judicial review of the agency's conciliation efforts, has been resolved generally to limit the constraint. Another matter, Congressional and White House oversight of the agency's activities, if anything has become more salient.



EEOC's report (Feldblum and Lipnic, 2016). The issue is how to get the organization to take more responsibility. Presumably an alternative to collaboration is to change the legal rules and force compliance. But there seems little in the offing that suggests a resurgence in EEO litigation, "game changing" new theories accepted by the courts to expand the scope of discrimination coverage, or a massive insurgence of resources for the EEOC to spearhead new initiatives.

A concern is a "magic bullet" view of social science research to which advocates of increased litigation may be drawn.<sup>59</sup> This appears in commentary where the plaintiffs would have prevailed in Wal-Mart but for the U.S. Supreme Court's disfavoring class actions with their social science and statistical arguments.

Selmi and Tsakos (2015) note that, in addition to the Court's attitude, perhaps there was an issue with how social science was used in Wal-Mart. Class claims based on subjective employment practices remain viable. But in Wal-Mart the social science testimony relating subjective employment practices and biased stereotypes was generic, essentially emphasizing a matter not in dispute: Decision-makers relied on their discretion. Perhaps that shows potential vulnerability to bias; it is not the same as showing evidence of discrimination.

---

<sup>59</sup> For example, Green's arguments are based on "the cognitive bias revolution" (2017, pp. 29-32), noting "what really blew cognitive bias onto the main stage" was the Implicit Associations Test (IAT). Many social scientists likely have awareness of cognitive bias research; whether many think there is a revolution in progress is another matter. More definite is the controversy regarding the IAT, with competent researchers on both sides arguing its implications. Someone ascribing such impact to the IAT might take note of the cumulative review of research with that instrument; Oswald, Mitchell, Blanton, Jaccard, and Tetlock (2015) have what, for the moment, is the latest word. Building public policy recommendations on a line of research with disputed practical meaning (not only the IAT) is the proverbial building on sand.

Selmi and Tsakos (2015) note that any of the following could have changed the outcome: a discriminatory policy or a clear culture of discrimination; smaller size of the class or the employer; or a specific practice or directive that informed the decision-making. Class actions are not dead, nor are social science arguments regarding “second generation” discrimination. But merely to assert ubiquitous bias with unknown effect likely will not gain traction.

Another problem was statistical evidence in support of a “pattern or practice” disparate treatment case where some units indicated that women were disfavored, but others indicated the opposite, and units with few employees could not produce a definitive statistical conclusion. Bielby and Coukos (2007) describe the issues pre-Wal-Mart; Gastwirth, Bura, and Miao (2011) in a post-Wal-Mart article noted alternative approaches to establishing statistical patterns. In particular, they noted the futility of arguing on the basis of how many locations showed a statistically significant disparity in favor of either men or women. Their re-analyses generally supported Dukes et al.

EEOC is not subject to private plaintiff class certification rules; it is still subject to presenting a persuasive case. This does not necessarily allow for filing a multitude of class cases. Walker and Monahan (1987) first wrote of “social framework” as an essentially neutral way to assist triers of fact. Monahan, Walker, and Mitchell (1998) criticized “social framework analysis” used in Wal-Mart not only for conflict with their concept, but also for being problematic on both legal and scientific grounds.

To get at the pervasive but relatively minor objectionable behavior, possibly the law might be changed. The U.S. Supreme Court, in its landmark harassment decision, noted that Title VII is not “a general civility code for the American workplace” (Oncale, 1998); possibly courts will spontaneously become more inclined to accept the seriousness of microaggressions. But neither possibility seems a probability.

Nothing here should be construed as opposition to developing and applying social science evidence in litigation, as appropriate, for either plaintiff or defendant. But the role of this evidence and expert testimony may at the moment be more for providing a coherent narrative that ties together anecdotal and statistical facts, rather than constituting the case-in-chief.

Moreover, simply forcing more charges into litigation without serious consideration of injunctive relief does not solve the structural issues noted here. But getting organizations on board with both avoiding liability and improving productivity just might work.

### **A View Toward Action**

- Meaningful injunctive relief. If litigation occurs and the plaintiffs prevail or obtain a settlement, there should be an effort to incorporate injunctive relief as appropriate for the situation. An additional consideration for enforcement agencies is whether to suggest changes when an investigation does not result in legal action. This can be risky; the agency would not want to be implicated in what the organization does that subsequently is found to be discriminatory. But with safeguards, there is the potential for effecting change beyond what currently results from such investigations.

- Employee involvement. The battle for inclusion is fought at ground level with the employees directly involved. After all, the organizations' culture is embodied by its workforce. Empower managers and workers to devise policies that reduce discrimination<sup>60</sup>.
- Comprehensive approach: Tactics are employed at the individual and organizational level. Behaviors likely to further inclusion are mutual psychosocial support and confrontation (calling out behaviors that undermine support. At the organizational level evidence-based practices that further the organization's mission and EEO are integrated. This includes HR practices.
- Metrics. One way to assess whether employees perceive themselves as included is to ask them. Surveys accompanied by action planning to address perceived issues constitute one tool at present. Morgan, Dunleavy, and DeVries (2016) foresee the possibility of Big Data applications. Patterns in recruitment, compensation, and development could be mined. Monitoring to record data in real time on the interactions of employees could shed light on intergroup dynamics. Email might be subject to "sentiment analysis." Prevalent stereotypes in the workforce could be identified. One might wonder if the authors' analyzing the "furrowed brow" of an employee in repeated video recorded interactions is a bit Orwellian. Perhaps less intrusive are interaction analyses that are used to foster collaboration and mentoring among employees with common business or professional interests.

---

<sup>60</sup> This and some of the following suggestion come from Nielson and Nelson (2005). Some of their suggestions (e.g., increasing the budget for EEOC) may not be likely in the near term; while desirable, such are omitted.

- Scientific consensus. There is a tendency of some legal scholars to ascribe their advocated positions to “social scientists,” implying a unified view in the scientific community. With issues relevant to employment discrimination, there are likely to be differences among researchers, and differences in the researched subject matter and methodology by academic discipline. Having a baseline of consensus regarding what science can say about discrimination and its prevention would be useful. Currently cumulative reviews of theory and research, and meta-analyses across quantitative studies, provide some of that baseline, but progress would be furthered by inter-disciplinary co-operation. The onus falls on the various professional associations, but some encouragement from other stakeholders (government, antidiscrimination advocacy groups, employer associations, the legal profession) would help to make it happen.
- Judicial education. The Reference Manual on Scientific Evidence (Federal Judicial Center, National Research Council, et al., 2011), a reference for judges, has mention of discrimination cases, but only as they illustrate statistical issues. Perhaps the time has come for discussion of the complexities of expert evidence in this area. It may also be time to re-boot Walker and Monahan’s initial use of social frameworks to assist triers of fact. Doing so implies that the scientific consensus described above has been realized to a sufficient degree.
- Some form of immunity. Employers are likely to be reluctant to allow a bunch of social scientists, particularly those disposed to find structural discrimination, free run of their organizations. It is an issue now for organizations performing self-audit to cloak their efforts in attorney-client privilege to avoid issuing invitations for litigation. But the

objective is to encourage employers to take proactive steps. This matter might involve discussions with enforcement agencies.

- New adjudication policies. As noted above, microaggressions are a continuum. Perhaps there is a need for some form of redress short of litigation, mediation or some other form of adjudication apart from internal grievance procedures<sup>61</sup>.

The desired end state is less discrimination. The metric for this is uncertain. Surely the number of organizations taking action is one possibility, but that alone ignores effectiveness considerations. A decline in the number of discrimination charges is an ultimate goal, but in the shorter term perhaps renewed emphasis on ending discrimination encourages claiming.

Good practice drives out bad, or so one might hope. Small organizations using facially neutral but discriminatory management procedures may never produce a viable impact charge. But

### **Caveats**

- “Old fashion” discrimination is still there. Inclusion might do much to alleviate less intentional forms of discrimination. But where the intent is there, the enforcement mechanisms may remain the same. Over time one would hope that, just as blatant discrimination has lost popularity with vigorous EEO enforcement and changing social attitudes, the less blatant will also wither.

---

<sup>61</sup> Silver-Greenberg and Gebeloff (2015) highlight problems with arbitration as an extra-judicial means of dispute settlement. Presumably any new alternatives would need to address the possibility of similar problems. There may be less of an issue where the disputes are on the less severe end of the discrimination continuum, rather than devising alternatives for handling allegations of unlawful practices.

- Inclusion is a “managerialist” solution. Effective workforce management is HR. Diversity efforts for inclusion are integrated into HR and other business practices. Conflict resolution, particularly for those conflicts that are not legally cognizable, is an internal process. Criticism that internal grievance processes have done little to get at causes of discrimination may still apply.
- The “validity-diversity dilemma” is still with us. This could still be a problem in hiring and promotion. A longstanding issue is that more job-relevant assessments of competency do not further diversity when, in fact, there are differences in level of competency across demographic groups. Apart from these technical matters, a basic challenge is to ensure that diverse demographic groups are equipped to participate equally in employment opportunity. A line of research has suggested that in many work situations the super performers make a far greater contribution than average performers; presumably in fairness these stars should be paid accordingly. Should demographic differences be linked to stardom, pay equity becomes an even more complex issue.
- The objection that diversity is but a pale substitute for affirmative action may remain. The impetus for inclusion is, of course, to help the previously excluded. But insofar as it aims to include everyone, it is not necessarily targeted to specific groups. On the other hand, special emphasis may not be contradictory to inclusion, depending on the specifics of the organization. A criticism of past diversity programs has been that a color-blind, rather than multi-cultural, approach led to increased subtle discrimination (Jones et al., 2013). Another complication is that attaining inclusion by fostering a superordinate team identity could backfire by forcing suppression of demographic identity.

- Bureaucratization of personnel policies may or may not affect managerial diversity. Formalization can limit discretionary bias, or it can lock in separate career pathing (Kalev and Dobbin, 2006).
- An issue with “civility training” is that it imposes a code of political correctness as to what can or cannot be said, and by whom<sup>62</sup>. This situation at best clamps down on visible bad behavior. At worst, it fosters resentment and could promote the form of aversive racism where inclusion is stifled for fear of offending.
- Any system can be gamed. Measuring is not the same as fixing. The lessons learned from previous organizational efforts with HR and diversity programs should not be forgotten.
- There are no magic bullets to shoot down prejudice, only tools in law and social science regarding EEO are continuously evolving. What that leaves us with is a lot of hard work.

---

<sup>62</sup> Feldblum and Lipnic (2016) noted that the National Labor Relations Board’s objections to civility rules as potentially suppressing collective action is a problem.



## References

- Bates, S. (2013, January 29) Beyond diversity: Employers focus on inclusion. SHRM. Retrieved from <https://www.shrm.org/resourcesandtools/hr-topics/behavioral-competencies/global-and-cultural-effectiveness/pages/from-diversity-to-inclusion.aspx>.
- Bendick, M. Jr., Egan, M.L., & Lanier, L. (2010). The business case for diversity and the pernicious practice of matching employees to customers. *Personnel Review*, 39, 468-486. Retrieved from <http://www.bendickegan.com/pdf/BendickEganLanierFinalJan2010.pdf>.
- Bendick, M. Jr., Egan, M.L., & Lofthjelm, S.M. (2001). Workforce diversity training: From anti-discrimination compliance to organizational development. *Human Resource Planning*, 24, 10-25.
- Bezrukova, K., Spell, C.S., Perry, J.L., & Jehn, K.A. (2016) A meta-analytical integration of over 40 years of research on diversity training evaluation. *Psychological Bulletin*, 142, 1227-1274.
- Bielby, W.T. (2005). Applying social research on stereotyping and cognitive bias to employment discrimination litigation: The case of allegations of systemic gender bias at Wal-Mart stores. In L.B. Nielsen & R.L. Nelson(Eds.), *Handbook of Employment Discrimination Research*. Dordrecht, The Netherlands: Springer.
- Bielby, W.T. & Coukos, P. (2007). "Statistical dueling" with unconventional weapons: What courts should know about experts in employment discrimination class actions. *Emory Law Journal*, 56, 1563-1612.
- Brenman, M. (2013, January 9). Diversity metrics, measurement, and evaluation. Workforce Diversity Network. Retrieved from [http://www.workforcediversitynetwork.com/res\\_articles\\_diversitymetricsmeasurementevaluation.aspx](http://www.workforcediversitynetwork.com/res_articles_diversitymetricsmeasurementevaluation.aspx).
- Clermont, K.M & Schwab, S.J. (2004). How employment discrimination plaintiffs fare in federal court. *Journal of Empirical Legal Studies*, 1, 429-458. Retrieved from [http://scholarship.law.cornell.edu/lrsp\\_papers/1](http://scholarship.law.cornell.edu/lrsp_papers/1).
- Clermont, K.M & Schwab, S.J. (2008). Employment discrimination plaintiffs in federal court: from bad to worse? *Harvard Law & Policy Review*, 3. Cornell Law Faculty Publications. Retrieved from [http://scholarship.law.cornell.edu/lrsp\\_papers/109](http://scholarship.law.cornell.edu/lrsp_papers/109).
- DeAngelis, T. (2009). Unmasking "racial micro aggressions." *Monitor on Psychology*, 40, 42. Retrieved from <http://www.apa.org/monitor/2009/02/microaggression.aspx>
- Dickson, M. W., Goldstein, H., Klein, A., Lundquist, K., Oliver, L. & Yusko, K. (2016, April). Implementing deep organizational change in HR-focused consent decree situations. Panel discussion conducted at the 31st Annual Conference of the Society for Industrial and Organizational Psychology, Anaheim, CA.

Dobbin, F. & Kalev, A. (2007) The architecture of inclusion: Evidence from corporate diversity programs. *Harvard Journal of Law and Gender*, 30, 179-301.

Dobbin, F. & Kalev, A. (2016, July-August). Why diversity programs fail. *Harvard Business Review*. Retrieved from <https://hbr.org/2016/07/why-diversity-programs-fail>.

Edelman, L.B., Krieger, L.H., Eliason, S.R., Albiston, C.R., and Mellema, V. (2011). When organizations rule: Judicial deference to institutionalized employment structures. *American Journal of Sociology*, 117, 888-954.

EEOC (2011). Job patterns for minorities and women in private industry, 2008 indicators over time. Author. Retrieved from <https://www.eeoc.gov/eeoc/statistics/employment/jobpat-eeo1/docs/indicators.html>.

Eisenberger, R., Malone, G.P, & Presson, W.D. (2016). Optimizing perceived organizational support to enhance employee engagement. *SHRM-SIOP Science of HR Series*. Retrieved from <http://www.siop.org/SIOP-SHRM/SHRM-SIOP%20POS.pdf>

Feldblum, C. R. & Lipnic, V. A. (2016). Report of the co-chairs of the EEOC Select Task Force on the Study of Harassment in the Workplace. EEOC. Retrieved from [https://www.eeoc.gov/eeoc/task\\_force/harassment/report.cfm](https://www.eeoc.gov/eeoc/task_force/harassment/report.cfm)

Federal Judicial Center, National Research Council, et al. (2011). Reference manual on scientific evidence (3<sup>rd</sup> ed.). Washington, DC: National Academies Press. Retrieved from <https://www.nap.edu/catalog/13163/reference-manual-on-scientific-evidence-third-edition>.

Gastwirth, J.L, Bura, E., & Miao, W. (2011). Some important statistical issues courts should consider in their assessment of statistical analyses submitted in class certification motions: Implications for *Dukes v. Wal-Mart*. *Law, Probability and Risk*, 10, 225–263.

Gilrane, V. L., McCausland, T. C., King, E. B., & Jones, K. (2012). The evaluation and effectiveness of workplace diversity programs. In M. A. Paludi & B. E. Coates (Eds.), *Women and Careers in Management*. Santa Barbara, CA: Praeger.

Goldstein, I.L. & Lundquist, K.K. (2010). A five-year journey with Coca-Cola. In J.L. Outtz (Ed.), *Adverse impact: Implications for organizational staffing and high stakes selection*. New York, NY: Taylor and Francis Group (Routledge).

Green, T.K. (2017). *Discrimination laundering: The rise of organizational innocence and the crisis of equal opportunity law*. Cambridge, United Kingdom: Cambridge University Press.

Hartstein, B. (2016). A review of EEOC's systemic initiative: Tracking its progress, current priorities, and key developments in FY 2016. Littler. Retrieved from [https://www.littler.com/files/report\\_on\\_eoocs\\_systemic\\_initiative\\_0.pdf](https://www.littler.com/files/report_on_eoocs_systemic_initiative_0.pdf)

Hegewisch, A., Deitch, C., & Murphy, E. (2011). *Ending sex and race discrimination in the workplace: Legal interventions that push the envelope*. Washington, DC: Institute for Women's

Policy Research. Retrieved from <https://iwpr.org/wp-content/uploads/wpallimport/files/iwpr-export/publications/C379.pdf>.

Heitner, K.L., Kahn, A.E., & Sherman, K.C. (2013). Building consensus on defining success of diversity work in organizations. *Consulting Psychology Journal: Practice and Research*, 65, 58-73.

Jones, K.P., Peddie, C.I., Gilrane, V.L, King, E.B., & Gray, A.L. (2013). Not so subtle: A meta-analytic investigation of the correlates of subtle and overt discrimination. *Journal of Management*, 42, 1588-1613.

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.

Kim, P.T. (2015). Addressing systemic discrimination: Public enforcement and the role of the EEOC. *Boston University Law Review*, 95, 1133-1154.

King, E.B., Dawson, J.F., Kravitz, D.A., & Gulick, L.M.V. (2012) A multilevel study of the relationship between diversity training, ethnic discrimination and satisfaction in organizations. *Journal of Organizational Behavior*, 33, 5-20.

King, E.B., Dunleavy, D.G., Dunleavy, E.M., Jaffer, S., Morgan, W.B, Elder, K., & Graebner, R. (2010). Discrimination in the 21<sup>st</sup> century: Are science and the law aligned? *Psychology, Public Policy, and Law*, 17, 54-75.

King, E.B & Gilrane, V.L. (2015). Social science strategies for managing diversity: Industrial and organizational opportunities to enhance inclusion. *SHRM-SIOP Science of HR Series*. Retrieved from [http://www.siop.org/SIOP-SHRM/SHRM\\_SIOP\\_Diversity.pdf](http://www.siop.org/SIOP-SHRM/SHRM_SIOP_Diversity.pdf).

King, E.B., Gulick, L.M, & Avery, D.R. (2010). The divide between diversity training and diversity education: Integrating best practices. *Journal of Management Education*, 34, 891–906.

Kirkpatrick, D.L. (1987). Evaluation of training. In R.L. Craig (Ed.), *Training and development handbook: A guide to human resource development* (3<sup>rd</sup> ed.). New York, NY: McGraw-Hill.

*Knight v. Nassau County Civil Service Commission*, 649 F. 2d 157 (2nd Cir. 1981).

Kuang, D. & Archer, M. (2014). How far we have come: Measuring the progress of women and minorities in official/manager and professional management positions at Title VII's 50th anniversary. *EEO Insight*, 6, 11-17.

Kulik, C.T. & Roberson, L. (2008). Common goals and golden opportunities: Evaluations of diversity education in academic and organizational settings. *Academy of Management Learning and Education*, 7, 309-331.

Kurtulus, F.A. (2016). The impact of affirmative action on the employment of minorities and women: A longitudinal analysis using three decades of EEO-1 filings. *Journal of Policy Analysis and Management*, 35, 34-66.

Maatman, G.L. (2017). 13<sup>th</sup> Annual Workplace Class Action Litigation Report. Chicago, IL: Seyfarth Shaw. Retrieved from <http://seyfarth-classaction.com/WCAR2017/>.

Maatman, G.L., Janice, C.M., & Karasik, A.W. (2016, July 15). After a decade of mixed results, EEOC rebrands its systemic discrimination litigation program. Seyfarth-Shaw Workplace Class Action Blog. Retrieved from <http://www.workplaceclassaction.com/2016/07/after-a-decade-of-mixed-results-eecoc-rebrands-its-systemic-discrimination-litigation-program/>.

McKay, P.F. & McDaniel, M.A. (2006). A reexamination of black–white mean differences in work performance: More data, more moderators. *Journal of Applied Psychology*, 91, 538–554.

Monahan, J., Walker L. , & Mitchell, G. (1998). Contextual evidence of gender discrimination: The ascendance of “social frameworks.” *Virginia Law Review*, 94, 1715-1749.

Morgan, W.B., Dunleavy, E. & DeVries, P.D. (2016). Using Big Data to create diversity and inclusion in organizations. In S. Tonidandel, E. King , & J. Cortina (Eds.), *Big Data at work: The data science revolution and organizational psychology* . Florence, KY: Taylor & Francis (Routledge).

Nielson, L.B. and Nelson, R.L. (2005). Rights Realized? An empirical analysis of employment discrimination litigation as a claiming system. *Wisconsin Law Review*, 2005, 663-711.

Nielson, L.B., Nelson, R.L., Lancaster, R., & Pedriana, N. (2008). *Contesting workplace discrimination in court: Characteristics and outcome of federal employment discrimination litigation 1987-2003*. Chicago, IL: American Bar Foundation.

Offerman, L.R. & Basford, T.E. (2014). Inclusive human resource management: Best practices and the changing role of human resources. In B.M. Ferdman & B.R. Deane, *Diversity at work: The practice of inclusion*. San Francisco, CA: Jossey-Bass.

*Oncale v. Sundowner Offshore Services, Inc.*, 523 U. S. 75 (1998).

Oswald, F.L. , Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P.E. (2015). Using the IAT to predict ethnic and racial discrimination: Small effect sizes of unknown societal significance. *Journal of Personality and Social Psychology*, 108, 562-571.

Pager, D. (2016). Are firms that discriminate more likely to go out of business? *Sociological Science*, 3, 849-859.

Rosen, E. (2017, April 3). Facebook pushes outside law firms on diversity. *New York Times*, B2.

- Schlanger, M. & Kim, P. (2014). The Equal Employment Opportunity Commission and structural reform of the American workplace. *Washington University Law Review*, 91, 1519-1590. Retrieved from [http://openscholarship.wustl.edu/law\\_lawreview/vol91/iss6/7](http://openscholarship.wustl.edu/law_lawreview/vol91/iss6/7).
- Selmi, M. (2003). The price of discrimination: The nature of class action employment discrimination litigation and its effects. *Texas Law Review*, 81, 1249-1335.
- Selmi, M. & Tsakos, S. (2015). Employment discrimination class actions after Wal-Mart v. Dukes. *Akron Law Review*, 48. Retrieved from <http://ideaexchange.uakron.edu/akronlawreview/vol48/iss4/4>
- SHRM (2014, April 8). SHRM Survey Findings: Diversity and Inclusion. Author. Retrieved from <https://www.shrm.org/hr-today/trends-and-forecasting/research-and-surveys/Pages/diversity-inclusion.aspx>.
- SHRM (2015, November 22). Managing equal employment opportunity. Author. Retrieved from <https://www.shrm.org/resourcesandtools/tools-and-samples/toolkits/pages/managingequalemploymentopportunity.aspx>
- SHRM (2016, August 30). EEO: General: What is the difference between EEO, affirmative action and diversity? Retrieved from [https://www.shrm.org/resourcesandtools/tools-and-samples/hr-qa/pages/cms\\_013810.aspx](https://www.shrm.org/resourcesandtools/tools-and-samples/hr-qa/pages/cms_013810.aspx).
- Silver-Greenberg, J. & Gebeloff, R. (2015, November 2). Arbitration everywhere, stacking the deck of justice. *The New York Times*. Retrieved from <http://www.nytimes.com/2015/11/01/business>.
- Stainback, K. & Tomaskovic-Devey, D. (2012). Documenting desegregation: Racial and gender segregation in private-sector employment since the Civil Rights Act. New York, NY: Russell Sage Foundation.
- Terpstra, D.E. & Honorée, A.L. (2016). Differences in the nature of employment discrimination litigation between private sector organizations and public sector organizations. *International Journal of Discrimination and the Law*, 16, 200-213.
- U.S. Office of Personnel Management (2015). Engaging the federal workforce: How to do it and how to prove it. Author. Retrieved from [https://admin.govexec.com/media/gbc/docs/pdfs/edit/engaging\\_the\\_federal\\_workforce\\_white\\_paper.pdf](https://admin.govexec.com/media/gbc/docs/pdfs/edit/engaging_the_federal_workforce_white_paper.pdf)
- Ulrich, D. (2016, March 8). The next “new thing”: How HR creates value from the outside in. Top Employee Institute & Human Resources Certification Institute. Retrieved from <https://www.hrci.org/>
- Vedantam, S. (2008, January 20). Most diversity training ineffective, study finds. *The Washington Post*. Retrieved from <http://www.washingtonpost.com/wp-dyn/content/article/2008/01/19/AR2008011901899.html>.

Walker, L., & Monahan, J. (1987). Social frameworks: A new use of social science in law. *Virginia Law Review*, 73, 559-598.

Wal-Mart v. Dukes 131 S. Ct. 2541 (2011)

**Chapter 4: Marking to Benchmark: Using Peer Comparison  
Standards to Measure Diversity and EEO Performance /  
*Pamela Coukos, CEO, Working IDEAL***

When and why does “what gets measured” turn into “what gets done?” There must be a process to understand and act upon the results of measurement, to know when or why a measure signals progress or concern or ambiguity and to make strategic or tactical decisions accordingly. Any performance measurement strategy or system requires a point of comparison to be meaningful. In the case of diversity metrics, there are still foundational questions about what to measure and how to measure it. But even when we choose a measure -- and a method -- we also need to choose a standard.

Constructing an EEO or diversity measure includes defining the standard for comparison. Options can range from exact parity to statistical deviation from expected outcomes to more aspirational approaches like high performance certification. These require certain tradeoffs. They also may reflect divergent measurement objectives. The proper point of comparison for imposing legal liability may be quite different than the one to use for setting internal performance goals. Regardless, establishing a proper comparison is an essential step in implementing any diversity measurement strategy.

This paper focuses on one possible standard – peer organization benchmarks - for core workplace pay and representation metrics for U.S.-based employers. The proposal is to use publicly available data to develop a free or low-cost tool based on broad cross-organizational comparisons for voluntary use. These accessible and transparent benchmarks could serve as a risk management tool and focus resources on the lowest performing organizations. These benchmarks could also serve as markers for industry leaders, conferring a competitive advantage

on higher performing organizations. Peer benchmarking is already a common workplace approach in areas ranging from occupational safety and health to salary offers, and could easily be adapted to support stronger diversity performance.

Peer organizations could include multiple axes of comparison. For example organization size might be an important marker of HR capacity, and would also account for differences in the scope of underlying data available to support the measures. Industry-based comparisons are an obvious opportunity for benchmarking, as indicators of potential similarity in the mix of jobs and types of employment practices. Public, private and nonprofit employer status is another, as well as regional, state, MSA, or similar geographic comparisons. Incorporating occupational measures would be helpful for employers seeking to set internal goals.

While some diversity benchmarking approaches already exist, they lack transparency, and have other barriers to wide implementation. For example, private certification standards or investment funds that screen based on diversity criteria may identify higher relative performance levels. However, disclosure of their criteria and underlying measures varies widely, there may be costs or other barriers to participation, and approaches may be limited to specific sectors.

With a public benchmarking tool, individual organizations could voluntarily test their data against the benchmarks and could use it to set their own specific internal diversity goals – much like typical affirmative action plans utilize Census and other availability data. Companies that score well are more likely to disclose their information to take advantage of the recruitment and brand-building benefits. Companies that score poorly may quietly engage in a more intensive assessment to identify risks and weaknesses, and should feel more pressure to undertake self-analysis and act on the results.



While far from a complete measurement strategy, establishing a free national database of peer benchmarks could confer broad social benefits. It would make it easier for a range of organizations to implement diversity metrics in the workplace. It would also provide external stakeholders useful information for engaging companies, universities, nonprofits and other organizations on their diversity performance. By identifying the highest performing and highest risk levels among a set of comparable employers, investors, regulators and consumer and worker advocates could use the information to highlight best practices and direct resources, pressure or technical assistance accordingly. Job seekers could incorporate this information into their decision making. Organizations hiring outside counsel could include this factor in their calculus. Stakeholder leverage would in turn create stronger incentives for employers to raise their relative performance level.

There are also legitimate questions and concerns with this approach. Since these measures prioritize comparability and ease of use they provide less information about any one organization's performance. The approach is an inadequate measure of legal compliance -- and in the case of extremely low performing industries could essentially grandfather in poorer performance as the benchmark value. It has potentially less salience for smaller organizations. Data limitations may make it difficult to measure diversity performance at the appropriate level of detail, such as analyses of individual race or ethnicity categories. Finally, high performance on these measures could easily obscure significant workplace concerns, while low performance could be inappropriately used to sanction.

However, as one element of a larger measurement strategy, peer benchmarking is a low-cost way to obtain relevant information across the employment spectrum. It can serve as a first step for

organizations that lack diversity program infrastructure. Coupled with appropriate quantitative and qualitative organization-specific metrics, diversity benchmarks enable cross-comparisons that employer-specific measures do not, and establish important baselines for measuring progress. And as the lowest performing organizations will have both the most incentive and potentially highest capacity to improve, over time the benchmark values should rise, creating upward pressure for high performers to continue improving to maintain their leadership. By focusing on leveraging voluntary private action and stakeholder engagement, a benchmarking tool like the one proposed here requires little new infrastructure and no specific mandates.

This paper serves as a thought experiment, framing the concept of a peer benchmarking database for further discussion and study. After reviewing the importance of establishing a point of comparison and reviewing alternatives, it proceeds to an examination of benchmarking and a sample approach, as well as potential benefits and drawbacks. Ultimately, it recommends using benchmarks that permit peer comparisons across a quartile distribution. Quartile measures allow employers to go beyond knowing whether they met, fell short of, or exceeded benchmark levels - and assess their performance relative to the higher and lower ends of the spectrum of comparable organizations.

Developing external standards to benchmark diversity and EEO performance is an obvious strategy for assessing progress against diversity measures, and one that is already in limited use. Expanding and formalizing this strategy will broaden its benefits.

## CONTEXTUALIZING EEO AND DIVERSITY OUTCOME MEASURES

Organizations create and implement workplace diversity programs for a range of reasons from symbolic to substantive.<sup>63</sup> When employers establish a strong and effective measurement strategy they are treating EEO outcomes more like core business objectives and less like a pure surface compliance or public relations exercise. This more explicit commitment to diversity and equity should yield more of the benefits of a fair and inclusive workplace, given the research that identifies measurement and accountability as particularly effective intervention.

Establishing specific performance metrics, and embedding those within transparent systems that promote meaningful accountability, can improve equal opportunity in the workplace. Measuring and reporting on progress helps interrupt common biases and in-group favoritism by making outcomes more visible.<sup>64</sup> There is good evidence that accountability and transparency can be more successful strategies than mandates or diversity training programs, and that collecting data and reviewing results is particularly critical.<sup>65</sup> Strong performance measurement programs balance benefits against reporting burdens, as well as privacy concerns, yielding information that helps managers and workers continue to improve workplace systems and practices, reduce bias, and build more inclusive and productive work environments.

---

<sup>63</sup> See generally Lauren B. Edelman, *Working Law: Courts, Corporations and Symbolic Civil Rights* (2016).

<sup>64</sup> Joel Nadler, et al, *Aversive Discrimination in Employment Interviews: Reducing Effects of Sexual Orientation Bias with Accountability*, *Psychology of Sexual Orientation and Gender Diversity* (2014); Patricia G. Devine, et al, *Long-term Reduction in Implicit Bias: A Prejudice-Breaking Intervention*, *Journal of Experimental Social Psychology* (2012); Christine Jolls and Cass Sunstein, *Debiasing Through Law*, *Journal of Legal Studies* (2006); Barbara Reskin, *The Proximate Cause of Employment Discrimination*, *Contemporary Sociology* (2000).

<sup>65</sup> Iris Bohnet, *What Works: Gender Equality by Design* (2016); Frank Dobbin and Alexandra Kalev, *Why Diversity Programs Fail, And What Works Better*, *Harvard Business Review* (July-August 2016); Frank Dobbin, Alexandra Kalev and Erin Kelly, *Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies*, *American Sociological Review* (2006); Michele E. A. Jayne and Robert L. Dipboye, *Leveraging Diversity to Improve Business Performance: Research Findings and Recommendations for Organizations*, *Human Resource Management* (Winter 2004).

While “the business case for diversity” can be overstated and sometimes oversimplified, an extensive literature across multiple disciplines supports the relationship between better EEO performance and business success. For example, several studies have linked gender diversity and increased innovation.<sup>66</sup> Research shows a link between diverse work teams and better outcomes, including improved decision-making, and higher productivity.<sup>67</sup> Building a diverse and talented workforce and a positive workplace culture can lead to quality improvements and higher efficiency and productivity.<sup>68</sup> Some evidence supports the view that diversity and other investments in human capital correlate with stronger financial performance.<sup>69</sup>

Poor performance, in turn, can increase risk. Worker exit drives unnecessary training and transition costs, including potential disruptions to production output and quality. Poor performance also increases the risk of public or private enforcement activity, including

---

<sup>66</sup> See, e.g., Cristina Díaz-García, Angela González-Moreno, Francisco Jose Sáez-Martínez, *Gender diversity within R&D teams: Its impact on radicalness of innovation*, *Innovation: Management, Policy & Practice* (2012) (study of R&D teams found a relationship between increased gender diversity and a greater likelihood of “radical innovation”); Sylvia Ann Hewlett, Melinda Marshall and Laura Sherbin, *How Diversity Can Drive Innovation*, *Harvard Business Review* (Dec. 2013), available at <https://hbr.org/2013/12/how-diversity-can-drive-innovation>.

<sup>67</sup> Adam Galinsky, et al, *Maximizing the Gains and Minimizing the Pains of Diversity: A Policy Perspective*, *Perspectives on Psychological Science* (2015), available at <http://groups.psych.northwestern.edu/spcl/documents/PerspectivesonPsychologicalScience-2015-Galinsky-742-8.pdf> (summarizing research); Katherine W. Phillips, *How Diversity Makes Us Smarter*, *Scientific American* (2014), available at <http://www.scientificamerican.com/article/how-diversity-makes-us-smarter/> (same); Vivian Hunt, Dennis Layton, Sarah Prince, *Diversity Matters*, McKinsey (2015), available at <http://www.mckinsey.com/business-functions/organization/our-insights/why-diversity-matters> (same); Lauren Foster, *The Business Case for Diversity*, CFA Institute (2015), available at <https://blogs.cfainstitute.org/investor/2015/06/30/women-in-investment-management-the-business-case-for-diversity/>; Catalyst, *Why Diversity Matters* (2013), available at [http://www.catalyst.org/system/files/why\\_diversity\\_matters\\_catalyst\\_0.pdf](http://www.catalyst.org/system/files/why_diversity_matters_catalyst_0.pdf).

<sup>68</sup> While some studies have found a more mixed or complex relationship between diversity and corporate success, substantial evidence supports the benefits of diversity for business, and the best current thinking is focused on how to manage diversity to maximize its positive impact. Galinsky, *supra* note 67; see also Michele E. A. Jayne and Robert L. Dipboye, *Leveraging Diversity to Improve Business Performance: Research Findings and Recommendations for Organizations*, *Human Resource Management* (Winter 2004), available at <http://www02.utm.edu/staff/mikem/documents/Diversity.pdf>; Iris Bohnet, *supra* note 65.

<sup>69</sup> See e.g., Niclas L. Erhardt, James D. Werbel and Charles B. Shrader, *Board of Directors Diversity and Firm Financial Performance*, *Corporate Governance, An International Review* (2003), available at [http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1005&context=management\\_pubs](http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1005&context=management_pubs); Catalyst, *supra* note 67 (summarizing studies showing link between gender board and senior management diversity and stronger firm financial performance).

potentially more likelihood of being targeted for a discrimination lawsuit or government compliance action, and less ability to successfully defend against them. Costs can go beyond direct costs of damages or other sanctions to include negative impacts on a corporate brand or workplace morale. And since recruiting a more diverse workforce is worth little if an employer cannot then retain that workforce, inclusion and equity must be sustained and meaningful commitments.

For these reasons, companies often affirmatively seek and promote reputations for diversity and inclusion when engaging with internal and external stakeholders -- to build their brand or expand the customer base as well as to recruit and retain employees and raise their productivity. Major employers seem to view diverse workplaces as a significant benefit in a global economy and an increasingly racially and ethnically diverse United States.<sup>70</sup>

Certain stakeholders have demonstrated clear interest in monitoring EEO performance and then using that information to make decisions. Workers are more likely to be aware of and to share their experiences and concerns, thanks to websites like Glassdoor, Hired, Payscale and many others, as well as more general social media networks. Investment fund researchers have studied gender diversity, attempting to target investments to companies with more women in leadership based on evidence of higher valuations, better financial results and/or less volatility.<sup>71</sup> And new

---

<sup>70</sup> The rising generation of U.S. workers are the most diverse in history. U.S. Bureau of the Census American Community Survey data, published in White House Council of Economic Advisers, *15 Economic Facts About Millennials*, (Oct. 2014), available at [https://www.whitehouse.gov/sites/default/files/docs/millennials\\_report.pdf](https://www.whitehouse.gov/sites/default/files/docs/millennials_report.pdf). A large number of major U.S. corporations identified business needs in the global economy and the critical importance of increased diversity in the highly skilled workforce, when asking the Supreme Court to preserve affirmative action remedies in education. *Grutter v. Bollinger*, 539 U.S. 306 (2003).

<sup>71</sup> Julie Dawson, Richard Kersley and Stefano Natella, *The CS Gender 3000: Women in Senior Management* (2014), Credit Suisse Research Institute, available at <https://www.calpers.ca.gov/docs/diversity-forum-credit-suisse-report-2015.pdf> (review of 3000 publicly held companies across the globe finds gender diversity on boards and in senior management is correlated with higher corporate valuations and stronger firm financial performance); Morgan Stanley, *Why It Pays to Invest in Gender Diversity* (May 11, 2016), available at

funds focusing specifically on gender equity suggest the potential for EEO and diversity performance to draw increased investment.<sup>72</sup>

Thus, internal and external stakeholders have some incentives to monitor diversity and EEO performance and reward or punish employers -- and employers have a variety of incentives to seek and promote their success and limit their risk of negative outcomes. This is akin to the case that been made for so-called “diversity report cards.” Scholars and advocates have promoted increasing mandatory corporate disclosures or utilizing government EEO data to identify outcomes at specific organizations.<sup>73</sup> Such a naming and shaming strategy leverages public pressure to generate performance improvements through some form of government intervention.

But even without a public disclosure risk or enforcement mandate, internal reporting mechanisms and performance management programs can still improve outcomes, to the extent they move employers from symbolic to substantive diversity work and leverage the beneficial effects of increased accountability. Given how much effort private companies already devote to measuring and benchmarking their performance in a range of core business functions, applying

---

<http://www.morganstanley.com/ideas/gender-diversity-investment-framework> (ranking of 1600 global stocks found those in the top third in terms of the percentage of women in the workforce generated higher equity returns); Linda-Eling Lee, et al, *Women on Boards: Global Trends in Gender Diversity on Corporate Boards*, MSCI ESG Research (November 2015), available at <https://www.msci.com/www/research-paper/research-insight-women-on/0263428390> (companies with stronger female leadership generated higher returns on equity and higher valuation).

<sup>72</sup> Michal Leibel, *Barclays Launches Women in Leadership Index and ETNs*, Reuters (July 9, 2014), available at <http://www.reuters.com/article/wealth-index-women-idUSL2N0PK29J20140710>; Bloomberg, *Bloomberg Financial Services Gender Equality Index*, available at [http://www.bbhub.io/professional/sites/4/BFGEI\\_Overview.pdf](http://www.bbhub.io/professional/sites/4/BFGEI_Overview.pdf); *State Street Global Advisors Launches Gender Diversity ETF to Help Investors Seek a Return on Gender Equity*, Business Wire, available at <http://www.businesswire.com/news/home/20160307005890/en/State-Street-Global-Advisors-Launches-Gender-Diversity>.

<sup>73</sup> Mary E. Graham and Julie Hotchkiss, *A More Proactive Approach to Addressing Gender-Related Employment Disparities in the United States*, *Gender in Management: An International Journal* (2009); Cyrus Mehri, Andrea Giampetro-Meyer and Michael B. Runnels, *One Nation, Indivisible: The Use of Diversity Report Cards to Promote Transparency, Accountability, and Workplace Fairness*, *Fordham Journal of Corporate and Financial Law* (2004), available at <http://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=1167&context=jcfl>; Note, *Should the SEC Expand Nonfinancial Disclosure Requirements?* *Harvard Law Review* (2002).

that framework to diversity and EEO metrics should be a worthy undertaking – provided an organization can develop the right metrics, measured in the right way, against the right standards.

### **Employment Terms and Activity Metrics**

Although workplaces can draw on a wide range of potential diversity measures, employers (and regulators) frequently use objective and quantifiable employment terms and activity metrics.

These can include applicants and hires by demographics, representation levels by demographics, pay by demographics and changes in representation over time (through hiring, promotion and termination). These measures are commonly reported or maintained on Human Resources Information Systems, especially because they are core elements of government reporting and recordkeeping requirements. Major labor market data sets also track these types of outcomes.

Most larger U.S. employers maintain and report employee representation metrics by demographics. For example, the U.S. Equal Employment Opportunity Commission (EEOC) requires private companies with 100 or more employees (and federal contractors with 50 or more) to submit regular EEO-1 reports with the count of workers by sex and race/ethnicity for each of ten broad occupational categories.<sup>74</sup> Corporate EEO-1 reports are based on individual establishments but also can be generated or aggregated at the companywide level. These reports are not public information, but some companies choose to disclose their EEO-1 reports, and EEOC at times releases aggregate data (based on industry, occupation or other categories).<sup>75</sup>

The U.S. Department of Labor’s Office of Federal Contract Compliance Programs requires entities that do business with the federal government holding government contracts above certain

---

<sup>74</sup> U.S. Equal Employment Opportunity Commission, EEO-1 report, <https://www.eeoc.gov/employers/eeo1survey/>.

<sup>75</sup> U.S. EEOC, Job Patterns for Women and Minorities in Private Industry, <https://www.eeoc.gov/eeoc/statistics/employment/jobpat-eeo1/>.

dollar thresholds to establish and implement an affirmative action program and an annual written Affirmative Action Plan.<sup>76</sup> These Plans include a summary listing of the count of workers by sex and race/ethnicity, as well as disability and veteran status, for each of the employer's established job groupings. OFCCP may periodically audit contractors and the agency regularly requests employment activity data by demographics, including applicants and hires, promotions and terminations information.

Through an updated version of the EEO-1 report, in the next year EEOC will require larger employers to provide summary data on pay by demographics. The EEOC's newly approved information collection (which supersedes an earlier OFCCP rulemaking) requires private employers with 100 or more employees (including federal contractors) to provide summary information based on W-2 wage data by gender and race/ethnicity using the 10 EEO-1 occupational categories.<sup>77</sup> Rather than specific pay amounts, the EEOC proposes reporting the number of workers within pay bands, as well as total hours worked. The agency anticipates beginning pay data collection with 2017 data, to be reported by March of 2018.

Major surveys conducted by the Bureau of Labor Statistics and the Bureau of the Census also collect data on many of these measures. The Current Population Survey and the American Community Survey ask individuals about income, occupation, industry, and demographics as well as employment status. While these surveys obtain information at the worker or household

---

<sup>76</sup> Executive Order 11246, requiring nondiscrimination on a range of grounds, and affirmative action on the basis of gender and race/ethnicity, Section 503 of the Rehabilitation Act of 1973 (Section 503) (requiring nondiscrimination and affirmative action on behalf of individuals with disabilities) and the Vietnam Era Veterans Readjustment Assistance Act of 1974 (VEVRAA) (requiring nondiscrimination and affirmative action on behalf of certain categories of veterans).

<sup>77</sup> U.S. EEOC, <https://www.eeoc.gov/employers/eeo1survey/2017survey.cfm>. For background on the OFCCP rulemaking, See Office of Federal Contract Compliance Programs, *Notice of Proposed Rule Making, Government Contractors, Requirement to Report Summary Data on Employee Compensation*, 79 F.R. 46562 (2014).



level, other federal and state data collection programs obtain information at the firm or employer level – and the Longitudinal Employer-Household Dynamics program combines several of these separate federal and state surveys.<sup>78</sup> It would be feasible to use this published data to generate summary measures of hiring and termination, workforce representation, and pay by demographics, aggregated across various factors, including occupation, industry and certain geographic breakdowns. Indeed, there are already examples in the research literature of these kind of industry-based averages or rankings.<sup>79</sup>

There are alternative approaches. Qualitative measures, attitudinal or survey data, metrics of timeliness or program maturity, counts of complaints lodged or resolved, government compliance activity, specific policy or program elements, and cost or efficiency measures can all provide important information on the strength, capacity and potential effectiveness of diversity programs. This proposal focuses on quantitative metrics based on employment activity data as one program element, an approach we could implement now at low cost with existing data. It does not discount the importance of measures of treatment, engagement, corporate culture, legal sanction, governance, inclusion, or process.

### **Compared to what?**

In addition to determining what aspects of diversity or EEO performance to capture, an EEO or diversity measure needs a point of comparison. Options can range from exact parity, to statistical deviation from expected outcomes, to more aspirational approaches like high performance

---

<sup>78</sup> U.S. Bureau of the Census, <https://lehd.ces.census.gov/>.

<sup>79</sup> Graham & Hotchkiss, *supra* note 73; Donald Tomaskovic-Devey, et al, *Private Sector Industry Disparities: A Report on Evidence of Systemic Disparities for Women, African Americans, Hispanics, Asians and Native Americans (2012)*, EEODataNet, <https://www.umass.edu/eeodatanet/private-sector-industry-disparities-report-evidence-systemic-disparities-women-african-americans>.

certification. All involve tradeoffs and reflect potentially divergent measurement objectives. The proper point of comparison for imposing legal liability may be quite different than the one to use for setting internal performance goals. Regardless, establishing a proper comparison is an essential step in implementing any diversity measurement strategy.

The simplest and most logical point of comparison is parity - equal outcomes for each group. Representation levels for all groups could be based on a population-level distribution like labor force participation or U.S. demographic representation. For example, under this type of approach, one would assume that women should hold approximately half of all management jobs and men should hold half of all clerical positions. This is a more radical framework for assessing diversity measures, and might even raise legal concerns to the extent it eliminated any consideration of availability, qualification or interest.<sup>80</sup> However, it is clearly an option to consider.

Typical economic or statistical measures, particularly in the single employer context, compare actual outcomes to those expected under a neutral process. An analysis of hiring shortfalls, pay differences or occupational segregation will yield a hypothetical distribution after accounting for relevant and appropriate factors or other refinements, and flag statistically significant departures as areas of potential concern that may pose a risk of legal liability. Such an analysis can be costly, and complex, requiring the use of experts and sufficiently robust data to generate useful and reliable comparisons of expected to actual outcomes. In addition, these comparisons are usually highly specific to a particular workforce or workplace and not always easily replicated from firm to firm. These are some simpler tools available, that compare applicants to hires and

---

<sup>80</sup> *Hazelwood School Dist. v. United States*, 433 U.S. 299 (1977).

determine whether any demographic group is being hired at a rate that differs from their representation in the applicant pool at a statistically significant rate. An even simpler rule of thumb applies a ratio analysis comparing one group's selection rate to that of the highest selected group, identifying concerns if that ratio is below 80 percent. In all of these cases, the expected outcome of a neutral process is the standard for comparison.

A similar but simpler mechanism relies on comparisons to local labor market representation (or internal labor markets) as the alternative (neutral process) outcome. Many affirmative action plans compare the distribution of workers by demographics to the qualified available pool, looking for shortfalls between actual and available qualified workers. Absolute numerical differences drive whether to set affirmative action goals, based on whether any underutilization exists. In addition, public or private enforcement can leverage these comparisons, challenging statistical deviations from availability measures, within certain legal limitations. This is essentially a benchmarking framework keyed to existing data, but limited to representation and some kinds of flow data, and based solely on departures from an absolute measure of comparison.

However, even though this form of benchmarking to local availability can be less costly and complex than multifactor statistical hypothesis testing against expected outcomes, it still requires some tailoring. Employers and regulators must employ technical and practical judgments on questions such as identifying the proper recruitment area, or the pool of relevant comparable workers, especially for more highly specialized occupations or highly under-represented groups. And the underlying benchmark values will "bake in" any existing structural discrimination, such

as occupational or housing segregation, access to educational or training opportunities, or other barriers to free and open competition.

In addition to estimating outcomes of a neutral process for a particular employer, or comparing outcomes to an aggregate measure of general availability, the standard for comparison could utilize a peer benchmark, like those developed in other contexts like workplace safety or productivity.<sup>81</sup> Many major employers rely on salary surveys to establish benchmark compensation levels for their positions. The goal is to identify the typical level of a position or job category that is paid by competitors, who are similar or peer employers, and then to adjust those levels based either on individual or firm level criteria. For example, firms may place entry level candidates within the range based on qualifications, or may decide to pay above benchmark in general - or for certain highly desirable candidates -- to beat the competition.

Using peer organizations for benchmark values, rather than generalized labor market availability data, would make the comparisons more tailored, even if they would still have some of the same limitations as the availability measures. It may be more meaningful to know that a particular company has a smaller gender gap than is typical for other employers of its size, or industry, than to compare to an overall average disparity level.<sup>82</sup> From an organizational perspective, one would expect peer organizations to operate under certain shared norms and congruent practices – and that among comparable organizations, any disparities would reflect at least to some extent

---

<sup>81</sup> See, e.g., Ron Z. Goetzel, et al, *Health and Productivity Management: Establishing Key Performance Measures, Benchmarks and Best Practices*, Journal of Occupational and Environmental Medicine (2001); Kathryn Means, Sean Whitaker, Rhona Flin, *Safety Climate, Safety Management Practice and Safety Performance in Offshore Environments*, Safety Science 2003.

<sup>82</sup> For an example of this framework using country-specific benchmarks, see Brian Levine, Min Park and Tom Jacob, *Driving Compensation Strategy Alignment: Using Analytics to Benchmark Practices from European Normative Data*, World at Work (2015).

common causes, shared networks or similar legal and cultural environments.<sup>83</sup> In that context, peer benchmarking would be more salient than the typical generalized workforce measures. The benchmark value measures the common level of disparity while the distance between that value and the organization's measure incorporates factors specific to that workplace.

There are a few existing examples of benchmarking to peer performance based on EEO-1 data. In the *Wal Mart* gender discrimination class action litigation, one of the plaintiffs' experts compared Wal-Mart's representation of women in management with similar large retail chains, opining that Wal-Mart's performance substantially lagged its competitors.<sup>84</sup> Assuming these companies all sought to fill similar jobs and drew from similar pools of available male and female workers, the implication was that something specific to Wal-Mart, rather than the qualifications, interest or availability of female workers, was driving the gender disparities – at least in part. Because the Supreme Court rejected the plaintiffs' request to certify the class, no court has tested the merits of this expert evidence.

One more potential basis for comparison is to use private certification standards that grade or evaluate organizations for a fee against a set of standards developed specifically for certifying higher performance levels. Those are sometimes based on existing peer performance and other times based on aspirational goals. EDGE is a global gender equity in the workplace certification standard that considers the level of women in leadership, pay equity and inclusive policies and

---

<sup>83</sup> Kevin Stainbeck and Donald Tomaskvic-Devey, *Documenting Desegregation: Racial and Gender Segregation in Private Sector Employment Since the Civil Rights Act* (2012); C. Elizabeth Hirsch, *Settling for Less? Organizational Determinants of Discrimination-Charge Outcomes*, *Law and Society Review* (2008); John J. Beggs, *The Institutional Environment: Implications for Race and Gender Inequality in the U.S. Labor Market*, *American Sociological Review* (1995); John R. Sutton, et al, *The Legalization of the Workplace*, *American Journal of Sociology* (1994); James Baron and William Bielby, *Bringing the Firms Back in: Stratification, Segmentation and the Organization of Work*, *American Sociological Review* (1980); Graham & Hotchkiss, *supra* note 73.

<sup>84</sup> Marc Bendick, *The Representation of Women in Management at Wal-Mart Stores* (2003), <http://www.walmartclass.com/staticdata/reports/r1.2.html>.

culture.<sup>85</sup> Publications such as Working Mother, Diversity, Inc. and Great Places to Work also rate companies on human capital, diversity and inclusion, identifying the highest performers in annual lists.<sup>86</sup> Generally these certification programs have limited transparency into the underlying ratings or scores and the databases used for comparison may be entirely proprietary.

An existing tool - the Census EEO Tabulation - provides a potential template for a public benchmarking database. The Census, through an interagency consortium that includes the Departments of Justice and Labor, the office of Personnel Management and the EEOC, provides look up tables that permit benchmarking to existing data, facilitating comparisons to the existing workforce based on demographics, age, occupation, and geographic location.<sup>87</sup> However, these measures only provide general measures of labor market availability. True peer performance benchmarking of diversity metrics requires access to proprietary data like individual EEO-1 affirmative action plans, so that one employer can directly compare performance to similar organizations. An alternative would be a database or tool that would permit simple lookups of hiring, pay and representation by demographics – with options to focus based on occupation, location and other strata of interest but also incorporating employer criteria (industry, size, etc.) Some organizations have created limited, proprietary versions of these data. With existing public data, one could generate comparisons of similar firms across occupations, geography and other strata of interest. That kind of tool could be available at low or no cost to all employers.

---

<sup>85</sup> EDGE Global Business Certification Standard for Gender Equality, see <http://www.edge-cert.org/our-impact/how-edge-creates-change-2/>.

<sup>86</sup> See, e.g. Diversity Inc. *Top 50 Companies for Diversity 2016*, <http://www.diversityinc.com/the-diversityinc-top-50-companies-for-diversity-2016/>; Working Mother Best Companies, <http://www.workingmother.com/best-companies>; Great Places to Work, *Fortune 100 Best Companies to Work For*, <https://clients.greatplacetowork.com/list-calendar/fortune-100-best-companies-to-work-for>.

<sup>87</sup> U.S. Bureau of the Census, <https://www.census.gov/people/eeotabulation/>.

Ultimately, diversity measures do not have much salience in isolation. Is a 12-cent gender pay gap doing well or doing poorly? How about when 8% of company managers are African-American? To focus on the relevance of these numbers in absolute terms, one might ask whether those differences are statistically significant departures from the pool of qualified and available or similarly situated workers. But there are questions that are more relative, such as where the employer is located and what kind of product or service it sells. It might also be important to know how those figures compare to diversity numbers from last year, or five years ago -- and whether these numbers being studied to assess compliance and liability risk or to develop outreach to millennial professionals. The basis for comparison ultimately drives whether that number represents progress or not.

While an employer can measure against absolute markers generated from average demographic values or certification standards, knowing how a measure ranks against benchmarks provided by peer organizations – ones that are more likely to have similar resources, opportunities, constraints and practices -- may be far more efficient and valuable.

### **Benchmarking in practice – options and alternatives**

At a basic level, a benchmark is simply a meaningful external comparison value. There are many kinds of benchmarks. Although they are usually grounded in some external real-world information, some are designed to reflect standard or average performance levels, while others are more aspirational in nature. Benchmarks can be absolute values or rely on a comparative framework. They can involve a single value or a distribution. And they can be established by the state in law or regulation, or through private action. Comparing three different examples of EEO benchmarking – the U.S. Department of Labor’s seven percent goal for the employment of

individuals with disabilities, the EEOC's interest in benchmarking of pay data for establishing enforcement priorities, and the "2020" initiative for increasing female representation on Boards of Directors illustrates the different ways benchmarking can be used for diversity measures and EEO goals.

In 2013, the Department of Labor's Office of Federal Contract Compliance Programs issued a regulation establishing a national 7% utilization goal for the employment of qualified individuals with disabilities.<sup>88</sup> This rule applies to companies that do business with the federal government who meet certain size and contract level thresholds. Contractors measure their progress through voluntary self-identification mechanisms and report the results in their affirmative action plans.

This "aspirational goal" is a diagnostic measure, not an enforcement threshold. Failure to meet a goal, or identifying job groups where qualified individuals with disabilities are under-represented compared to their availability to fill those jobs, does not violate any OFCCP regulation. Nor, as in the case of any OFCCP affirmative action goal based on race, sex or other criteria, does it trigger any mandate to hire or promote an individual belonging to a particular demographic group. Rather, it triggers an obligation to understand why certain groups are underutilized and to implement a specific strategy to address the situation. *See Associated Builders and Contractors v. Shiu*, 30 F.Supp. 3d 25 DDC (2014), *aff'd*, 636 F.3d 650 (D.C. Cir. 2014).

The regulation established this goal using existing public data. In this case, the agency used American Community Survey data on the national level of employment of individuals with disabilities across geographic and occupational differences. As the agency explained in its rulemaking, data limitations precluded establishing goals specific to geographic locations or

---

<sup>88</sup> 41 C.F.R. § 60-741.45.



occupation. Nevertheless, having a standard to evaluate a contractor's efforts in employing individuals with disabilities provides valuable insight into progress and barriers:

OFCCP recognizes that the 7 percent figure is less precise than the geographically specific availability information that contractors are familiar with under the Executive Order 11246 program, and that for some jobs in some locations availability of qualified individuals may be less than 7 percent. Furthermore, we recognize that the ACS data is based on a definition of disability that is narrower than that used under section 503. . . . While not perfect, the goal will provide a yardstick against which contractors will be able to measure the effectiveness of their equal employment opportunity efforts. It is our belief that the goal will enable contractors to think critically about their employment practices, including their outreach, recruitment, and retention efforts, and help them to assess whether and where any barriers to equal employment opportunity for individuals with disabilities remain. If barriers are identified, then the contractor can move to take corrective action.<sup>89</sup>

This approach is an example of using existing public data to establish a benchmark value to be used to measure performance, for voluntary assessment purposes. In this case the government established the benchmark, but only required contractors to assess their internal performance against that measure and take appropriate action in light of results. There was no requirement to achieve a specific level of performance.

Federal pay data collection proposals have explicitly incorporated a peer comparison approach for benchmarking EEO performance. In 2014, OFCCP proposed a rule to require federal

---

<sup>89</sup> Office of Federal Contract Compliance Programs, *Affirmative Action and Nondiscrimination Obligations of Contractors and Subcontractors Regarding Individuals with Disabilities*, 78 FR 58681, 59706 (2013).

contractors to report aggregate pay data. Now all larger employers, under the EEOC's revised EEO-1 report, will provide pay data to the two agencies. Employers will report pay data by sex and race/ethnicity, and by EEO-1 job group and establishment.

Throughout the rulemaking process, both agencies have proposed using peer performance comparisons to identify more likely violators for further investigation, as well as releasing aggregate measures to support voluntary compliance.<sup>90</sup> For example, the OFCCP's proposed regulation included a detailed discussion of using pay data reports (potentially combined with external public data sources) to generate "objective industry standards" – measures of the pay gap by industry and occupation. The EEOC explained that it would update its existing investigator tool to facilitate comparisons between a particular employer or establishment's gender or race-based pay disparities to those of similar employers in that labor market. Those comparisons could be based on both EEO-1 pay data reports and publicly available data. Both agencies also stated they would provide reports that aggregated data on gender and race-based pay disparities by industry, occupation and labor market, among others. Here again the government would provide the "benchmarks" but in this case would also use them for preliminary assessments of potential enforcement action. In addition employers could use the published aggregate data to assess their own performance and engage in voluntary compliance or estimate the risk of government enforcement.

---

<sup>90</sup> U.S. EEOC, *Agency Information Collection Activities; Notice of Submission For OMB Review, Final Comment Request: Revision of The Employer Information Report (EEO-1)*, 81 FR 45479, 45490-91 (2016); OFCCP, *Notice Of Proposed Rule Making, Government Contractors, Requirement To Report Summary Data On Employee Compensation*, 79 F.R. 46562, 46571-72 (2014).

The 30 Percent Coalition and Women on Boards 2020 are private campaigns that promote a benchmark number for women on corporate boards of directors of publicly traded companies through voluntary action. These nonprofit organizations have each established a numerical standard for female membership (30 percent overall or 20 percent by the year 2020 respectively).<sup>91</sup> In this case the benchmark sets a single aspirational goal to raise women's representation. These campaigns rely on stakeholder engagement and public pressure – including from institutional investors - to improve corporate diversity performance.

These different examples of EEO benchmarking include both government and private mechanisms, and a range of tools from purely voluntary to enforcement-driven. In two cases they involve single absolute measures, while the pay equity measures specifically contemplate a peer comparison framework. All of them provide some external yardstick to measure or assess progress by specific organizations. But peer comparisons provide specific benefits when thinking about potential explanations for variation in individual results. By considering how disparities compare across employers with similar characteristics, this kind of benchmark can reduce some of the potential sources of variation. It provides a more realistic potential basis for setting organization-specific goals.

### **Benchmarking voluntary diversity measures – leveraging the power of peer comparisons**

Regardless of the standard for comparing an outcome, there will be variation -- and in some cases substantial variation -- in how any individual employer performs. Typical disputes over EEO disparities in hiring rates, pay or other employment practices tend to turn on how to explain what is driving the difference -- and whether it can be assigned to an employer or employee

---

<sup>91</sup> <http://www.30percentcoalition.org/> and <https://www.2020wob.com/>.

responsibility. For purposes of enforcement it is important to assign fault or liability. Even in the context of voluntary diversity measures, understanding the basis for an outcome also matters, because it can suggest whether and how performance could be improved.

However, some factors that may affect differences in diversity outcomes can be hard to measure, while others are hard to isolate from potentially tainted effects. Common factors attributed to employees include skills and qualification, work effort, personal constraints and interests. On the employer side, location, position, tenure and business practices are major elements that may affect EEO and diversity outcomes. Statistical analysis at the firm level allows us to “control” for the effects of these factors, but data limitations, expense, and the challenge of properly modeling real-world behavior can leave gaps or create potential constraints.

Employee level factors are particularly challenging to measure independent of structural constraints. So, for example, pay gap measures constantly struggle with how much gender differences represent exogenous individual determinations as opposed to some biased societal or employer process. Are work hour differences between men and women purely a matter of individual choice – and are those choices affected by cultural expectations? Are women less likely to be offered longer hours because of stereotyped assumptions about what they will choose? Elements that could be labeled “choices” might be a complex mix of individual preferences and external constraints.

In addition to employee-level factors, there are also organizational characteristics that may affect disparities. Organizational scholars have addressed the impact of the larger economic, cultural and regulatory environments within which firms operate. These factors include industry-level practices and policies, organizational structures, legal interventions, corporate culture or

professional networks.<sup>92</sup> These factors can interact with more specific firm or worker-level effects, or can track across organizations to influence diversity and EEO outcomes.

If individual employers vary in their EEO performance, but have common structural or organizational characteristics that could align with firm or employee level differences, we can take advantage of this overlap. Benchmarking against peer organizations is essentially a mechanism that controls for common patterns of variation among similar entities. It establishes a typical level of common disparity across these organizations companies, isolating the remaining variation as more worthy of consideration.

Thinking back to the Wal-Mart example, large retail chains typically have a similar mix of jobs, and are likely to be hiring and promoting workers who are similar in terms of factors like education or other qualifications. The available pool of qualified workers based on gender, race or ethnicity should look broadly consistent across these employers. If some of these companies are more successful in terms of their representation of women in management, it may be worth looking to see if they have developed better practices, or removed barriers to advancement, or been more resistant to stereotype. Those who are underperforming in relative terms may be at a higher risk of finding systemic discrimination or at least substantial barriers to female advancement. While it is certainly plausible that all large retail chains are underperforming to some extent on this measure, considering who is performing better or worse in a relative sense is still highly significant information.

---

<sup>92</sup> Stainbeck and Tomaskovic-Devey, *supra* note 83; Graham and Hotchkiss, *supra* note 73; Beggs, *supra* note 83; Frank Dobbin, Soohan Kim and Alexandra Kalev, *You Can't Always Get What You Want: Organizational Determinants of Diversity Programs*, American Sociological Review (2011); Susan Bisom-Rapp, *Bulletproofing the Workplace: Symbol and Substance in Employment Discrimination Law*, Florida State University Law Review (1999); Lauren B. Edelman, et al, *Internal Dispute Resolution: The Transformation of Rights in the Workplace*, Law and Society Review (1993).

Benchmarking seems particularly useful as a relative measure of risk. Employers who are doing substantially worse than their peers on these diversity measures might be more vulnerable to litigation, government enforcement, or other adverse events. Organizations with larger gender- or race/ethnicity-based disparities compared to peers are less likely to find those gaps explained entirely by the common factors used to isolate potential discrimination from legitimate explanations. To the extent that these factors are correlated with peer employer characteristics (size, industry, or location for example), their explanatory power in this particular case would be correspondingly reduced. These larger gaps also mean potentially more workers could be aware of and raise discrimination concerns.

### **Using benchmark distributions – A sample matrix**

This paper proposes a basic matrix of relative performance and relative risk. Much like a salary survey, the benchmarks would include a median performance level and measures of the 25<sup>th</sup> and 75<sup>th</sup> percentile. Benchmark data could come from large existing data sources, private or proprietary data sets, non-public government data, surveys, or some combination of these sources.

Going beyond simple comparisons to single peer organization benchmarks, and creating even a limited distribution like the quartile measures, would have additional benefits. An organization could do a preliminary risk assessment knowing not just what the average value is, but where they fall relative to the highest and lowest performers. A company in the bottom quartile for its industry or size might be at a higher relative risk compared with a company in the top quartile, and would have a basic indicator suggesting the need to do a more extensive self-analysis. A top quartile company might be able to promote itself as an industry leader, especially where better

diversity performance can be tied to workplace practices, corporate culture or exemplary leadership.

Using publicly available data, it is feasible to create a lookup tool or a set of tables that display the median, 25<sup>th</sup> and 75<sup>th</sup> percentile measures for hiring and termination, representation and pay by demographics – and then to further differentiate these benchmarks by industry, location or other strata. Below is a sample concept for this kind of table.

**Table 1: Sample Benchmark Matrix**

Top 25%	High Performance
25-50%	Above Median Performance
50-75%	Above Median Risk
Bottom 25%	High Risk

Developing industry-based benchmarks seems especially useful, given the potential for similarity in terms of the occupational mix and employment practices. The occupational mix for management positions in healthcare is likely more similar across individual firms, and more distinct from the occupational mix for management positions in IT, education, hospitality or manufacturing. Factors that drive compensation decisions, and the relative mix of compensation components – such as bonuses, overtime, commission shift pay or piecework – should also align to some degree with industry. Larger labor market trends, like levels of education, occupational differences or rates of part time work by demographics, should operate in a broadly similar

fashion to EEO outcomes at peer firms who are drawing from the same pools of workers.

Workplace disparities vary across industries and industry is an important factor in explaining divergent outcomes.

Other key organizational characteristics could serve as markers of similarity. Organization size would track to some extent HR capacity and the relative amounts of data available for analysis. Benchmarking by MSA would take local labor market differences into account – something that may matter for certain industries and occupations more than others – while benchmarking by state addresses potential differences in the relevant legal standards or regulatory environment. Occupational benchmarks allow for more specific comparisons of similar workers. In cases where sufficient data exists to cross-tab these values, the benchmarks could incorporate that option – like occupation by location, or industry by size. As always, data limitations may foreclose the ability to benchmark at a particular level – or for a particular worker demographic, as in the example of the disability employment goal.

**Table 2: Sample Benchmark Matrix by Industry**

Finance	Healthcare	Retail	Legal Services	
Top 25%	Top 25%	Top 25%	Top 25%	High Performance
25-50%	25-50%	25-50%	25-50%	Higher than Median Performance
50-75%	50-75%	50-75%	50-75%	Higher than Median Risk
Bottom 25%	Bottom 25%	Bottom 25%	Bottom 25%	High Risk



**Table 3: Sample Benchmark Matrix by Employer Size/Type**

Private firm 500 or more	Private firm 100 or more	Non-profit organization 100 or more	Public employer 100 or more	
Top 25%	Top 25%	Top 25%	Top 25%	High Performance
25-50%	25-50%	25-50%	25-50%	Above Median Performance
50-75%	50-75%	50-75%	50-75%	Above Median Risk
Bottom 25%	Bottom 25%	Bottom 25%	Bottom 25%	High Risk

**BENEFITS AND LIMITATIONS of Benchmark Data**

Even those these measures are relatively simplistic, the ability to compare across firms yields important benefits. First and foremost, these benchmarks provide an easy way for individual organizations to measure themselves and set organization-specific internal goals. Existing laws and regulations either require companies to implement regular self- analysis, or create strong risk

management incentives to do so.<sup>93</sup> However, progress remains uneven. Even employers who seek in good faith to comply with legal mandates and social norms face challenges in integrating regular self-analysis into their EEO compliance functions – including data limitations and a lack of sufficient in-house technical HR capacity. Social scientists who study the workplace have found that formal legal requirements may not matter as much as the structures, systems and practices companies establish to comply with those rules – and without strong and substantive programs those formal rules may not in practice lead to meaningful diversity improvements.<sup>94</sup> Making a basic tool available at low or no cost could fill some of these gaps.

Much like the Census EEO tab facilitates the efforts of federal contractors to set their goals for affirmative action plans, companies could use the benchmark values to establish their own internal goals for representation and pay, and assess progress over time. Providing an external industry benchmark that is most relevant to an individual employer, and even tailored to factors like location or occupation, makes it easier to set goals that are meaningful targets. Studies of federal contractors have identified a relationship between affirmative action programs -- which require companies to establish written plans, review data, set goals and monitor progress -- and progress in the workplace for women and workers of color.<sup>95</sup>

Because the benchmarks include distributional information, rather than the absolute values contained in the EEO tab, organizations could also use their position vis a vis peers to determine

---

<sup>93</sup> Covered federal contractors must include regular self-analysis of employment practices by race and gender as part of their EEO programs, *see* 41 C.F.R. §60-2.17, and all employers are potentially subject to public or private enforcement actions under federal or state laws banning discrimination.

<sup>94</sup> Edelman, *supra* note 63.

<sup>95</sup> *See, e.g.,* Fidan Ana Kurtulus, *Affirmative Action and the Occupational Advancement of Women and Minorities 1973-2003*, *Industrial Relations* (2012); Jonathan S. Leonard, *The Impact of Affirmative Action on Employment*, *Journal of Labor Economics* (1984).

relative risk. Those organizations with lower performance relative to peers might have more vulnerability and could be less confident that a detailed analysis with factors substantially explains the disparity. Those with higher performance could try to leverage their advantage in terms of recruitment and outreach or brand-building.

And although this is an entirely voluntary mechanism, external stakeholders can create opportunities and incentives for action. Organized stakeholders, including institutional investors, consumer and worker advocates, local elected officials, and the general public, can exert leverage on employers to disclose their diversity performance and to meet or exceed benchmark values. Investor pressure<sup>96</sup> and union engagement<sup>97</sup> as well as an increasing culture of transparency are already increasing voluntary disclosure. According to the site Open Diversity Data, more than a dozen major tech employers including Apple, Cisco and Google have voluntarily released EEO-1 data, and more have provided other disclosures of their diversity data and measures.<sup>98</sup> A few major employers have even released the results of their pay equity audits: for example Amazon, Microsoft, Salesforce and the Gap have publicly disclosed their findings and plans.<sup>99</sup> The potential for disclosure in turn increase the incentives to effectively measure

---

<sup>96</sup> Lisa Hayles, Boston Common Asset Management Comment to U.S. EEOC on Proposed Revision of the Employer Information (EEO-1) Report to Include Collection of Pay Data, available at <https://www.regulations.gov/document?D=EEOC-2016-0002-0240>; Susan Baker and Brianna Murphy, Trillium Asset Management Comment to U.S. EEOC on Proposed Revision of the Employer Information (EEO-1) Report to Include Collection of Pay Data, available at <http://www.trilliuminvest.com/wp-content/uploads/2016/03/EEOC-Comment-Letter-3.9.16.pdf>; Katie Johnson, *She's Pressing Top Companies on Pay Equity*, Boston Globe (May 21, 2016), available at <https://www.bostonglobe.com/business/2016/05/20/she-pressing-top-companies-pay-equity/tA0XUQep7QCRGj6NTG82pL/story.html>.

<sup>97</sup> Alexander C. Kaufman and Emily Peck, *Wall Street Journal Vows to Fix Pay Gap for Women and Minorities*, Huffington Post (March 24, 2016), [http://www.huffingtonpost.com/entry/wall-street-journal-pay-gap\\_us\\_56f44629e4b0143a9b47bc4d](http://www.huffingtonpost.com/entry/wall-street-journal-pay-gap_us_56f44629e4b0143a9b47bc4d).

<sup>98</sup><http://opendiversitydata.org/>.

<sup>99</sup> Diversity at Amazon, [https://www.amazon.com/b/ref=tb\\_surl\\_diversity/?node=10080092011](https://www.amazon.com/b/ref=tb_surl_diversity/?node=10080092011); Salesforce, *Equality at Salesforce: The Equal Pay Assessment Update* (March 8, 2016), available at <https://www.salesforce.com/blog/2016/03/equality-at-salesforce-equal-pay.html>; Cora Lewis, *These Companies are Eliminating Their Gender Pay Gaps*, BuzzFeed (March 14, 2016), available at

diversity performance and to address shortcomings – as well as creating opportunities to use good performance affirmatively. Regulators could presumably also use benchmarks as the EEOC proposes to strategically target enforcement resources based on EEO-1 data.

The most common objection to using benchmark values to judge specific outcomes in the workplace is that aggregate data may obscure important explanations for individual variation, and unfairly penalize a firm for disparities with legitimate explanations. Certainly any individual's outcome is a function of a broad range of personal characteristics, choices and constraints, along with organizational and broader economic factors. But these individual-level differences are not always relevant to explain gender- or race-based disparities. Measuring benchmarks at an industry level effectively limits the impact of those explanations to the extent they reflect common industry practices or worker characteristics. And they can isolate firm-specific differences for further study.

Another important concern is that this could validate a poor performance level of an entire industry or accept the impact of a problematic practice. For example, the finance industry has been the subject of numerous systemic discrimination lawsuits. It also features particularly high race and gender disparities in compensation and advancement. Being a top quartile finance firm may not exonerate that employer from a close examination of its practices and EEO vulnerabilities.

---

[https://www.buzzfeed.com/coralewis/companies-are-eliminating-their-gender-pay-gaps?utm\\_term=.ek115WEXv#.nkGVy6reK](https://www.buzzfeed.com/coralewis/companies-are-eliminating-their-gender-pay-gaps?utm_term=.ek115WEXv#.nkGVy6reK).

Even with these challenges, peer benchmarks could allow employers seeking to make preliminary low or no-cost assessments of their diversity performance when compared with their peers.

**Part II: Innovations in Metrics and Organizational Performance**

**Chapter 5: How effective are organizational diversity initiatives?**  
**An illustrative evaluation of the National Football League's Rooney Rule /**  
*Presented by Christopher Rider, Assistant Professor*  
*of Strategy at Georgetown University*

Christopher I. Rider  
Georgetown University  
[chris.rider@georgetown.edu](mailto:chris.rider@georgetown.edu)

James B. Wade  
George Washington University  
[jbwade@gwu.edu](mailto:jbwade@gwu.edu)

Anand Swaminathan  
Emory University  
[anand.swaminathan@emory.edu](mailto:anand.swaminathan@emory.edu)

Andreas Schwa  
Iowa State University  
[aschwab@iastate.edu](mailto:aschwab@iastate.edu)

July 23, 2017

\* Prepared for the Research Group on Legal Diversity's 2016 Conference on Metrics, Diversity, and Law. We gratefully acknowledge the invaluable research assistance of Kim Wheeler, Keith Harrison, and Scott Bukstein, scholars at the Univ. of Central Florida's Institute for Diversity & Ethics and at the Research & Preservation Center of the Pro Football Hall of Fame

## **How effective are organizational diversity initiatives? An illustrative evaluation of the National Football League's Rooney Rule.**

### **Abstract**

Many organizations implement diversity initiatives to align employee demography with workforce demography. We propose that typical implementations render it difficult, if not impossible, to evaluate the effectiveness of such initiatives and highlight two specific statistical challenges: (1) diversifying candidate pools and (2) small numbers of targeted positions. To support our argument, we analyze demographic data for all National Football League coaches from 1985 to 2015. Our analyses focus on a league-wide policy – the “Rooney Rule” – that was implemented in 2003 to increase racial minority representation at the head coach level. Probing the continuing debate of the rule’s effectiveness, we demonstrate how initiative evaluations necessitate clearly-stated, ex ante expectations absent intervention (i.e., counterfactuals) and outcomes that are insensitive to single events (i.e., small numbers). Aiming to influence diversity management best practices, we offer initiative design and implementation guidelines that enable fairly straightforward evaluations of effectiveness.



An organization's members are often less demographically diverse than are residents of the communities in which it operates (e.g., McPherson and Smith-Lovin, 1987; Reskin, McBrier and Kmec, 1999). Because organizational employment practices shape inequalities across demographic groups (Baron and Bielby, 1980, Castilla, 2011), organizations that wish to adhere to prevalent norms of equal employment opportunity often implement diversity initiatives intended to align organizational demography with workforce demography (e.g., Dobbin, Kim, and Kalev, 2011). Yet, despite implementations of numerous and various initiatives across many kinds of organizations, diversity initiative effectiveness is often empirically ambiguous and, relatedly, popularly debated (e.g., Dobbin, Schrage, and Kalev, 2015).

Ideologically-diverse populations are bound to debate the need for organizational diversity initiatives. But, even those who share beliefs about the need for organizational interventions find it difficult to agree on what is effective or just how effective interventions are. For example, recent research suggests that bureaucratic personnel practices highlighted by Weber (1978) and advocated by prominent sociologists (e.g., Edelman, 1990; Bielby, 2000; Reskin, 2000) may not alleviate – much less eliminate – organizational influences on inequality (e.g., Kalev, Dobbin, and Kelly, 2006; Castilla, 2011; Kalev, 2014; Rissing and Castilla, 2014). And while research on diversity initiatives that do or do not “work” accumulates (e.g., Dobbin, et al., 2015; Apfelbaum, Stephens, and Reagans, 2016), skepticism about their effectiveness persists (e.g., Bregman, 2012; Stephens and Apfelbaum, 2016).

Despite such continuing debate and absent clear evidence of effectiveness, some diversity initiatives are nonetheless diffusing as organizations seek to adopt best practices. In particular,

the National Football League’s “Rooney Rule” is regularly highlighted as a model diversity initiative that should be adopted in various organizations, professions, and industries. The Rooney Rule is a hiring process intervention implemented in 2003 to increase racial diversity in organizational leadership positions. The Rule basically ensures a minimum level of minority representation among head coaching candidates by requiring franchises to interview at least one racial minority for every such vacancy. Similar initiatives have been adopted for jobs in academia (e.g., the University of Texas System), in technology (e.g., Facebook, Pinterest) and in municipal employment (e.g., City of Portland, Oregon) while still others call for implementation in Hollywood (e.g., Verrier, 2016), Silicon Valley (e.g., Alba, 2015), English football (e.g., BBC, 2016), Corporate America (e.g., Palmeri and Womack, 2013) and U.S. Congressional staff members (O’Keefe, 2017). Yet, even the Rule’s supporters debate whether or not it continues to work – or ever worked. Paul Tagliabue – the former National Football League (NFL) Commissioner who oversaw the rule’s implementation – publicly stated “I don’t think the Rooney Rule has done as much as anyone hoped it would” (Florio, 2016).

We propose that the Rooney Rule’s effectiveness continues to be debated because its implementation is representative of many organizational diversity initiatives. In particular, evaluations of the Rooney Rule’s effectiveness are clouded by several statistical challenges. Typically, such initiatives aim to increase the representation of underrepresented groups in a particular position (or set of positions) by an unstated amount over an indeterminate period of time by uniformly applying a policy to all hiring decisions. Uniform applications render it difficult to estimate what would have been observed in the absence of the initiative’s implementation. Furthermore, failing to clearly state specific objectives and a timeframe for

achieving them is bound to invite disagreement on whether progress has been made or is being made fast enough.

We focus on two statistical challenges that complicate effectiveness evaluations. First, the U.S. workforce has diversified considerably over the past few decades and continues to do so. For example, the Center for American Progress projects that by 2050 there will be no racial or ethnic majority in the United States.<sup>100</sup> If the pool of potential candidates for positions targeted by an initiative is diversifying over time then one might expect higher-level positions to diversify in the future independently of any organizational interventions. Despite diversifying pipelines, diversity initiatives are rarely (if ever) implemented with a clearly-stated expectation of what would be observed absent implementation or an explicit plan for estimating such “What if...?” scenarios (i.e., statistical counterfactuals). Further complicating evaluation is the tendency of an initiative’s advocates to claim impact for the entire before-after difference in the metric of interest (e.g., female representation increased from 15% to 25% after implementation). In such cases, at least part of the claimed difference is probably attributable to a diversifying workforce.

Second, diversity initiatives often focus on an organization’s most visible and highest-authority positions which, in most organizational hierarchies, are few in number. Consequently, change and representation figures are heavily influenced by single hiring events that exert disproportionate influence on small denominators. For example, consider an organization with five Vice-Presidents. In this organization, increasing the number of female Vice-Presidents from

---

<sup>100</sup> <https://www.americanprogress.org/issues/economy/reports/2012/07/12/11938/the-state-of-diversity-in-todays-workforce/>

one to two increases the number of female Vice-Presidents by 100% and female representation from 20% to 40%. Of course, symmetric decreases result from replacing a female Vice-President with a male. In such cases, the law of small numbers (i.e., generalizing broadly from a small number of observations) exerts excessive influence on evaluations of initiative effectiveness.

Evaluations of the NFL's Rooney Rule are clearly complicated by these factors. First, the rule was implemented to ensure that racial minorities would be considered for head coaching positions but no specific objective or metric was stated at implementation. Second, assistant NFL coaches were becoming more and more diverse (and gaining experience) in the decade before the Rooney Rule's implementation. Third, at any point in time only 32 people hold an NFL head coach position, with a fraction of those positions turning over in a typical year (ranging from 7% to 37% per year with a mean of 21% and a median of 22%). This context, therefore, enables us to demonstrate the challenges of evaluating initiative effectiveness by analyzing the evolving demography of NFL head coaches and assistant coaches between 1985 and 2015. Drawing upon academic studies as well as popular press articles, we marshal illustrative claims and quotes from the Rooney Rule's proponents and opponents that highlight the difficulty of producing unequivocal evidence of initiative effectiveness (or ineffectiveness).

We conduct a program evaluation of the Rooney Rule's impact on minority representation among NFL head coaches using basic statistical techniques. Our guiding question is "Did the Rooney Rule cause minority representation among NFL head coaches to increase after its implementation in 2003?" Our quantitative analysis of this initiative establishes several facts. First, representation of racial minorities at the head coach level increased from 6% in 2002 to

19% in 2015 (and a high of 25% in 2011). Second, the pool of assistant coaches most likely to be promoted to the head coach position increased steadily from 9% minority representation in 1985 to 30% in 2002 and to 32% in 2015. Third, evaluations of the rule's effectiveness hinge on one's beliefs about what the expected representation trend would have been between 2003 and 2015 if the rule had not been implemented; we demonstrate how slight changes in statistical assumptions substantially affect whether the Rooney Rule is considered successful or not.

Despite controversy surrounding its implementation, ongoing debate about its effectiveness, and our inconclusive evaluation, the Rooney Rule is widely considered an organizational best practice for diversity management. With the goal of influencing best practices, we offer design and implementation guidelines for organizations seeking to implement diversity initiatives that produce credible evidence of what does and does not work. We demonstrate how the ongoing debate about the Rule's effectiveness might have been avoided if the NFL had followed these guidelines when introducing this rule in 2003. As more and more organizations adopt initiatives similar to the Rooney Rule, it is our hope that our guidelines will help organizations achieve greater returns on their investments in diversity.

### **Diversity in Context: the NFL's Rooney Rule**

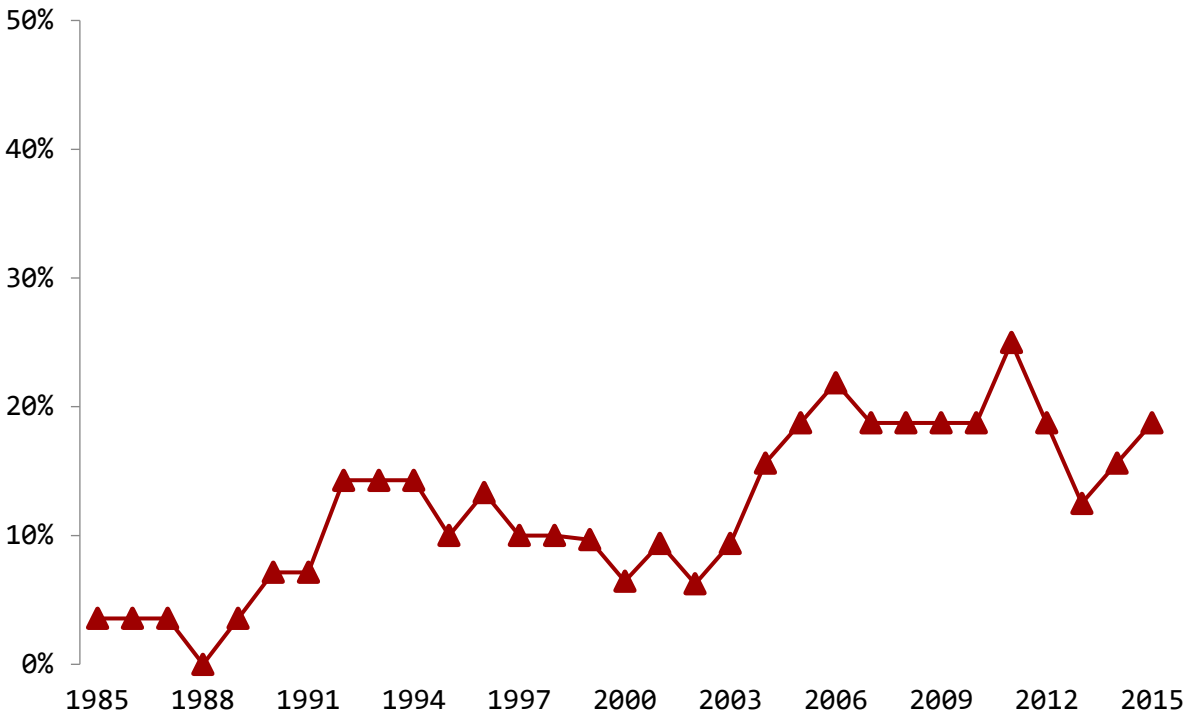
The National Football League consists of 32 American football franchises (i.e., teams) distributed across the continental United States and organized in two conferences with three divisions each. The highest authority position on a typical NFL franchise coaching staff is the head coach. During our sample period the league expanded from 28 to 32 teams, creating four additional head coaching positions between 1995 and 2002. The next highest positions are two

coordinators tasked with coaching offensive and defensive players, respectively. Several lower-level position coaches are overseen by the offensive and defensive coordinators. These assistants are responsible for coaching position players like quarterbacks, linebackers, wide receivers, and defensive backs.

Most coaches begin their NFL coaching careers as a position coach (or an assistant to a position coach). They can be promoted to higher-level positions by their current employer or by being hired by another franchise. Franchise owners and/or personnel staff typically hire head coaches, who then hire coordinators and assistant coaches. More details on promotions within NFL coaching hierarchies can be found in Rider, et al.'s (2017) study, which examines racial differences in promotion rates among NFL coaches using extensive coaching career history data.

Fritz Pollard, an African American football player and coach, was the first racial minority to serve as a head coach of a professional football franchise. Pollard coached the Akron Pros of the American Professional Football Association (a precursor to the NFL), as well as NFL teams in Milwaukee and Indiana, in the 1920s before the NFL became segregated in 1926. In 1979, the Oakland Raiders hired Tom Flores, a Hispanic man, to be the first racial minority head coach since Fritz Pollard. Flores served as Raiders head coach until 1987. It was not until 1989 that a second African-American was hired as a head coach, when the (then) Los Angeles Raiders hired Art Shell. Flores was again hired as the Seattle Seahawks head coach in 1992. Figure 1 depicts the time trend in racial minority representation among NFL head coaches between 1985 and 2015, when the percentage of racial minorities at the head coach level increased from 3.6% to 18.8%.

**Figure 1. Minority representation at the head coach level**



The hiring of head coaches changed dramatically between the 2002 and 2003 seasons, when the Rooney Rule was implemented. It mandates that at least one racial minority candidate must be interviewed before an NFL franchise hires a head coach. The Rooney Rule is named after Pittsburgh Steelers owner Dan Rooney, who chaired the league-wide committee tasked with addressing diversity issues. The rule's implementation was largely motivated by the 2002 firings of two black head coaches, Tony Dungy and Dennis Green, when Dungy had a winning record and Green had only one losing season in ten years. Attorneys Cyrus Mehri and Johnnie Cochran Jr., as well as former NFL player John Wooten, led a campaign that resulted in the voluntary adoption of the Rooney Rule by all NFL franchises on December 20, 2002. Duru (2011) details the substantial and continuing efforts to institute and enforce the Rooney Rule.

Proponents of the Rooney Rule typically argue that hiring biases make it difficult for racial minority coaches to attain head coaching positions in the NFL. Some empirical studies have supported this position. For example, one study of pre-2002 coaching performance found that African American head coaches earned more wins and were more likely to qualify for the playoffs between 1990 and 2002 than were white coaches (Madden, 2004). This evidence is consistent with minorities being held to a higher standard than whites in head coach hiring decisions. Revisiting the analysis and including coaches fired mid-season, a follow-up study affirmed this result and maintained the position that prior to the Rooney Rule minority coaches had to be better coaches to obtain the same position as white coaches (Madden and Ruther, 2009). Another follow-up study found that this racial disparity was no longer observable after 2002, leading the authors to conclude that the Rooney Rule “likely eliminated” racial disparity in head coach hiring (Madden and Ruther, 2011: 140).

Even before such analyses were technically feasible, others deemed the Rooney Rule an “uncharted success” because it forced decision makers to expand their searches to candidates that they would not have otherwise considered (Collins, 2007: 870). Consistent with this position, Figure 1 illustrates that the highest minority representation figure prior to the Rooney Rule’s implementation was 14.3% but that in only one post-implementation season did representation fall below that pre-implementation maximum (12.5% in 2013). Difference-in-means tests confirm that the representation of minority head coaches is significantly higher and that of white coaches lower after the Rooney Rule’s implementation (pre-RR mean of 8.1% versus post-RR mean of 17.8%;  $p < 0.01$ ). Proponents point to this statistically significant, post-intervention increase in minority representation as evidence of the Rule’s effectiveness.



Despite these observations, support for the Rooney Rule is not universal and has not remained consistent among its supporters over time. One study of NFL coordinators between 1970 and 2009 found no discernible difference in promotion rates to head coach, conditional on a coach attaining a coordinator position (Solow, Solow, and Walker, 2011). The authors concluded that there was no empirical basis for inferring that the Rooney Rule “increased the number of minority head coaches” (Solow, et al., 2011: 332). A more recent and extensive analysis reproduces the finding that there was no racial disparity in promotion rates to head coach, conditional on attaining a coordinator position, but that there is a persistent disparity in promotion rates to coordinator that is observed between 1985 and 2015 – both before and after the Rooney Rule (Rider, et al., 2017).

Others who initially deemed the Rooney Rule a “spectacular success” later claimed that within a decade the rule had “lost whatever effectiveness it once had” while also acknowledging the difficulty of distinguishing single-year statistical anomalies from a trend reversal (Burke, 2013). Such contrasting claims are almost inevitable given the small number of head coaches in the NFL. Substantial changes in the percentage of minority representation can occur when one or two minority head coaches are replaced by white coaches or vice-versa. In short, there is considerable debate concerning the Rooney Rule’s effectiveness in increasing minority representation among NFL head coaches.

It is important to note the various statistical approaches used to address the question of whether or not the Rooney Rule has been effective at achieving its aim. Some compare the performance of minority and white coaches before and after the Rooney Rule’s implementation (Madden,

2004; Madden and Ruther, 2009). Others compare promotion rates (Solow et al., 2011; Rider, et al. 2017). To the best of our knowledge, no studies explicitly compare the actual time trend in representation figures to baseline figures implied by either the Rooney Rule's advocates or by statistical estimation. We are, therefore, the first to do so and our analysis is intended to make it apparent why such comparisons are not often made.

Because the Rooney Rule applies to all NFL franchises, it is feasible to evaluate its effectiveness at increasing minority representation among all NFL head coaches. It would not, however, be feasible to evaluate its effectiveness at increasing representation at any single franchise over the 12-year post-implementation period, given that each franchise employs a single head coach and most coaches hold the position for multiple years.<sup>101</sup> This is a statistical challenge of implementing a diversity initiative that applies to a very limited number of positions and especially those occupied by a single individual (e.g., head coach, Chief Executive Officer, President). An analogous situation is a diversity initiative implemented across a large group of employers (e.g., a university system, a municipality, a profession). Below, we evaluate the Rooney Rule in a way that demonstrates the difficulty of answering a fairly straightforward question: “Did the Rooney Rule increase minority representation among NFL head coaches?”

### **Evaluating the Rooney Rule's Effect on Minority Representation**

We start by describing an ideal – “ideal” from an econometric perspective – diversity initiative implementation. The econometric principles underlying this ideal are detailed in Angrist and

---

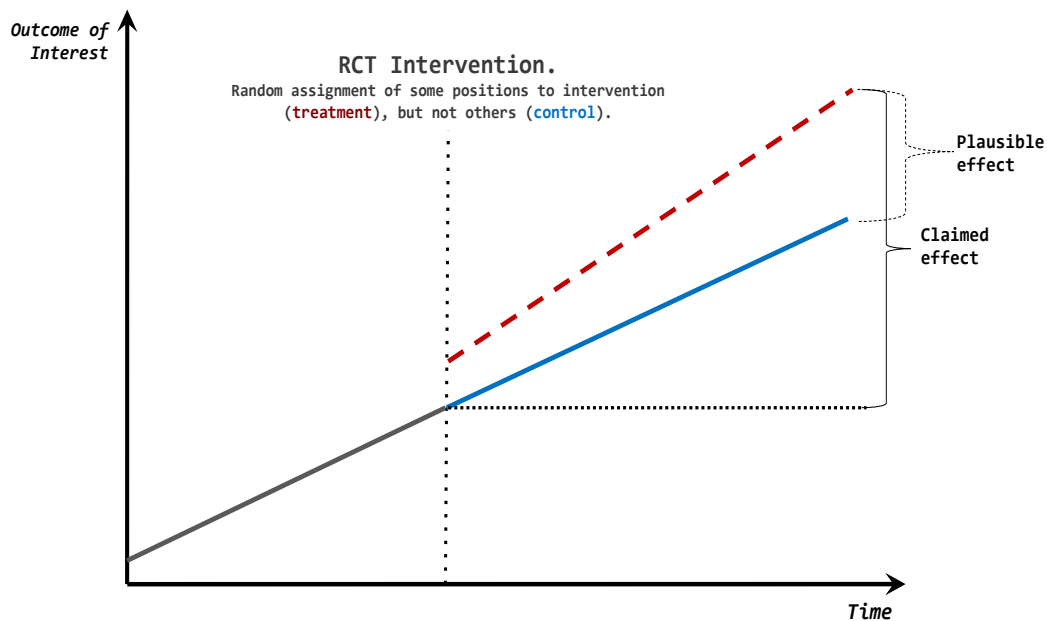
<sup>101</sup> In such an analysis it would not be possible to control for time-invariant omitted factors like market size, tradition, ownership, etc. Moreover, some franchises exhibit no variance in minority representation at the head coach level.

Pischke (2009: 221-247). In a nutshell, in our view the ideal implementation is a randomized controlled trial (RCT) in which some arbitrarily-determined sub-set of positions was subject to organizational intervention while another sub-set of observationally-equivalent positions was not. The second sub-set provides a plausible estimate of what is known as the counterfactual scenario (i.e., what would have happened absent the intervention) by serving as a “control” group. Conversely, positions subject to the intervention serve as the “treatment” group. If the treatment group positions are randomly-determined, then one can typically treat the outcome of interest among the control group as a valid comparison for those positions subject to the intervention. A fairly straightforward estimate of the initiative’s effect is, therefore, the difference in the mean outcome of interest between the two groups.

A proper evaluation of the Rooney Rule depends on one’s expectations of how representation would have evolved absent the Rooney Rule’s implementation. To illustrate our ideal scenario, consider an alternative implementation of the Rooney Rule in which only half of NFL head coach positions would be subject to the Rooney Rule after 2003 (i.e., the “treatment group”) while the other half (e.g., the “control group”) would not be subject to the Rule’s requirements. To satisfy the observationally-equivalent criteria, the positions subject to the rule would need to be unbeknownst to NFL franchises prior to the decision to dismiss their current coach. Post-implementation, we could estimate the Rooney Rule’s “treatment effect” on minority representation by computing the difference in representation between the treatment and control groups.

This ideal scenario mirrors the general scenario depicted in Figure 2. As this figure depicts, the difference in before-implementation and after-implementation representation figures represents the effect typically claimed by an initiative’s advocates (i.e., the claimed effect). For example, Rooney Rule proponents might claim that the change in interviewing policy increased minority representation among head coaches from 6.3% to 18.8% between 2002 and 2015. But, such estimates are valid only if representation would not have increased absent the initiative’s implementation. In contrast, the RCT design enables one to estimate the initiative’s causal effect as the difference between Figure 2’s dashed red and solid blue lines at the end of the evaluation period (i.e., the plausible effect). This plausible effect is substantially more modest than the effect typically claimed (i.e., the difference between the dashed red and dotted gray lines).

**Figure 2. Intervention effects: Typically-claimed versus plausibly-caused**



We consider this “difference-in-differences” represented by the plausible effect in Figure 2 an econometrically credible estimate of a diversity initiative’s effect for several reasons. First, because of random assignment there is no ex ante reason to believe that the treatment (dashed

red) and control groups (solid blue) differed substantively prior to implementation. This assumption can and should, of course, be validated by comparing the two groups on observable characteristics including the outcome of interest prior to implementation. Second, if the treatment group positions and control group positions are subject to the same rules prior to implementation then one can assume (again subject to validation) that they followed the same pre-implementation trend in the outcome of interest. Third, all treatment organizations would comply wholly with the rule. Subject to satisfying these three assumptions, this ideal implementation enables one to estimate the diversity initiative's causal effect on the outcome of interest as the difference between the mean treatment group outcome and the mean control group outcome in post-implementation time periods (i.e., the plausible effect represented in Figure 2).

Actual implementation of the Rooney Rule diverged substantially from this ideal intervention. There were no randomly-determined treatment and control groups. Rather, the Rooney Rule was applied uniformly to all NFL franchises simultaneously between the 2002 and 2003 seasons and, moreover, franchises did not uniformly adhere to the Rooney Rule. For example, the Detroit Lions were penalized for ignoring the rule entirely before hiring Steve Mariucci and the Dallas Cowboys merely granted Dennis Green a telephone interview prior to hiring Bill Parcells as head coach. The statistical upshot is that one can only evaluate the Rooney Rule by assuming a counterfactual that is not estimated according to standard econometric techniques (e.g., differences-in-differences estimation).

It is also not possible to evaluate the Rooney Rule's effect based on initiative objectives or expectations because neither was made explicit at the time of implementation. Cyrus Mehri

states that “When we created the Rooney Rule in 2002-03, what we were trying to do was give guys who otherwise got overlooked a fighting chance” (Wilner, 2017). Such objectives are not easily translated to statistical measurement. It is consequently challenging to estimate the solid blue line and much easier to estimate the dashed red and dotted grey lines depicted in Figure 2. It is not surprising, then, that the difference between dashed red and dotted grey is often claimed as the causal effect of a typical diversity initiative. For example, Mehri noted in 2017 that he was “very happy” that “now we are at the highest number of clubs” employing minority head coaches (Wilner, 2017), essentially pointing to the maximum point of the dashed red line in Figure 2.

We make several attempts to estimate the solid blue line in Figure 2 in order to estimate a plausible causal effect of the Rooney Rule on minority representation at the head coach level. To do so, we follow two approaches to estimating a statistical counterfactual. The first approach assumes that in the absence of the Rooney Rule minority representation would have continued to increase at the same rate (e.g., percentage per year) as it increased prior to implementation. This possibility is most reasonable if one believes increasing diversity is simply a matter of time needed to change norms, alleviate bias, or otherwise address the underlying cause of the racial gap in head coach representation.

The second approach assumes that minority representation would have increased proportionately with minority representation among the candidates for head coaching positions. This possibility is most reasonable if one believes that increasing diversity is simply a matter of the pipeline of qualified coaching candidates diversifying. In other words, one might expect head coaches to be as racially diverse as coordinators and assistant coaches.

Both approaches require some critical (and debatable) assumptions – the first about the functional form of the representation-as-a-function-of-time relationship and the second about the definition of candidates “at-risk” of promotion to head coach. For our purposes, it is less important that our analyses are framed by the most justifiable set of assumptions than it is that we demonstrate how sensitive the evaluation is to slight changes in those assumptions. We, therefore, make fairly simple assumptions sufficient to serve this purpose. Our first approach is informed by the assumption that minority representation increases linearly with time and the second approach is informed by the qualifications of the “typical” NFL head coach hire. Below, we describe this definition after first describing the data and methods for our empirical analysis.

## **Data**

We use a cross-sectional time-series dataset consisting of all individual coaches – head coaches, coordinators, and assistants – employed by NFL franchises between 1985 and 2015. The observation period was determined by the availability of NFL annual directories that provided detailed coaching staff information (Official National Football League Records and Fact Book, 1985-2012). These sources specify all coaches for the upcoming season. These data are also used in Rider, et al. (2017), where its collection is described extensively.

Based on job titles, we assigned all NFL coaches to four hierarchical levels common to most NFL coaching staffs: (1) head coaches, (2) assistant head coaches and coordinators, (3) position coaches (e.g., linebackers, receivers, special teams), and (4) assistants to position coaches and quality control coaches. Our evaluation of the Rooney Rule focuses primarily on minority

representation at the first level (i.e., head coach). In some analyses, we consider a sub-set of coaches at the lower levels to be “at-risk” of being promoted to head coach.

In sum, our data set consists of 14,741 coach-years for 1,780 coaches, which represents the full population of NFL coaches between 1985 and 2015. For some analyses, we drop the lower-level coaches to restrict our sample to the candidate pool of coaches “at-risk” of being hired as head coaches. We define the at-risk set based on an analysis of all head coach hires observed in our data. Descriptive statistics reveal that approximately 82% of NFL head coach hiring events involve first-time head coaches and that, furthermore, 84% of first-time head coaches have more than six years of NFL coaching experience. We, therefore, defined the at-risk candidate pool as all NFL coordinators and assistant coaches with at least seven years of NFL coaching experience but no NFL head coaching experience. These restrictions yield 7,253 coach-years and 901 coaches.

## **Method**

We first consider statistics describing the time trend in minority representation among NFL head coaches, before and after the Rooney Rule’s implementation between the 2002 and 2003 NFL seasons. We then consider what the most appropriate comparison is for this time trend, which informs a series of “counterfactuals” estimated by statistical analysis. Note that we focus on percentage representation as the outcome of interest, as opposed to minority head coach counts, because NFL expansion events in 1995, 1999, and 2002 added four NFL head coaching jobs.



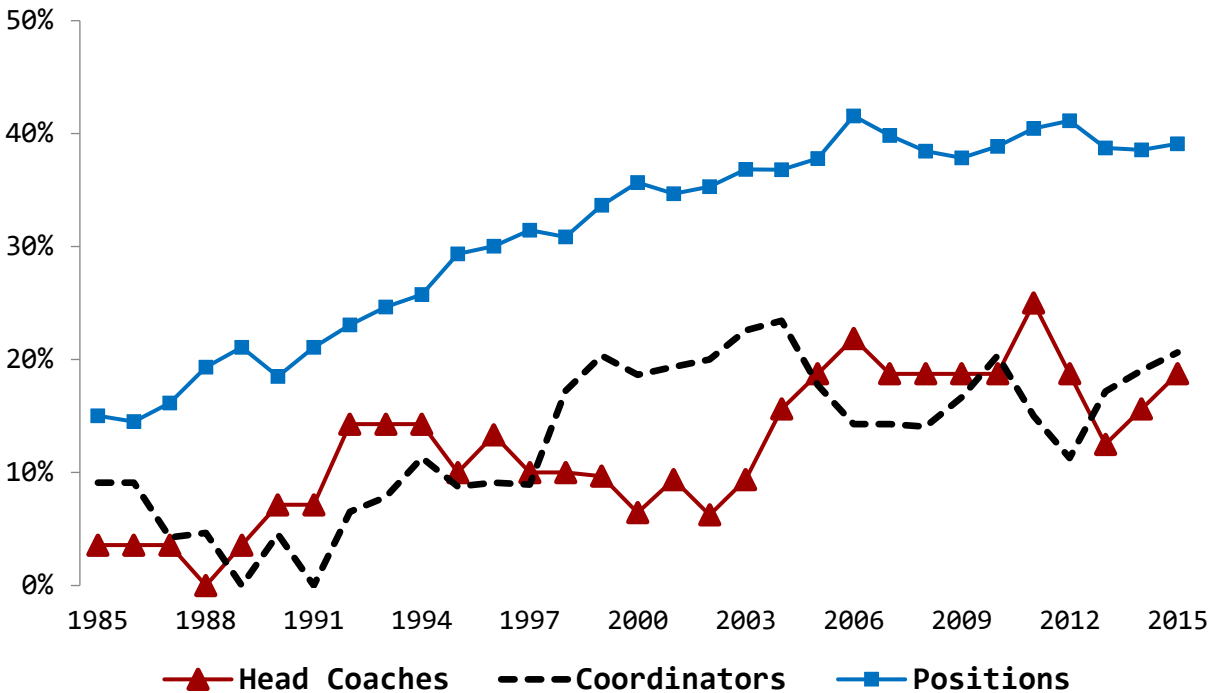
As Figure 1 depicts, minority representation among NFL head coaches increased significantly from 8.1% during the years from 1985 to 2002 to 17.8% between the years from 2003 to 2015 ( $p < 0.01$ ). But, such before-after comparisons of minority representation are most informative if the composition of the candidate pool is fairly constant throughout the observation period. As we demonstrate below, though, it was not.

Figure 3 illustrates that before 2003 minority representation among coordinators exceeded head coach representation but also that this relationship reversed after the Rooney Rule's implementation.<sup>102</sup> Many observers treat this reversal as evidence that the Rooney Rule increased the representation of minority coaches in the head coaching ranks by eliminating a bottleneck in the minority coaching pipeline (e.g., Freeman, 2013; Fox, 2015). Another analysis of coordinator promotion rates, however, concluded that there was "...little support for the proposition that...the Rooney Rule has successfully increased the number of minority coaches" (Solow, et al., 2011: 333). Moreover, initial enthusiasm for the Rule diminished in the eyes of many who pointed to the flattening of the upward, post-Rule trend. For example, one observer noted that by 2006 the Rule was a "spectacular success" but that by 2013 it "lost whatever effectiveness it once had" (Burke, 2013).

---

<sup>102</sup> Madden and Ruther (2010) as well as DuBois (2015) also document this reversal in the trend of the numbers of black coaches and coordinators in the NFL.

**Figure 3. Minority representation at head, coordinator, and assistant coach levels.**



As noted above, these observers base their comments on different comparisons (e.g., coaching performance, promotion rates of coordinators, before/after representation figures). In other words, various observers implicitly invoke different counterfactuals in their assessments of the Rooney Rule to proxy for the solid blue line in Figure 2. An important factor neglected in nearly all of these assessments is the fact that, as Figure 3 depicts, minority coach representation at the head coach, coordinator, and position coach levels was increasing prior to 2003. So, at least some of the representation gains depicted in Figure 1 might have been achieved in the absence of the Rooney Rule.

Motivated by these possibilities, we estimate two associated counterfactual trends to serve as comparisons to the observed representation trend. Two keys to estimating a credible counterfactual are (1) projecting the pre-implementation time trend from 2003 to 2015 based on

observations in 1985 through 2002 and (2) identifying which of the coordinators and position assistant coaches are reasonably at-risk of being hired as NFL head coaches.

First, we consider the possibility that the minority representation trend simply increased linearly with time as it had prior to the Rooney Rule’s implementation. To estimate this “stay the course” counterfactual, we estimated an ordinary-least-squares regression in which the dependent variable is the observed representation figure for minority head coaches from 1985 to 2002 and the single explanatory variable is calendar year. We present summary statistics for these variables in Table 1 and regression results in Model 1 of Table 2. As Table 1 indicates, the outcome of interest is positively and significantly correlated with year ( $p < 0.05$ ).

**Table 1. Summary statistics and correlations of variables in regression analyses.**

		Mean	St. Dev.	Min	Max	(1)	(2)
(1)	% minority rep. among NFL HCs	0.08	0.04	0.00	0.14	--	
(2)	Year	1993.5	5.3	1985	2002	0.48	--
(3)	% minority rep. in at-risk pool	0.19	0.07	0.09	0.30	0.37	0.98

**Table 2. OLS regressions of minority representation among NFL head coaches.**

	(1)	(2)	(3)	(4)
Year	0.004 *		0.007 **	
	(0.002)		(0.002)	
% minority representation in at-risk pool		0.242 †		0.563 **
		(0.131)		(0.158)
Constant	-7.58 *	0.036	-14.7 **	-0.012 **
	(3.02)	(0.028)	(3.25)	(0.029)
R-squared	0.229	0.137	0.505	0.383
F-statistic	4.75 (1, 16)	3.39 (1, 16)	20.6 (1, 13)	12.7 (1, 13)
Years included	1985 to 2002	1985 to 2002	1985 to 1999	1985 to 1999

Robust standard errors in parentheses.

\*\*  $p < 0.01$ ; \*  $p < 0.05$ ; †  $p < 0.10$ ; two-tailed tests.

The regression of 18 annual NFL observations between 1985 and 2002 summarized in Model 1 of Table 2 indicates that minority representation was increasing linearly at a rate of 0.4% per year between 1985 and 2002. Based on this figure, we generated annual predicted values and then extrapolated the predictions through 2015. Figure 4 depicts these predicted values (solid blue lines), bound by 95 percent confidence intervals (dashed black lines), along with the observed values for each year between 1985 and 2015 (solid black circles).

**Figure 4. Rooney Rule evaluation with counterfactual based on linear time trend as of 2002.**

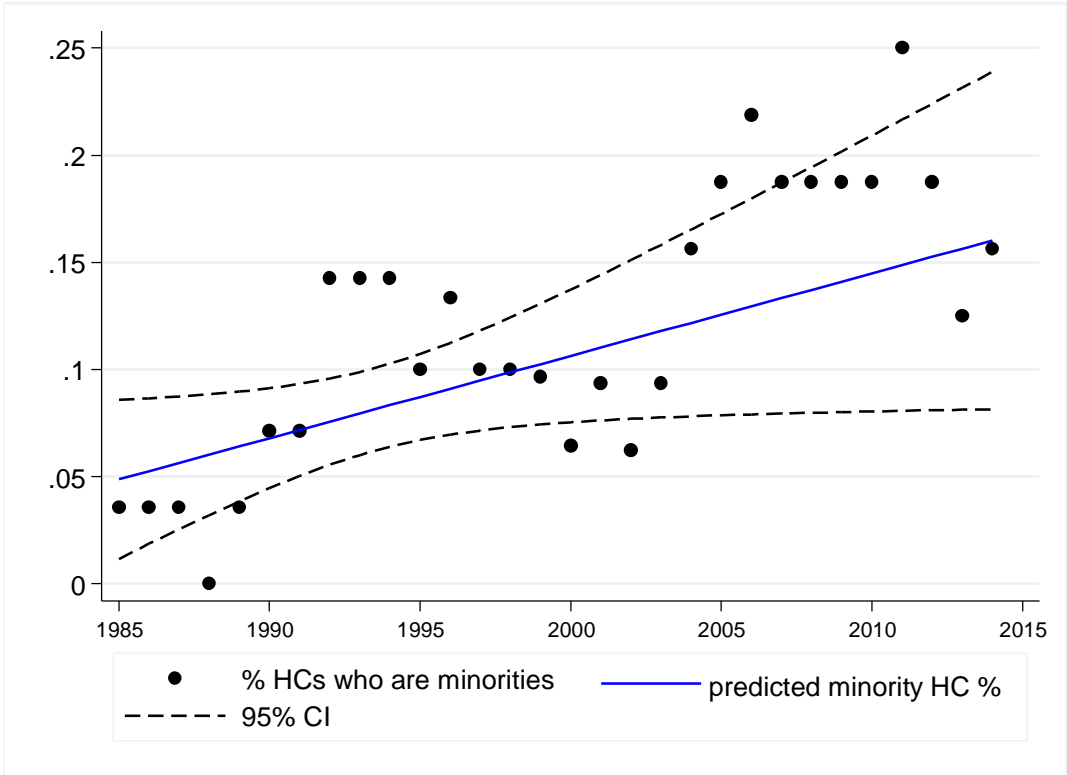
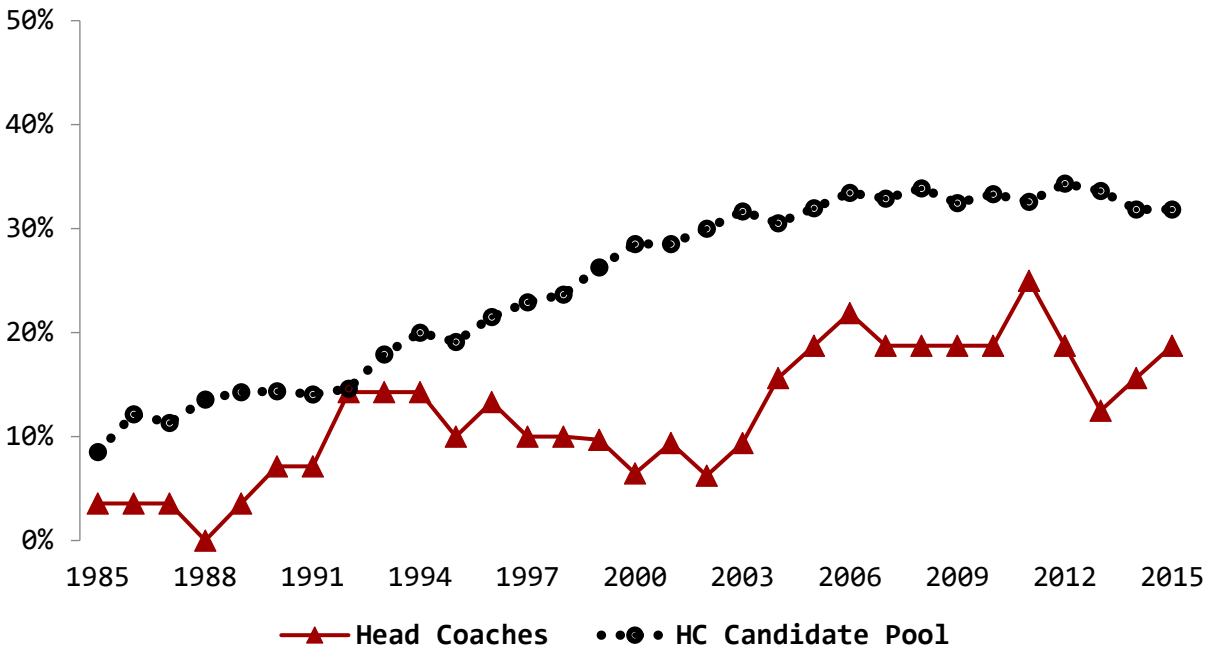


Figure 4 demonstrates that between 2003 and 2015, nine of the twelve observed values are above the predicted value and only three observed values are below that line. This is consistent with representation outpacing the pre-implementation trend after the Rooney Rule’s implementation. In other words, observed minority representation among NFL head coaches was greater than

expected given our observations of the minority head coach representation trend from 1985 to 2002. Proponents of the Rooney Rule can interpret these results as supportive of the position that the Rooney Rule increased minority representation among NFL head coaches.

Second, we consider the possibility that the minority representation trend increased not with calendar year but, rather, with the composition of the candidate pipeline. To do this, we must characterize the candidate pool of assistant coaches “at-risk” of promotion to head coach. Based on the descriptive statistics described previously, we defined the candidate pool as all NFL coordinators and assistant coaches with at least seven years of NFL coaching experience but no NFL head coaching experience. We depict the minority time trend for this candidate pool in Figure 5, which basically represents a combination of the coordinator and position groups depicted in Figure 3. As seen in Figure 5, minority representation in the head coach candidate pool increased steadily from 9% in 1985 to 32% in 2015. Furthermore, as Table 1 indicates, the outcome of interest is positively but not significantly correlated with this at-risk pool variable.

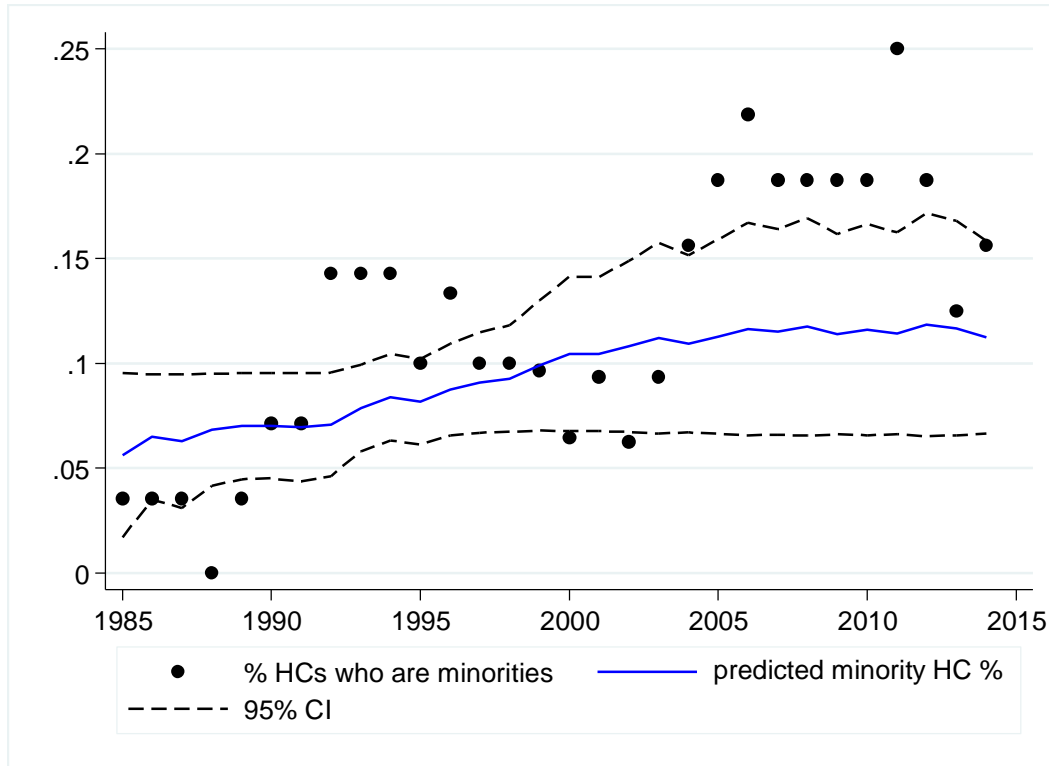
**Figure 5. Minority representation at head coach level versus candidate pool.**



To estimate a counterfactual trend based on this “diversifying pipeline” we estimated an ordinary-least-squares regression in which the dependent variable is the observed representation figure for minority head coaches from 1985 to 2002 and the single explanatory variable is the minority representation figure for the head-coach candidate pool. The results of this regression are presented in Model 2 of Table 2. One can interpret the coefficient as follows: every 1% increase in minority representation within the head coach candidate pool is associated with an increase of approximately 0.2% in minority representation at the head coach level. We again generated annual predicted values based on observations between 1985 and 2002. We then extrapolated these predictions through 2015. Figure 6 depicts these predicted values (solid blue lines), bound by 95 percent confidence intervals (dashed black lines), along with the observed values for each year between 1985 and 2015 (solid black circles). This is an even more favorable

estimate of the Rooney Rule's effect on minority representation than the estimate depicted in Figure 4.

**Figure 6. Rooney Rule evaluation with counterfactual based on diversifying candidate pool.**



In Figure 6, all but one of the post-implementation figures is above the predicted time trend represented by the solid blue line. Moreover, nine of the 13 observed values between 2003 and 2015 are above the 95% confidence interval obtained from the regression analysis. This analysis indicates that the observed minority representation among NFL head coaches after 2002 was greater than we might have expected given the composition of the head coach candidate pool.

Thus far, we have produced two counterfactual estimates consistent with the position that the Rooney Rule increased minority representation among NFL head coaches. We now seek to

demonstrate how fragile this support is by changing the assumptions generating the counterfactuals. To do so, we simply generate predicted values based on observations from 1985 to 1999 instead of 1985 to 2002. Why? Admittedly, this choice is somewhat arbitrary. There was no change in interviewing policy implemented in 2000. Minority representation among the candidate pool did, however, reach its highest point to-date at 25%. We do not, however, attach any particular importance to the year 2000. We change our estimation simply to demonstrate how sensitive our evaluation of the Rooney Rule is to subjective statistical assumptions.

We estimated an ordinary-least-squares regression in which the dependent variable is the observed representation figure for minority head coaches from 1985 to 1999 and the single explanatory variable is calendar year. The results of this regression are presented in Model 3 of Table 2. These results indicate that minority representation was increasing linearly at a rate of 0.7% per year between 1985 and 1999. This figure is notably higher than the 0.4% estimate implied by Model 1. Importantly, the regression presented in Model 3 produces a much better model fit than the regression summarized in Model 1 (as indicated by the higher variance explained, as indicated by the R-squared figure and the F-statistic).

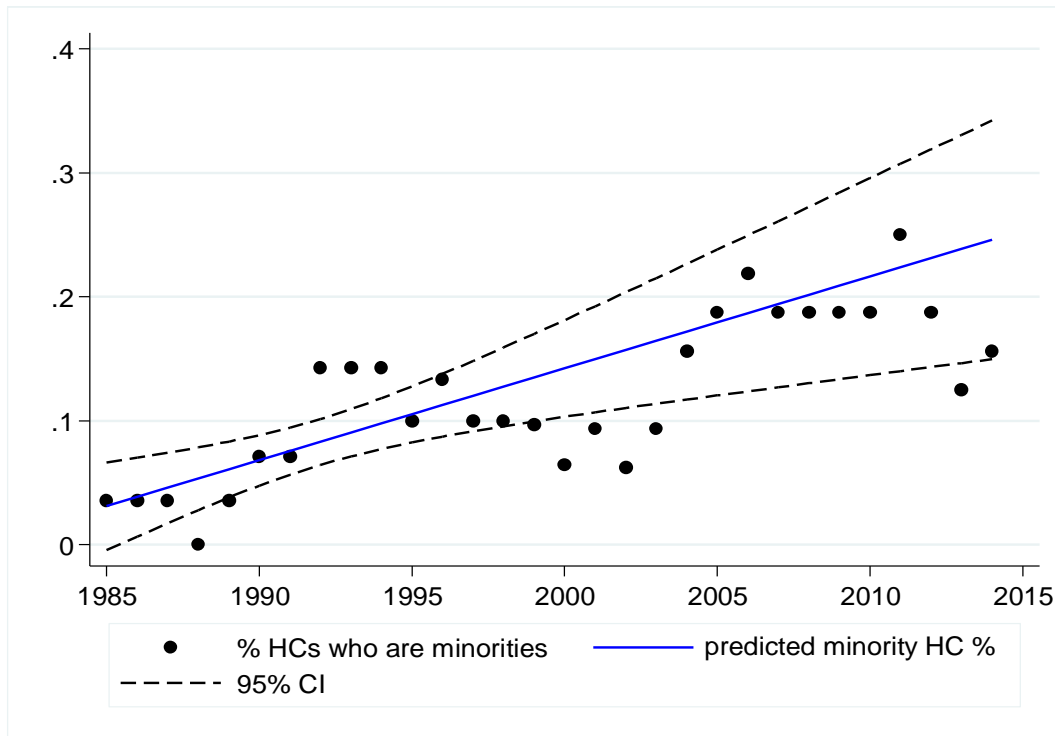
We also estimated an ordinary-least-squares regression in which the dependent variable is the observed representation for minority head coaches from 1985 to 1999 and the single explanatory variable is the minority representation in the head coach candidate pool. The results are presented in Model 4 of Table 2. Model 4 also produces a better model fit than the regression summarized in Model 2 (also indicated by the higher R-squared and F statistics). One can interpret the coefficient as follows: every 1% increase in minority representation within the head



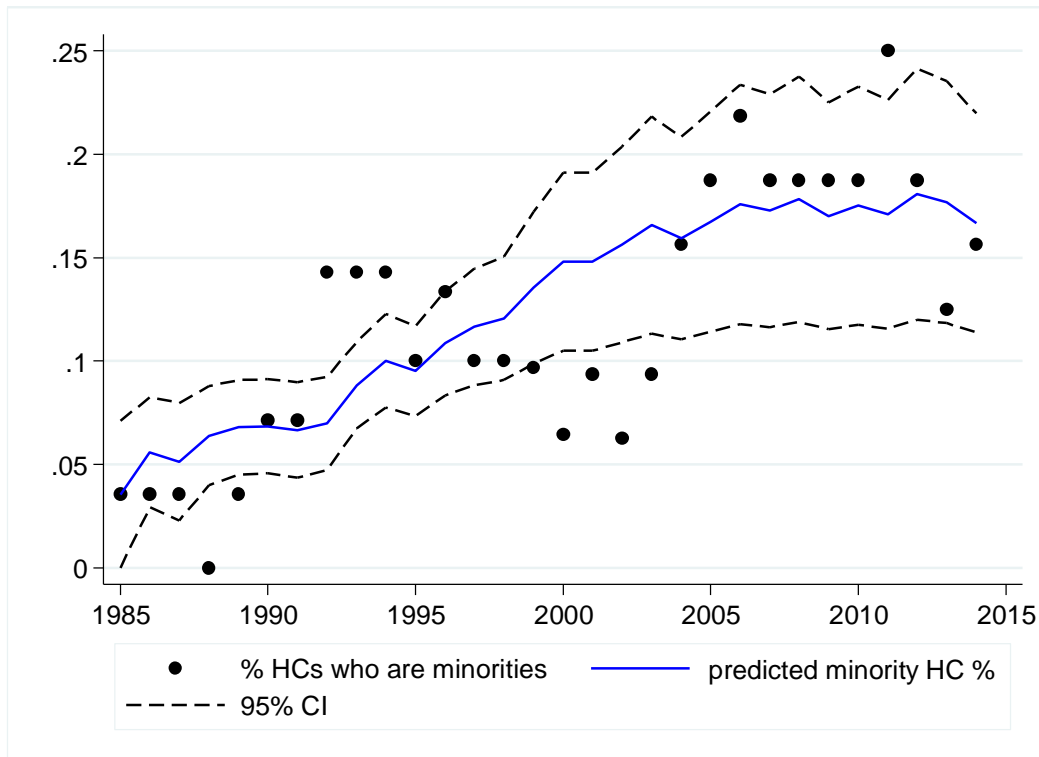
coach candidate pool is associated with an increase of approximately 0.6% in minority representation at the head coach level.

For both Model 3 and Model 4, we again generated annual predicted values based on observations between 1985 and 1999. We then extrapolated these predictions through 2015. Figures 7 and 8, respectively, depict predicted values (solid blue lines), bound by 95 percent confidence intervals (dashed black lines), along with the observed values for each year between 1985 and 2015 (solid black circles). In both cases, the evaluation is much less favorable for the Rooney Rule. Between 2000 and 2015, all but one of the observed values that are outside of the 95% confidence interval are actually below the lower bounds in Figure 7 and in Figure 8. If anything, these observations imply that the Rooney Rule decreased representation below what would have been expected in 1999. This possibility is, however, strongly counterbalanced by the preponderance of observed values within the 95% confidence interval. We might infer from this exercise that the Rooney Rule had no effect on minority representation among head coaches.

**Figure 7. Rooney Rule evaluation with counterfactual based on linear time trend as of 1999.**



**Figure 8. Rooney Rule evaluation with counterfactual based on based on diversifying candidate pool as of 1999.**



How can one reconcile the differences between Figures 4 and 7 and Figures 6 and 8, respectively? The only difference between the comparisons is the choice of year in which we draw the line between observation and prediction (i.e., 2002/2003 versus 1999/2000). Therefore, the years 2000, 2001, and 2002 are critical to generating predictions. What happened in those years that skews our inferences? Reviewing the hiring and firing of minority coaches in those years underscores why evaluations of the Rooney Rule are so sensitive to statistical assumptions.

In 2000, Ray Rhodes – an African American – was fired as head coach by the Green Bay Packers. In 2001, Herm Edwards – another African American, was fired as head coach of the New York Jets. In 2002, as mentioned previously, Dennis Green was fired by the Minnesota Vikings and Tony Dungy was fired by the Tampa Bay Buccaneers. But, Dungy was subsequently hired as head coach of the Indianapolis Colts in 2002. Hence, the total count of minority NFL head coaches fluctuated from three in 1999 to two in 2000 to three in 2001 and to two in 2002 (this can be seen in Figure 1). After minority representation increased from approximately 4% in 1985 to approximately 10% in 1999, representation dropped to 6% in 2000 and 2002 (9% in 2001). In short, the big difference between extrapolating from pre-2000 data versus pre-2003 data is the net reduction of a single African American head coach in the intervening years.

This statistical exercise demonstrates the difficulty of inferring a credible effect of a high-level diversity initiative on minority representation in a focal set of positions: inference is highly sensitive to the race of a single hiring or firing. As a useful reference point, the number of head coaching changes varies between three and eleven in a given year between 1985 and 2015 with a

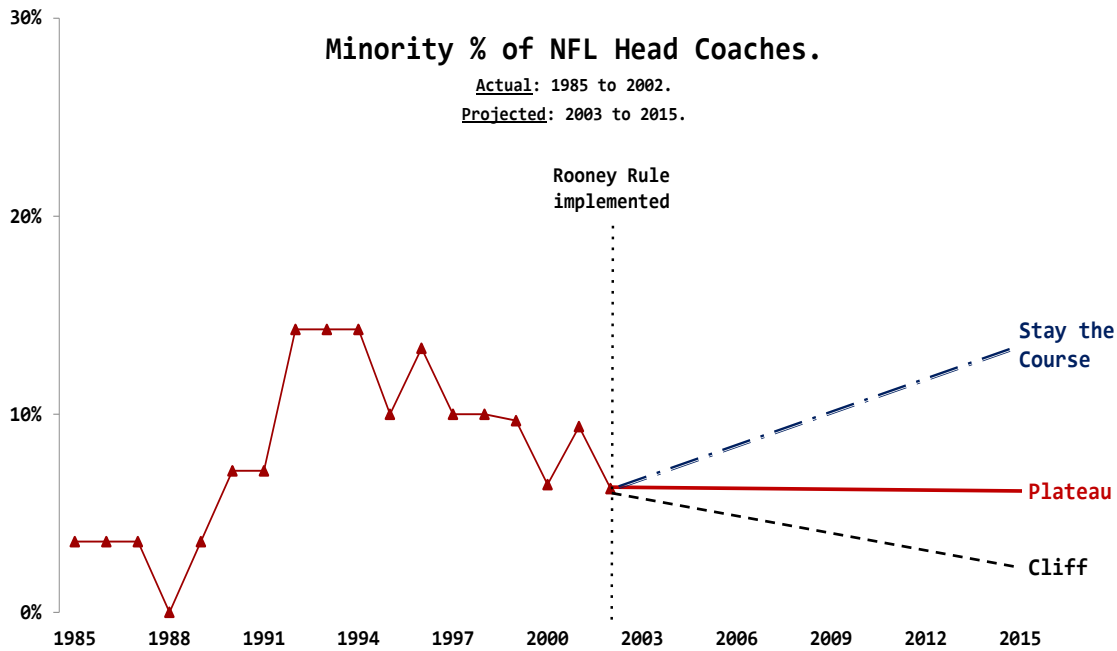
median of seven. Consequently, single cases substantially influence the observed representation values and subsequent predictions. The NFL context is especially sensitive to this small numbers issue because racial minorities constitute such a small proportion of head coaches (ranging from 4% to 25% during our observation period).

Our effort to estimate the causal effect of the Rooney Rule on minority representation highlights the difficulty of evaluating diversity interventions designed to increase representation in positions with small numbers of observations. It is also worth noting that the NFL consists of 32 franchises; evaluations of single organization interventions are even more susceptible to single cases biasing inferences. The fact that the candidate pool is also diversifying over time renders evaluation even more difficult. Basically, the underlying statistical challenge is that these evaluations are trying to hit a percentage target for which both numerator and denominator are small and our best predictor is changing over time.

### **Discussion and Interpretation**

Did the Rooney Rule increase minority representation among NFL head coaches? Our analysis does not enable us to answer this question with confidence. We think it appropriate to encourage readers to decide for themselves but to provide guidance on how their beliefs about “What might have been?” can inform their conclusion. Because our statistical analyses suggest that the years 2000, 2001, and 2002 are critical, we focus in particular on how one interprets the representation time trend and its implications for the post-implementation trajectory (i.e., the solid blue line in Figure 2). Figure 9 depicts these rhetorical counterfactuals graphically.

**Figure 9. Three stylized counterfactual trends, as articulated at Rooney Rule’s implementation.**



1. A plateau? If one believes that minority representation among NFL head coaches had risen as high as it could rise absent intervention prior to the Rooney Rule (e.g., a statistical plateau), then one can logically infer that the Rooney Rule caused representation to increase. This belief motivates the solid red horizontal line drawn in Figure 9 at the representation level observed between 1985 and 2002. In other words, the difference between this projected line and actual representation values between 2003 and 2015 would be the Rooney Rule effect estimate. Such a line is also consistent with the belief that minority representation among NFL head coaches is unrelated to rising minority representation among head coach candidates.
  
2. A cliff? If one believes that minority representation among NFL head coaches had not plateaued but, rather, begun decreasing to persistently lower levels in 2002 (or 2000), then one could infer that the Rooney Rule caused minority representation among NFL

head coaches to increase. This belief supports the downward-sloping, gray dotted line in Figure 9 so that the difference between this projected line and actual representation values between 2003 and 2015 would be the Rooney Rule effect estimate.

3. Staying the course? If one believes that minority representation among NFL head coaches had not plateaued but, rather, was increasing prior to 2002 (or 2000) either because of the diversifying pipeline of candidates or the passage of time (or any other plausible factor), then one could infer that the Rooney Rule did not cause minority representation among NFL head coaches to increase. This position supports the upward-sloping dashed blue line in Figure 9 – similar to those drawn in Figures 4, 6, 7 and 8 – so that the difference between this projected line and actual representation values between 2003 and 2015 would be the Rooney Rule effect estimate. It is worth considering that it would be difficult to maintain this belief and draw projected line that is significantly different in slope from the line that fits observed representation values.
4. What happened between 2000 and 2002? As our analysis shows, how one interprets the events of 2000, 2001, and 2002 influences one’s evaluation of the Rooney Rule. If one believes that the net reduction of one minority head coach during this period represented a statistical anomaly, then one should infer that the Rooney Rule had no effect on minority representation among head coaches. If, however, one believes that something systematic was influencing head coach employment and that the Rooney Rule addressed this influence then one can conclude that the Rooney Rule “righted the ship.” As we’ve tried to impress upon readers throughout, the small numbers statistical issue makes this determination highly subjective. In this light, it is clear why opinions on the Rooney Rule’s impact are so diverse.

## **Implementation Guidelines**

Is it possible to produce more definitive evidence of a diversity initiative's effect than the evidence produced by the Rooney Rule? We are optimistic that organizations can do so and offer some suggestions below on how they might. We start by describing the ideal design and implementation and then discuss next-best alternatives, with the caveat that next-best is unambiguously inferior to the ideal from a statistical evaluation perspective.

The ideal diversity initiative meets the following requirements:

1. A clearly-stated objective. The objective should be specific, quantifiable, and time-bound to aid statistical evaluation. This informs pre-intervention and post-intervention measurement decisions.
2. A clearly-defined and sufficiently large population or sample of interest (e.g., NFL head coaches). Defining the population or sample is necessary in order to randomize the intervention among participants assigned to "treatment" and "control" groups. A large sample is necessary to obtain statistical power in estimating differences in outcomes between those subject and not subject to the intervention.
3. A randomized-controlled trial (RCT). Although there is no unambiguous "gold standard" for statistical inference (Deaton and Cartwright, 2016), the medical sciences have accumulated a large body of knowledge through the use of RCTs. Principles of RCTs are described in great detail in many sources, as are methods for evaluating them. We suggest readers consult Gallo (2016) for the basic details, Bloom, et al. (2014) or Chatterji, et al. (2016) for guidance on conducting RCTs in the field, and Angrist and Pischke (2009) for

detailed econometric details. The key requirement for diversity professionals is that some sub-set of the population or sample of interest is randomly assigned to the intervention while the remainder is not. This random assignment is essential to estimating in order to estimate the counterfactual (i.e., the blue line in Figure 2).

4. Time series data on the outcome(s) of interest. Whatever it is that the organization seeks to change should be measured prior to and also after the intervention.
5. A skilled statistical analyst. The analyst need not be a Ph.D.-level statistician. Many interventions can be evaluated with a simple difference-in-means test, but the construction of counterfactuals and assessments of robustness of findings may require a minimum level of statistical expertise (e.g., Are treatment group values truly different from control group values?).

From our perspectives, these are not particularly onerous requirements. Any diversity initiative worth implementing should be allocated such investments so that the returns on diversity investments can be measured. Yet, it is rare that diversity initiatives follow these simple guidelines. For example, the Rooney Rule does not meet requirements 1 or 3 above. Despite our attempts to supply 4 and 5 (and even 1), we cannot credibly evaluate the Rule's effectiveness. But, it should be feasible for other organizations to do so. For example, large technology companies, the University of Texas System, and the U.S. Democratic Party are implementing or considering implementation of interviewing policies modeled after the Rooney Rule. It is feasible to apply these implementation guidelines to the large number of employee and staff positions that these organizations fill each year, as they employ far more than 32 people (as the NFL does). Doing so promises to produce more credible evaluations of the Rooney Rule's effectiveness.



The NFL, too, can implement these guidelines. In recent years, the NFL has announced Rooney Rule policies to increase the representation of women in front-office positions and the representation of minorities in coordinator positions. There are many more front-office positions than head coach positions and (approximately) twice as many coordinators as head coaches. Implementing randomized applications of the Rooney Rule should be relatively more feasible for these positions than for head coaching positions. Researchers like us would be pleased to collaborate with the NFL (or any organization) in designing and implementing diversity initiatives to enable credible policy evaluations. Despite our ambiguous evaluation of the Rooney Rule, we acknowledge that the NFL is at the forefront in recognizing and addressing racial disparity in employment opportunities. Moreover, nothing in our analyses suggests that the Rooney Rule has done anything to widen the racial gap in representation.

What if an organization cannot implement diversity initiatives that meet our guidelines? We also offer next-best alternatives for each guideline below.

1. If an organization cannot state the objective in the way we propose, then its next-best alternative is probably to allow others to debate its effectiveness using whatever means they view as appropriate. This is akin to the Rooney Rule's implementation. Objectives like "leveling the playing field" or "offering a fighting chance" do not inform statistical analysis and leave evaluations vulnerable to rhetorical manipulation by both proponents and opponents. We strongly advise organizations to clearly state objectives at the outset.

2. Most organizations will clearly identify the population or sample of interest so we do not identify a next-best alternative.
3. In the absence of a RCT design or other statistical approach to identifying treatment effects (due to small numbers or a lack of organizational enthusiasm or political obstacles or other reasons), we strongly advise organizations to state the counterfactual prior to implementation (i.e., a “rhetorical counterfactual”). That is, take the time trend of outcomes prior to implementation and try to draw Figure 2 so that the expectation absent intervention is clear ex ante even if it cannot be measured ex post (also our attempt at doing so in Figure 9). Ideally, this line will be informed by multiple parties using different methods and arriving at something of a consensus belief. The challenge of doing this only strengthens the case for a statistical approach to estimating the counterfactual with a RCT.
4. If a time series of pre-intervention data is unavailable then measurement at time of intervention can serve as a second-best alternative. Such measures do not enable analysts to validate similar pre-intervention trends among treatment and control groups but it does allow for validation at the time of intervention.
5. If the organization does not employ a suitable statistical analyst, then one can usually be hired on an hourly basis. The statistical analyses are fairly basic and difference-in-means tests are even more so. Many social science graduate students with quantitative statistical training should be capable of conducting such tests.

## **Conclusion**

Many organizations implement diversity initiatives to align employee demography with workforce demography. We proposed that typical implementations render it difficult, if not impossible, to evaluate the effectiveness of such initiatives and highlighted two statistical challenges: (1) diversifying candidate pools and (2) small numbers of targeted positions. To support our argument, we analyzed demographic data for all National Football League coaches from 1985 to 2015. Our analyses focused on a league-wide policy – the “Rooney Rule” – that was implemented in 2003 to increase racial minority representation at the head coach level. Our analyses make it clear that interpreting the evidence is difficult if not impossible given the way in which the Rooney Rule was implemented.

Considering the continuing debate of the Rule’s effectiveness, we demonstrated how initiative evaluations necessitate clearly-stated, ex ante expectations absent intervention (i.e., counterfactuals) and outcomes that are largely insensitive to single events (i.e., small numbers). Aiming to influence diversity management best practices, we also offered initiative design and implementation guidelines for enabling more credible evaluations of effectiveness. It should be feasible for most organizations to follow these guidelines, with or without the assistance of academic researchers. We hope that our work will inform and assist practitioners in accumulating even more persuasive evidence of what does and does not work in diversity management.

## References

- Apfelbaum, E. P., Stephens, N. M., & Reagans, R. (2016). Beyond one-size-fits-all: Tailoring diversity approaches to the representation of social groups. *Journal of Personality and Social Psychology*, 111(4), 547-566. doi: 10.1037/pspi0000071
- Alba, D. 2015. The NFL is Showing Silicon Valley How to Be More Diverse. *Wired*. October 26<sup>th</sup>. Last accessed: January 15, 2017. <https://www.wired.com/2015/10/tech-silicon-valley-nfl-rooney-rule-diversity/>
- Angrist, J. D. and J. S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton Univ. Press: Princeton, NJ.
- Baron, J. N., W. T. Bielby. 1980. "Bringing the firms back in: Stratification, segmentation, and the organization of work." *American Sociological Review*, 45: 737-765.
- BBC, 2016
- Bielby, W. T. 2000. Minimizing Workplace Gender and Racial Bias. *Contemporary Sociology*, 29(1):120-29.
- Bloom, N., J. Liang, J. Roberts, and Z. J. Ying. 2014. Does Working From Home Work? Evidence from a Chinese experiment. *Quarterly Journal of Economics*. 165-218.
- Bregman, P. 2012. Diversity Training Doesn't Work. March 12<sup>th</sup>. *Harvard Business Review Blog*. Last accessed: January 15, 2017. <https://hbr.org/2012/03/diversity-training-doesnt-work>
- Burke, M. 2013. Why The NFL's Rooney Rule Matters. *Forbes*. Last accessed: January 15, 2017. <http://www.forbes.com/sites/monteburke/2013/01/26/why-the-nfls-rooney-rule-matters/#20ccb9736a2d>
- Castilla, E. J. 2011. Bringing Managers Back In: Managerial Influences on Workplace Inequality. *American Sociological Review*, 76 (5): 667-694.
- Chatterji, A., M. Findlay, N. Jensen, S. Meier and D. Nielson. 2016. Field Experiments in Strategy." *Strategic Management Journal*, 37: 116-132.
- Collins, B. W. 2007. Tackling Unconscious Bias in Hiring Practices: The Plight of the Rooney Rule. *NYU Law Review*, 82(3): 870-912.
- Deaton, A. and N. Cartwright. 2016. Understanding and Misunderstanding Randomized Controlled Trials. *NBER Working Paper*, No. 22595.
- Dobbin, F, S Kim, and A Kalev. 2011. "You Can't Always Get What You Need : Organizational Determinants of Diversity Programs". *American Sociological Review*, 76:386.

Dobbin, F., D. Schrage, and A. Kalev. 2015. "Rage against the Iron Cage: The Varied Effects of Bureaucratic Personnel Reforms on Diversity". *American Sociological Review*, 80(5):1014-1044.

Duru, N. Jeremi. 2011. *Advancing the Ball: Race, Reformation and the Quest for Equal Coaching Opportunity in the NFL*. New York: Oxford University Press.

Edelman, Lauren B. 1990. "Legal environments and organizational governance: The expansion of due process in the American workplace," *American Journal of Sociology* 95:1401-1440.

Florio, M. 2016. Paul Tagliabue unhappy with performance of Rooney Rule. NBC Sports. Last accessed: January 15, 2017. <http://profootballtalk.nbcsports.com/2016/12/07/paul-tagliabue-unhappy-with-performance-of-rooney-rule/>

Fox, A. 2015. "How the Rooney Rule succeeds...and where it falls short." 19 May 2015. Last accessed: June 19, 2015. < [http://espn.go.com/nfl/story/\\_/id/12867233/rooney-rule-opened-doors-minority-head-coaching-candidates-do-more](http://espn.go.com/nfl/story/_/id/12867233/rooney-rule-opened-doors-minority-head-coaching-candidates-do-more)>

Freeman, M. 2013. "The Rooney Rule 10 Years Later: It's Worked...Usually, and we Still Need It." October 24. Last accessed: June 19, 2015. <<http://bleacherreport.com/articles/1822988-the-rooney-rule-10-years-later-its-worked-usually-and-we-still-need-it> >.

Gallo, A. 2016. A Refresher on Randomized Controlled Experiments. Harvard Business Review Blog. Last accessed: January 15, 2017. <https://hbr.org/2016/03/a-refresher-on-randomized-controlled-experiments>

Kalev, A. 2014. How You Downsize Is Who You Downsize: Biased Formalization, Accountability, and Managerial Diversity. *American Sociological Review*, 79(1):109–135.

Kalev, A, F Dobbin, and E Kelly. 2006. "Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies". *American Sociological Review*, 71:589-617.

Landry, B. 1987. *The new black middle class*. Berkeley, CA: University of California Press.

Madden, Janice F. 2004. "Differences in the Success of NFL Coaches by Race, 1990-2002: Evidence of Last Hire, First Fire." *Journal of Sports Economics* 5: 6-19.

Madden, Janice F. and Ruther, M. 2009. "Reply to: Differences in the Success of NFL Coaches by Race: A Different Perspective." *Journal of Sports Economics* 10: 543-550.

Madden, Janice F. and Ruther, M. 2011. "Has the NFL's Rooney Rule Efforts "Leveled the Field" for African American Head Coach Candidates?" *Journal of Sports Economics* 12: 127-142.

McPherson, J. M., L. Smith-Lovin. 1987. "Homophily in voluntary organizations: Status distance and the composition of face-to-face groups." *American Sociological Review*, 52(3): 370-379.

National Football League. 1985-2002. *The official National Football League record and fact book*. New York: Workman Publishing.

O'Keefe, E. 2017. Under pressure, Schumer wants Democrats to focus on minority hiring — by using an NFL tool. *Washington Post*. Last accessed: January 15, 2017.

[https://www.washingtonpost.com/news/powerpost/wp/2017/01/06/under-pressure-schumer-wants-senate-democrats-to-adopt-the-rooney-rule-for-staff-hires/?utm\\_term=.af1ad66cc723](https://www.washingtonpost.com/news/powerpost/wp/2017/01/06/under-pressure-schumer-wants-senate-democrats-to-adopt-the-rooney-rule-for-staff-hires/?utm_term=.af1ad66cc723)

Palmeri, C. and B. Womack. 2013. BET Founder Sees Business 'Rooney Rule' Adding Minorities. *Bloomberg*. Last accessed: January 15, 2017.

<https://www.bloomberg.com/news/articles/2013-07-09/bet-s-johnson-wants-minorities-interviewed-for-all-senior-jobs>

Reskin, B. F. 2000. The Proximate Causes of Employment Discrimination. *Contemporary Sociology: A Journal of Reviews*, 29(2):319–28.

Reskin, B. F., D. A. McBrier, J. Kmec. 1999 "The determinants and consequences of workplace sex and race composition." *Annual Review of Sociology*, 25: 355-261.

Rider, C. I., J. B. Wade, A. Swaminathan, and A. Schwab. 2017. Racial disparity in leadership: Performance-reward bias in promotions of National Football League coaches. Working paper, Georgetown University. Available on SSRN: <https://ssrn.com/abstract=2710398>.

Rissing, B. A. and E. J. Castilla. 2014. House of Green Cards: Statistical or Preference-Based Inequality in the Employment of Foreign Nationals. *American Sociological Review* 79 (6): 1226-1255.

Solow, B. L., J. L. Solow, and T. B. Walker. 2011. Moving on up: The Rooney Rule and minority hiring in the NFL. *Labour Economics*, 18: 332-337.

Stephens, N. and E. Apfelbaum, 2016. The Real Reasons Diversity Programs Don't Work. *Fortune*. Last accessed: January 15, 2017. <http://fortune.com/2016/08/16/diversity-workplace-lessons/>

Verrier, R. 2016. What Hollywood can learn from the NFL in addressing its diversity problem. *Los Angeles Times*. Last accessed: January 15, 2017. <http://www.latimes.com/entertainment/envelope/cotown/la-et-ct-on-the-record-white-20160303-story.html>

Weber, M. 1978. "Basic Sociological Terms." Pp. 3–62 in G. Roth and C. Wittich (eds.) *Economy and Society*. Berkeley, CA: University of California Press.

Wilner, B. 2017. Rooney Rule: 8 Minority Head Coaches for 2017. ABC News. Last accessed: January 15, 2017. <http://abcnews.go.com/Sports/wireStory/rooney-rule-minority-head-coaches-2017-4476779>

**Chapter 6: Solving the Legal Diversity Problem / *William D. Henderson,*  
*Professor of Law and Van Nolan Faculty Fellow at Indiana University*  
Published in the Indiana University Maurer School of Law, Legal Studies Research Paper  
Series: Research Paper Number 335, PD Quarterly (February 1, 2016)**

The legal profession's lack of progress on diversity stems from a systems problem, not a lack of moral resolve, and applied research suggests there are ways that law firms can address the systems problem.

Here is a familiar fact pattern in large U.S. law firms.

- Time 1. Partners come together and agree that diversity is part of their firm's core values; they review the firm's bleak statistics, particularly at the partnership level, and agree they can and will do better.
- Time 2. Through significant time and expense, they successfully recruit a diverse class of incoming associates.
- Time 3. A disproportionately large number of female and diverse associates leave the firm.
- Time 4. The remaining associates eligible for partner are primarily white men.
- Time 5. Partners come together and agree that diversity is part of their firm's core values; they review the firm's bleak statistics, particularly at the partnership level, and agree they can and will do better.

Why does this cycle repeat itself? As a long-time law firm re-searcher who has seen this cycle play out over several iterations, I can tell you that it is easy for a group of lawyers, especially those new to leadership, to convince themselves that they can solve the profession's diversity



problem through greater moral resolve. Yet, if the root causes are not moral in nature, we won't make much progress.

In this article, I ask readers to consider the possibility that the profession's lack of progress on diversity is a systems problem rather than a failure of moral resolve.

What does it mean to have a systems problem? Every firm has a system of recruitment, selection, development, feedback, evaluation, and promotion that enables law graduates to enter as legal novices and, through years of effort, acquire the skills, knowledge, and experience necessary to become partners. At most law firms, however, this system is driven more by tradition and past practice than science. Further, the system seldom places explicit or rigid demands on partner-owners because partner-owners prize their autonomy and are given the greatest rewards for bringing in business. To the extent the system relies on measurement, the quality of the data is uneven and under-analyzed. Stated another way, the "system" for creating successful lawyers and partners is not much of a system at all. And in this ignorance lies the cause of our diversity problem.

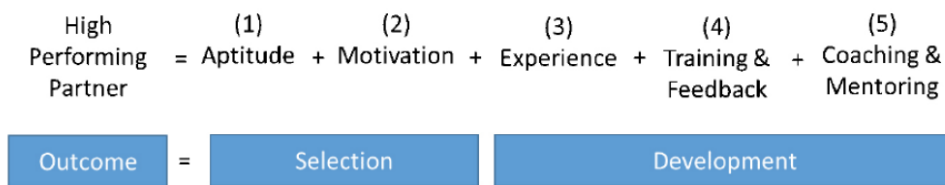
For the last several years, I have shifted my focus from academic to applied research. Although academic ideas can be elegant, compelling, and important, their major limitation is that we don't really know if they will work in actual practice. Applied research attempts to sort this out, usually through social scientists hired by organizations that are hungry for a competitive advantage. The goal of applied research is to find solutions to important problems and then make them cheap and simple to implement. Law has a shortage of applied researchers, partially

because the profession has been so prosperous for so long (what's there to fix?) and partially because lawyers tend to be uncomfortable with data and statistics. Yet, these background factors are starting to change.

In this article, I am going to share what I have learned through my applied research as it bears on the problem of law firm diversity. The bottom line is that the problem is fixable. If we design and implement a better system, out the other side will flow successful diverse attorneys in roughly the same proportion as the number we managed to hire several years earlier. Further, the stakes are hardly academic. Organizations with a reliable system for creating diverse lawyers will have a competitive advantage for attracting clients and the best entry-level talent. Likewise, esteem and accolades await the leaders who finally make a breakthrough on law firm diversity.

### **You Have to Start with a Theory**

An intelligent system is invariably built upon a theory drawn from multiple sources. One high quality source is published empirical research. A second is one's own professional work experience: "When I have tried X, Y usually happens" — so we rely on X. Finally, a subset of our theories will be based on pure reason: "Based on our collective knowledge and experience, this is the best approach for this problem." Figure 1 is a summary of my own theory for creating high-performing partners.



**Figure 1. Elements Needed to Create a High-Performing Partner.**

In narrative form, I am saying that the creation of high-performing partners is influenced by five factors: (1) aptitude, also known as cognitive ability; (2) motivation, which is primarily a function of values alignment between the lawyer and the substance of his or her work; (3) the type and quality of work experience that a lawyer receives during his or her early career; (4) the quality, quantity, and timeliness of training and feedback; and (5) the presence and quality of a mentoring or coaching relationship.

The model can also be broken down into selection and development components. A law firm optimizes elements (1) and (2) through a process of accurate selection at the point of hiring. The less accurate the selection, the higher the lawyer attrition due to poor fit for aptitude and motivation. A firm can optimize (3), (4), and (5) by designing and implementing systems for professional development. The better the design and execution of the interconnected systems, the faster and higher the lawyer’s growth trajectory.<sup>103</sup>

---

<sup>103</sup> This is, fundamentally, a system for maximizing human potential as a lawyer. This is the source of success for the greatest brands in professional services, such as Cravath, Swaine & Moore, McKinsey, and Goldman Sachs. See Robert T. Swaine, *The Cravath Firm and Its Predecessors Vol. II* (1948) (the purpose of the Cravath system was to create “a better lawyer faster”); Marvin Bower, *Perspective on McKinsey* (1979) (explaining how the original McKinsey was modeled on the associate-partner model of Jones Day and setting forth the refined McKinsey model from entry-level to partner); Charles D. Ellis, *The Partnership: The Making of Goldman Sachs* (2008) (explaining the origins of the teamwork ethos at Goldman Sachs and how it became the basis for recruitment criteria and explaining the origins of the firm’s values, which emphasize the necessity of continuous professional development).

What is the relative importance of these factors? This is a good question that no one can answer with any degree of precision, primarily because we are in the early days of applied research within the legal profession and the required data have not yet been collected and analyzed. The best we can do is to start with a theory that is consistent with the data we do have and continuously improve our knowledge through measurement.

It has been my experience, however, that lawyers often have strong opinions on what does and doesn't matter. These views on lawyer selection and development essentially create a series of default settings based on conventional wisdom and past practice. I have enough knowledge of the social science literature and enough experience doing sophisticated applied research in law firms to conclude that many of these default settings are wrong.

Below is a summary of what I know about each of the five components in my five-factor model. One by one, and cumulatively, these model components provide me with optimism that law firm diversity can be dramatically improved, particularly at the partnership level.

### **(1) Aptitude**

Intelligence is an obvious baseline requirement for successful lawyers, including high-performing partners. Yet, what quantum of intelligence, or IQ, is needed to perform at a high level? There is a natural presumption that more intelligence is better. Thus, many legal employers favor highly selective law schools on the theory that these schools only admit students

with high LSAT scores and strong undergraduate GPAs. If a candidate graduated from a non-elite law school, law school grades tend to become much more important in hiring decisions.

These academic filters tend to have a negative impact on the ability of law firms to recruit more diverse entry-level candidates. Yet, a more fundamental question is whether heavy reliance on academic proxies truly produces a better candidate pool.<sup>104</sup> If lawyers and law firms operate on the presumption that the answer to this question is yes when the actual empirical answer is no, all sorts of negative consequences follow:

- Partners and professional staff waste time and resources on selection criteria that don't matter
- Higher quality candidates who excel on dimensions such as motivation and values alignments are never interviewed at all;
- Developmentally rich work assignments are disproportionately allocated toward majority associates because, as a group, they tend to perform better on academic proxies.

The heavy emphasis on academic markers may be misplaced because law school graduates are already part of a heavily filtered population. To become a licensed lawyer, the typical path is to complete a four-year college degree, obtain an LSAT score high enough for admission into an ABA-accredited law school, complete three years of law school, and pass a state bar examination. This is what psychologists refer to as a “range restricted” population — compared

---

<sup>104</sup> For a more exhaustive treatment of this topic, see William D. Henderson, “Successful Lawyer Skills and Behaviors,” in *Essential Qualities of the Professional Lawyer* (Paul Haskins, ed., 2013); and William D. Henderson, “Law Firm Strategies for Human Capital: Past, Present, Future,” in *Studies in Law, Politics and Society* (Austin Sarat, ed., 2010).

to the general population, this group has very high cognitive ability. This fact raises a simple, testable empirical question: Do marginal gradations on a handful of academic measures reliably signal greater potential to become a high-performing lawyer?

Several years ago, this question was taken up by Professors Marjorie Shultz and Sheldon Zedeck of the University of California at Berkeley,<sup>105</sup> and, remarkably, the answer was no. Drawing upon the methodology of industrial and organization-al (IO) psychology (Zedeck's area of expertise), the researchers identified a set of 26 lawyer effectiveness factors (see Figure 2 next page). Behaviorally anchored rating scales (BARS) were then created to measure lawyer effectiveness on a 1 to 5 scale, with increments defined by specific, concrete examples of lawyer behaviors. The researchers then used the BARS to obtain peer and supervisor evaluations for over 1,100 law alumni of UC Berkeley and UC Hastings and approximately 200 UC Berkeley law students. In turn, these measurements were correlated with participants' undergraduate GPAs, LSAT scores, and law school grades.

---

<sup>105</sup> See Marjorie M. Shultz and Sheldon Zedeck, Identification, Development, and Validation of Predictors for Successful Lawyering (Sept. 2008), [www.law.berkeley.edu/files/LSACREPORTfinal-12.pdf](http://www.law.berkeley.edu/files/LSACREPORTfinal-12.pdf).

<p style="text-align: center;"><b>Intellectual and Cognitive</b></p> <ul style="list-style-type: none"> <li>• Analysis and Reasoning</li> <li>• Creativity and Innovation</li> <li>• Problem Solving</li> <li>• Practical Judgment</li> </ul>	<p style="text-align: center;"><b>Conflict Resolution</b></p> <ul style="list-style-type: none"> <li>• Negotiation Skills</li> <li>• Able to See the World through the Eyes of Others</li> </ul>
<p style="text-align: center;"><b>Research and Information Gathering</b></p> <ul style="list-style-type: none"> <li>• Researching the Law</li> <li>• Fact Finding</li> <li>• Questioning and Interviewing</li> </ul>	<p style="text-align: center;"><b>Client/Business Relations: Entrepreneurship</b></p> <ul style="list-style-type: none"> <li>• Networking and Business Development</li> <li>• Providing Advice and Counsel, and Building Relationships with Clients</li> </ul>
<p style="text-align: center;"><b>Communications</b></p> <ul style="list-style-type: none"> <li>• Influencing and Advocating</li> <li>• Writing</li> <li>• Speaking</li> <li>• Listening</li> </ul>	<p style="text-align: center;"><b>Working with Others</b></p> <ul style="list-style-type: none"> <li>• Developing Relationships within the Legal Profession</li> <li>• Evaluation, Development, and Mentoring</li> </ul>
<p style="text-align: center;"><b>Planning and Organization</b></p> <ul style="list-style-type: none"> <li>• Strategic Planning</li> <li>• Organizing/Managing One's Own Work</li> <li>• Organizing/Managing Others (Staff/Colleagues)</li> </ul>	<p style="text-align: center;"><b>Character</b></p> <ul style="list-style-type: none"> <li>• Passion and Engagement</li> <li>• Diligence</li> <li>• Integrity/Honesty</li> <li>• Stress Management</li> <li>• Community Involvement and Service</li> <li>• Self-Development</li> </ul>

Figure 2. Shultz and Zedeck – 26 Lawyer Effectiveness Factors.

The results of the Shultz-Zedeck study suggest that academic factors are not very reliable proxies for future lawyering potential. Among the law school graduates in the sample, factors such as Analysis and Reasoning, Researching the Law, Writing, and Problem Solving showed modest, positive correlations with grades and LSAT scores (between 0.10 and 0.15,  $p > .05$ ). Yet, LSAT scores and first-year grades were also negatively correlated at statistically significant levels with Networking (-.122) and Community Service (-.096). In the student sample, undergraduate GPA was positively correlated with no effectiveness factors but negatively associated with Practical Judgment (-.169), Seeing the World through the Eyes of Others (-.170), Developing Relationships (-.195), Integrity (-.189), and Community Service (-.152).

The Shultz-Zedeck study also documented numerous job-relevant markers of future success as a lawyer that are likely crowded out by over-reliance on academic proxies. For ex-ample, Shultz and Zedeck correlated lawyers' BARS scores with performance on the Hogan Personality

Inventory (HPI), an established, off-the-shelf personality assessment that has been validated on a large sample of knowledge workers. Overall, the HPI provides much stronger signals for lawyer effectiveness:

- On the HPI Adjustment construct, which measures emotional stability and steadiness under pressure, alumni lawyer scores were positively correlated at statistically significant levels with 22 of the 26 effectiveness factors (ranging from .072 to .220) and negatively correlated with none.
- On the HPI Prudence scale, which measures self-control and conscientiousness, alumni lawyer scores were correlated with 18 effectiveness factors (ranging from .071 to .189) and negatively correlated with none.
- On the HPI Ambition scale, which measures achievement and leadership orientation, alumni lawyer scores were positively correlated with 14 effectiveness factors (ranging from .076 to .239) and negatively correlated with none.

These correlation patterns strongly suggest ample opportunities for law firms to engage in better selection by relaxing their emphasis on academic proxies and improving their focus on job-relevant behaviors. Moreover, on these broader measures of lawyer effectiveness, Shultz and Zedeck documented no performance gap based on race and gender.

These findings may surprise some law firm partners who are wedded to the idea that pedigree — perhaps the pedigree that they possess — is a strong proxy for lawyer potential. Yet, the Shultz-Zedeck findings are broadly consistent with what we have found on numerous client projects.



For example, I have been a part of numerous internal law firm studies designed to build success profiles based on résumés and transcripts. The purpose of these success profiles is to identify future high performers.<sup>106</sup> One of the most persistent results across virtually all of these studies is that attendance at an elite law school is seldom a marker of future success and often a slight negative predictor.

Another persistent finding is that law school grades tend to predict future performance within a law firm. Yet, grades are a function of both aptitude and effort. Since law graduates at more elite law schools tend to have higher LSAT scores, and attending an elite law school does not predict future high performance, the predictive power of law school grades within law firms is probably attributable to higher levels of motivation.

In summary, the proportion of law school graduates with the requisite aptitude to become high-performing partners is probably larger than most law firm partners would tend to believe.

Reducing a firm's reliance on academic proxies will increase the number of candidates who could be considered for hiring, which in turn expands the number of diverse applicants who might be eligible for an interview.

## **(2) Motivation**

---

<sup>106</sup> In the field of IO psychology, these studies are referred to as biographical inventories, though we tend to refer to them as Moneyball studies.

In my five-factor model for creating high-performing partners, my only other selection criterion is motivation. As noted above, I view motivation as primarily a values alignment between the lawyer and the substance of his or her work.

My theory runs as follows: High performance inside a law firm requires relentless focus on other people's problems, typically legal problems mixed together with business, professional, and personal dimensions. Is a candidate motivated to solve these types of problems for 50+ hours per week? If the answer is yes, he or she will be sufficiently self-directed to take advantage of the developmental opportunities that the firm will provide. Indeed, it is through these opportunities that the lawyer becomes intelligent and capable. The reason is deliberative practice rather than native intelligence.<sup>107</sup> If the answer is no, there is a values misalignment and the candidate is better off in a work environment where the quality and quantity of problems are a better fit.

The importance of motivation as a selection criterion flows from my applied work with law firms. In these projects, we frequently use an assessment tool called the Achievement Motivation Inventory (AMI). The AMI measures 17 dimensions of achievement motivation, which are grouped into three broad categories: (1) Ambition, (2) Self-Assurance, and (3) Self-Control. We also use another assessment called the Management Development Questionnaire (MDQ), which

---

<sup>107</sup> There are many books on deliberative practice. For a good overview of the topic, see Geoff Colvin, *Talent Is Overrated* (2009). Many lawyers operate on the assumption that intelligence is genetically based and fixed. However, in cognitive science, intelligence is broken down into fluid and crystallized intelligence. The former is the ability to think logically, identify patterns, and solve complex problems independent of prior knowledge. In contrast, crystallized intelligence is the ability to use skill, knowledge, and experience that are accumulated over time. See generally Paul Kline, *Intelligence: The Psychometric View* (1991).

measures various work behaviors that map onto a competency model similar to the Shultz-Zedeck 26 effectiveness factors.

One of the most persistent findings we observe is that high-performing partners tend to score significantly higher on the AMI, particularly on dimensions related to self-assurance and self-control. The advantage of these higher levels of motivation is that the lawyer has the emotional and mental staying power to acquire, over a long period of time, job-relevant skills and knowledge. The advantage of this greater staying power can be observed on the scores of the MDQ. On a 200-point scale, the typical law student scores a 101 compared to 107 for an associate, 115 for a client service partner, and 123 for a high-performing partner, mirroring the progression within a law firm.

Assessment tools like the AMI and MDQ are useful tools for professional and organizational development, but they should not be relied upon as a primary selection tool for entry-level employment.<sup>108</sup> This is because high-aptitude candidates will try to figure out what the employers want to see and hear. A better alternative is a well-structured behavioral interview that asks questions that map onto a set of predefined skills and competencies necessary to be successful at the firm.

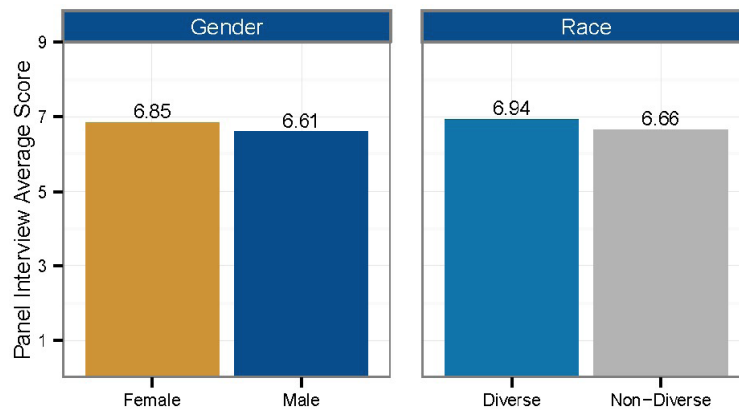
The core principle underlying a behavioral interview is that past behavior is a relatively good predictor of future behavior. If you are looking for someone with initiative, or practical problem

---

<sup>108</sup> The AMI and MDQ and similar personality based assessments can be useful in the hiring context, but they are not a substitute for a well-structured interview.

solving skills, or team work, ask questions that elicit specific, concrete examples of these behaviors and score the candidates based on the quality and quantity of the examples they provide. The scores on these various job dimensions tend to cluster together, which I interpret as motivation to learn, improve, and perform at an overall high level.

I have helped construct these types of interview systems for several law firms. These projects typically include statistical analysis to monitor the impact on diverse candidates. Figure 3 is a



**Figure 3. Scores on Structured Panel Interviews at Am Law 200 Firm, by Gender and Minority Status.**

chart that compares interview scores of entry-level candidates at one Am Law 200 firm over three years of recruiting (n = 350). Scoring is on a nine-point scale where scores of 7.0 or higher were usually necessary to receive an offer. Each interview was identical in format, structure, and content. The interviews were scored in a panel format by law firm partners. The typical number of interviewers was four.

The most heartening aspect of Figure 3 is that the familiar performance gap for female and minority candidates is not present. In fact, in this relatively large sample, female and minority candidates tend to do slightly better than their white, male counterparts. These results are also

consistent with the findings of the Shultz-Zedeck study. Simply put, if a law firm applies job-relevant criteria in a uniform, structured way to a diverse array of entry-level candidates, they should expect a roughly equal proportions of diverse candidates to receive scores in the “highly qualified” range.

Not surprisingly, candidates who have been put through this structured panel interview (SPI) process are much more likely to accept an offer of employment. As a result of the SPI process, the firm in Figure 3 increased its yield from 33% to 48% over a three-year period despite becoming more selective in making offers. Minority yield rates were significantly higher. Through a subsequent debriefing process, minority candidates told our researchers that they perceived the process as being more thorough and fair, thus making them more confident that the firm had an overall plan for their professional development.

### **(3) Experience**

Parts (1) and (2) of my model suggest that proper selection criteria and valid and reliable selection methods will yield a group of high-potential entry-level female and diverse lawyers in roughly the same proportion as the overall law school population. The next step is bringing that high potential to fruition.

There is strong empirical evidence documenting that the quantity and quality of work assignments are a major driver of professional development. Those who receive these opportunities early in their careers tend to thrive, as they can immediately leverage their enhanced skill set to obtain even better assignments in the future. Conversely, those who are

assigned more repetitive tasks – often to serve the immediate needs of the firm, a partner, or a specific client – are often put at a permanent disadvantage, as they lack the skills, experience, and accomplishments to compete with lawyers of the same age and billing rate who received better early assignments.

The exact pattern described above was documented in a large 1,000-lawyer U.S. law firm. The research was done by Professors Forrest Briscoe of Penn State University and Katherine Kellogg of MIT and published in the *American Sociological Review*, the leading peer-reviewed journal in the field of sociology.<sup>109</sup> Their core research question was focused on the circumstances under which a knowledge worker could use a work-family program that offered a reduced or flex-time policy (often to accommodate child care responsibilities) without it have a damaging impact on the worker's long-term career progression and pro-motion prospects. The work context happened to be a law firm.

Through a carefully designed study covering 958 associates who entered the firm between 1997 and 2005, including 71 who participated in the work-family program, the researchers learned that program participants were more likely to fall behind in pro rata performance pay and more likely to exit the firm. This finding is after controlling for law school rank, race, gender, undergraduate grades, office location, and practice department.

---

<sup>109</sup> See Forrest Briscoe and Katherine C. Kellogg, “The Initial Assignment Effect: Local Employer Practices and Positive Career Outcomes for Work-Family Program Users,” 76 *American Sociological Review* 291-319 (2011).

Yet, the researchers also found that the negative outcomes associated with the work-family program were significantly reduced when the associates were exposed to high-quality work assignments during their first few months at the firm. For all associates, these types of early career assignments were associated at statistically significant levels with higher performance pay and longer tenure. Yet, for program participants, it had a significantly larger impact, essentially inoculating them from the negative consequence experienced by their peers. Observed Briscoe and Kellogg, “Exposure to powerful initial supervisors helps employees gain access to reputational-building project opportunities, which in turn allows them to build a significant track record with a wide range of supervisors and clients by the time they use the reduced-hours program.”<sup>110</sup>

The implication of the Briscoe-Kellogg study is that high-quality work assignments are an essential ingredient to a lawyer’s career progression. An associate who has successfully climbed the first one or two rungs of this career ladder is both more visible and attractive to other partners. Under the free market assignment systems that so many law firms embrace as part of their culture, associates who get early breaks are also set up for future career-enriching opportunities, largely because more experienced junior associates help partners build their practices.

Now, what is the likelihood that these networks of developmentally rich early career assignments are blind to both race and gender? Our work with law firms has given us a window to explore this question. For example, we have built several data-driven success profiles of in-coming

---

<sup>110</sup> Id. at 297.

associates. The concept is simple. Identify high-performing senior and mid-level associates and, using multivariate statistical analysis, work backward to identify entry-level attributes on résumés and transcripts that can be used to predict future success.

When possible, we like to include demographic information as a control variable to ensure that a negative race or gender impact is not baked into a firm's success profile. Because the legal profession has such a bleak record on diversity, it should come as no surprise that a statistical analysis of associate records is going to document some unpleasant facts, such as lower evaluations and higher rates of attrition for specific subgroups. But these bad outcomes are not the same thing as discrimination, as they could result from poorly constructed systems.

When the data are available, we dig deeper and try to understand why females and non-white men are getting lower performance scores. To date, we have never encountered systemic racial or gender bias: partners, regardless of their race or gender or the race or gender of the associate being evaluated, tend to agree — “This associate is excellent,” “This associate is below average,” etc. Yet, in large datasets involving over 250,000 associate hours, we have observed very large gender- and race-based patterns in which junior associates gravitate toward supervisors who match their own race and gender.<sup>111</sup> Although this dynamic may be comfortable for associates, their female or minority supervisors tend to control fewer important client relationships and thus have fewer developmentally rich assignments to allocate. In contrast, white males who attend elite law schools are ideally situated to receive these types of

---

<sup>111</sup> See also William D. Henderson, “Diversity by the Numbers,” NALP Bulletin (July 2012) (documenting the large racial subgroup effect in explaining associate-level diversity in which Blacks attract Blacks, Hispanics attract Hispanics, Asians attract Asians, etc.)



opportunities because their comfort zones overlap with the power center of most large law firms.

The longer law firms ignore the profound effect played by work assignments, the longer the profession will be plagued with a diversity problem. Again, this is not a problem of a moral deficit; it's a systems problem. Yet, I am among the group of lawyers who believe that it is morally wrong not to fix this system. Moreover, it is just bad business.

#### **(4) Training and Feedback**

To reiterate my basic claims thus far, the formula for a high-performing partner is (1) a lawyer with a high cognitive aptitude who (2) possesses a high motivation to learn and be successful in a legal service organization and (3) is given the opportunity to do progressively more challenging work so

as to signal his or her capabilities to colleagues and clients. Components (1) and (2) raise issues of selection. Component (3) is entirely developmental. Through work opportunities, law students and junior lawyers become someone who appears to others as smart or naturally gifted. Yet, few law firm partners understand that a system is at work. Hence, the system is poorly tooled and tends to produce results that are not in alignment with our professional values and business goals.

The fourth component of my model is training and feedback. Experience alone is unlikely to round out a lawyer's professional development. For issues of quality, cost, or both, there will be times when skills and knowledge are best taught through formal training. Feedback is a communication loop that aids a lawyer in connecting together tacit information embedded in

both lawyer training and experience, thus enhancing judgment and overall lawyer effectiveness. Feedback accelerates professional development. Feedback is also very expensive because it requires a practice master to closely observe performance and communicate subtle points in a manner that the less experienced lawyer is able to hear and absorb.<sup>112</sup>

Because legal service organizations lack a systems perspective, they routinely confuse the cost of feedback, which tends to be short-term and personal, with the value of feedback, which is short-, medium-, and long-term and affects the competitiveness of the entire enterprise. Specifically, an organization composed of lawyers who are too busy to provide high-quality feedback to high-potential people is an organization with fewer high-performing lawyers. Further, the problem only compounds and worsens over time.

The best way to illustrate the importance of training and feedback to professional development, including the large implications for female and diverse lawyers, is to relate a clear success story. In my research, the most compelling knowledge worker example I have come across is the Bell Labs study, which was written up nearly 20 years ago in the Harvard Business Review and expanded upon in a best-selling business book.<sup>113</sup>

The basis for the study was an antitrust consent decree between the Department of Justice and AT&T, which forced divestiture of AT&T operating units and created the seven “Baby Bells.”

---

<sup>112</sup> See generally William D. Henderson, “Supercharging Lawyer Development Through Feedback,” NALP Bulletin (June 2014).

<sup>113</sup> See Robert Kelley and Janet Caplan, “How Bell Labs Creates Star Performers,” Harvard Business Review (July-August 1993); Robert E. Kelley, *How to Be a Star at Work* (1998).

One of the implications of the settlement was that AT&T would no longer have the monopoly profits to subsidize the famed Bell Laboratories, which had funded the research of several Nobel Prize winners. Despite decades of success, there was no guarantee that Bell Labs could survive in the private sector as a pure science think tank.

To increase their odds of success, Bell Lab executives committed the organization to an internal study that would enable them to identify top-performing engineers. The reasoning was simple: If the organization could hire and develop more “ten- or twenty-for-ones,” productivity would skyrocket and the organization would become a magnet for paid client work. To run this project, Bell Labs engaged IO psychologist Robert Kelley, whose prior work specialized in productivity assessments in the emerging “gold collar” sector.

In Kelley’s study, an engineer was designated a top performer if he or she was identified as such by peers, managers, and (eventually) the organization’s clients. Kelley polled managers and workers to generate theories of success based on various cognitive, psychological, and social factors. In turn, these theories were tested using a large sample of engineers and a two-day battery of tests designed to measure 45 alleged attributes of success. Yet, despite the tremendous expenditure of time and resources, Kelley and his colleagues came up empty: There was no appreciable relationship between status as a star performer and any of the cognitive, psychological, social, or environmental factors. Attempts to reanalyze the data were equally fruitless.

Puzzled by these findings, Bell Lab executives extended the study so that Kelley's research team could generate new theories of star productivity based on observation rather than self-reporting. For the next two years, Kelley and his researchers examined the work habits and strategies of Bell Lab engineers. At the end of this process, they identified nine work strategies that distinguished star performers from the middle-of-the-road engineers. In relative order of importance, they included:

1. Taking Initiative. Top performers took responsibility above and beyond their stated jobs, volunteering for new activities and promoting new ideas;
2. Networking. Top performers were deft at tapping into coworkers' expertise and shared their own knowledge with those who needed it;
3. Self-Management. Top performers were very good at regulating their own work commitments, time, performance level, and career growth;
4. Perspective. Top performers understood their jobs within the larger context of the organization and could analyze problems from the viewpoint of customers, managers, and team members;
5. Followership. Although perceived by others as leaders, top performers excelled at setting aside their own agendas and using their talents to help other leaders accomplish the organization's goals;
6. Teamwork. Top performers were more willing to assume joint "ownership" of goal setting, group commitments, work activities, schedules, and defusing conflict among group members;

7. Leadership. Top performers had the ability to formulate, state, and build consensus on common goals and then work to accomplish them;
8. Organizational Savvy. Top performers recognized and thus could navigate competing interests within the organization;
9. Show-and-Tell. Top performers typically had the ability to present their ideas persuasively in written or oral form.

One of the most striking features of Kelley's research was the propensity of average workers to draw the wrong lessons from the success of top performers. Average performers tended to invert the order of priority and thus focus on organizational savvy and show-and-tell, which they surmised was the key — based on the success of the stars — to impressing management.

Similarly, middle performers viewed networking as staying “in the loop” on office gossip and getting to know people who could help their careers. Top performers, in contrast, viewed networking as a bartering system in which the cost of admission was technical expertise and staying in the loop required a sincere commitment to reciprocity. Kelley reported that star performers got their phone calls returned faster than their middle-performing peers, who were typically receiving bad answers slowly — hardly a recipe for career success.

Yet, the most remarkable finding of the Bell Labs study was that the star performer work strategies were found to be teachable — an outcome verified using the controlled experiment methodology, which is the gold standard for empirical research. The study documented that engineers who received star performer training (one day per week for several weeks) tended to

post statistically significant gains in productivity over the next year as compared to the control group that did not get the training. Yet, the gains were the most dramatic among female and minority engineers — four times larger than for white males. In contrast, within the untreated (or control) group, female and minority engineers' performance tended to deteriorate on several dimensions over the next several months.

According to Kelley, there are two main reasons why the productivity levels of female and minority engineers disproportionately soared after receiving the star performer training. First, these engineers undertook proactive measures to break into knowledge networks that were based on expertise rather than gender or race (success factor #2). Second, with the benefits of the training, these engineers engaged in better self-management to deal with incoming requests from coworkers (success factor #3). In many cases, the purpose of the requests was to showcase the company's diversity rather than tapping into the engineer's developed skill set. When played out over several iterations, the disparities between white males and female and diverse knowledge workers gets larger, not because of a gap in innate ability, but because of a systems failure in training and feedback.

## **(5) Coaching and Mentoring**

The fifth and final component in my model is coaching and mentoring. A strong coach and mentor is often the vehicle through which a young lawyer receives developmentally rich work experience (3) and high-quality training and feedback (4). Yet coaching and mentoring is its own freestanding component because when it is done well it becomes an intense personal connection where talented professionals choose to allocate their valuable time and resources toward the

success of others. Conversely, understanding the nature of the investment being made, the person being mentored experiences a mixture of heightened motivation and gratitude that enables him or her to persevere through virtually any professional hardship in order to reach long-term goals.

One of the best examples of the power of mentorship is New York City business lawyer Walter Carter, who served as a mentor to many of the leading corporate lawyers of the early 20th century. Carter's accomplishments on this front were chronicled in a 1954 book entitled *Walter S. Carter: Collector of Young Masters*. According to the book's author, Otto Koegel, Carter's gift was spotting promising young talent and bringing them along as corporate lawyers who were capable of counseling executives of large financial and industry enterprises.

An appendix at the back of Koegel's book is a folded poster with a family tree of Carter's lawyer progeny. One of the first nodes on the family tree is Paul Cravath, who worked for Carter as a junior lawyer. The subsequent branches document Cravath's departure and movement to a firm that would later become Cravath, Swaine & Moore, where Cravath designed and implemented the "Cravath system." According to the firm's history, the Cravath system is largely credited with the firm's eventual leadership position among Wall Street firms. The firm history also cites Walter Carter's training principles as the basis for the system. Other branches on the Carter family tree connect founders or leaders at many familiar powerhouse firms of the 21st century, including Milbank Tweed, Willkie Farr, Cadwalader, Shearman & Sterling, and Hughes Hubbard.

I have also observed something similar to Carter's impact on future leading lawyers, albeit within the context of a government agency. Colleagues in the securities bar have observed the phenomenon of "Sporkin's kids," referring to the many influential lawyers who worked under Stanley Sporkin during his long and distinguished tenure at the Securities and Exchange Commission (SEC). Many of Sporkin's SEC protégés lacked the pedigree of an elite law school, yet they went on to become some of the most sought after and influential securities litigation lawyers of their generation. They include Edward Herlihy of Wachtell Lipton (George Washington Law), William McLucas of WilmerHale (Temple Law), and Ralph Ferrera of Proskauer (Cincinnati Law).

After two decades at the SEC, Sporkin became general counsel of the CIA and then a prominent federal judge. In preparation for writing this article, I contacted Judge Sporkin to ask him about this track record of mentorship. He commented that his philosophy was to look for intelligent young lawyers who would approach their jobs "with enthusiasm." In Sporkin's view, the law school attended was a poor proxy for these intangibles (Sporkin himself attended Yale). Further, according to Sporkin, it was critical that there be values alignment between the young lawyer and the mission of the agency. Otherwise, the lawyer could not keep up with the demands of working in his office. (Compare Sporkin's observations to the Motivation factor outlined in this article's five-factor model.) Judge Sporkin expressed gratitude for the lack of bureaucracy in the 1960s, 1970s, and 1980s, which enabled him to hire so much raw talent according to his own criteria. He related the story of meeting a young Ralph Ferrera, who pleaded with Sporkin for an opportunity to work at the agency. Sporkin lacked the budget to hire him, so Ferrera worked for free until a formal staff position became open. The rest, as they say, is history.



In my experience, law firms undervalue the importance of coaching and mentorship. Carter and Sporkin had the power to make these investments on their own. Yet, today's modern law firm emphasizes the production of revenues. The cost of nonbillable time can be readily calculated; the same cannot be said, however, about the value of nonbillable time. Partners who have given little thought to the power of professional development are most likely to resist large investments. They lack the systems perspective of Paul Cravath. I have studied lawyer development for over a decade. I think these partners are trading dollars for pennies.

## **Conclusion**

The purpose of this article is to create a roadmap for solving the legal profession's longstanding diversity problem. The solution is to end the moral handwringing and to create a system for selecting and developing lawyers. Yes, it will be expensive in time, money, and political capital, but not nearly as costly as wasting raw human potential. Glory, and possibly organizational riches, will accrue to the law firm leaders and general counsel who are brave enough and wise enough to demand that we go down this road. The time has come to fix this problem once and for all.

## **Chapter 7: Measuring Employee Health: Balancing Costs, Control, and Controversy / *Jennifer Bennett Shinall, Assistant Professor of Law at Vanderbilt University Law School\****

In spite of the increasing presence of improved, lower-cost technology to measure health metrics, employers have made little use of employee health data collected in the context of a workplace wellness program. For many years, and especially since 2000, employers have been discouraged from analyzing these data by Equal Employment Opportunity Commission (EEOC) regulations and guidance under the Americans with Disabilities Act. Out of privacy and confidentiality concerns for disabled workers, the former agency guidance strongly discouraged the analysis of employee health data and prohibited the matching of these data to personnel records. In May 2016, however, the EEOC released new regulations and interpretive guidance that opened the door to employer health data analysis. Contrary to many disability rights advocates and scholars, this article argues that collecting employee health data—and in particular, health data matched to personnel records—has the potential to benefit workers affected by health conditions. The present dearth of employee health data matched to personnel records has done nothing to ameliorate the stereotypes that continue to exist about the abilities of disabled workers. Only data analysis of employee health data can advance current understandings of the relationship (or lack thereof) between employee health and productivity.

### **Introduction**

Wellness programs are the panacea for all workplace ills—at least, according to popular belief.

Workplace wellness programs vary widely in scope, structure, and content. They may require a

---

\* 131 21<sup>st</sup> Avenue South, Nashville, TN 37203. (615) 343-9622. [jennifer.shinall@vanderbilt.edu](mailto:jennifer.shinall@vanderbilt.edu). The author extends a special thanks to participants in the 2016 Research Group on Legal Diversity's Metrics, Diversity, and Law Conference for their helpful feedback.

once-a-year checkup as part of open enrollment, or they may have an everyday presence within the work environment. They may or may not rely on an incentive structure, and incentives may be financial or nonfinancial in nature. They may include tobacco cessation, medical, nutritional, or exercise components.<sup>114</sup> Despite the broad variation between programs, they are widely touted as a way to decrease employer insurance costs, reduce employee absences, increase productivity, improve worker morale, and promote employee camaraderie.<sup>115</sup> As such, workplace wellness has ballooned into a six billion dollar industry.<sup>116</sup> Such promises of an improved bottom line likely explain why so many employers have jumped on the workplace wellness bandwagon. According to the 2015 Kaiser Employer Health Benefits Survey, four out of five employers with more than 200 workers offer a wellness program, and approximately half of smaller employers offer one.<sup>117</sup> Yet popular belief does not match the available evidence on workplace wellness programs.

Despite the fact that many business leaders have bought into the existence of a causal link

---

<sup>114</sup> For a discussion of the extensive variation in wellness programs, and the lack of empirical evidence surrounding them, see Jessica L. Roberts & Elizabeth Weeks Leonard, *What Is (and Isn't) Healthism?*, 50 GA. L. REV. 1, 70-72 (2016).

<sup>115</sup> See, e.g., Jacqui Cook, *The Benefits of Workplace Wellness Programs*, Center for Association Leadership, American Society of Association Executives (2016), [https://www.asaecenter.org/resources/articles/an\\_magazine/2012/july/the-benefits-of-workplace-wellness-programs](https://www.asaecenter.org/resources/articles/an_magazine/2012/july/the-benefits-of-workplace-wellness-programs); see also Rival Health, *Solutions: Employers* (2016), <http://www.rivalhealth.com/corporate-wellness-programs-and-employee-engagement-activities> (marketing its corporate wellness programs as a way to “[l]ower long-term health care costs,” improve “focus” and “productivity,” “increase retention and employee satisfaction,” and “[d]rive internal competition to break the performance plateau”); TotalCare Wellness, *TotalCare Wellness Programs Reduce Health Care Costs and Improve Productivity* (2016), ESI GROUP, available at <http://www.totalcarewellness.com/wp-content/uploads/2016/01/TCW-Sales.pdf> (marketing its corporate wellness programs not only as able to “reduce health care costs and improve productivity,” but also as able to “reduce absenteeism, presenteeism and the risk of on-the-job injuries”).

<sup>116</sup> RAND Corporation, *Do Workplace Wellness Programs Save Employers Money*, RAND RESEARCH BRIEF (2014), available at [http://www.rand.org/content/dam/rand/pubs/research\\_briefs/RB9700/RB9744/RAND\\_RB9744.pdf](http://www.rand.org/content/dam/rand/pubs/research_briefs/RB9700/RB9744/RAND_RB9744.pdf).

<sup>117</sup> See Kaiser Family Foundation & Health Research & Educational Trust, *Employer Health Benefits: 2015 Annual Survey* (September 2015), available at <http://files.kff.org/attachment/report-2015-employer-health-benefits-survey> (finding that 81 percent of large employers have a wellness program, and 49 percent of small employers have one).

between wellness programs and improvements in productivity and costs,<sup>118</sup> the actual research on this link is limited and, at best, inconclusive.<sup>119</sup> Improvement in any given health metric does not necessarily improve an employee's productivity, nor does it automatically reduce employers' costs. In fact, one of the most rigorous studies to date on workplace wellness, the 2014 RAND Wellness Programs study, concluded,

The press and trade publications strongly endorse workplace wellness programs as a good investment for employers, and even the normally skeptical academic world has joined the bandwagon. . . . [Our] Study, which included almost 600,000 employees at seven employers, showed that wellness programs are having little if any immediate effects ....<sup>120</sup>

The widespread implementation of workplace wellness programs juxtaposed with the scarcity of evidence on their efficacy raises the question: Why are there so many potential wellness programs to study, but so few actually studied? These programs are costly,<sup>121</sup> and since many

---

<sup>118</sup> Society for Human Resources Management, *Health Productivity, and Performance Study Committee, Exploring the Value Proposition for Workplace Health* (February 2015), available at [https://www.shrm.org/ResourcesAndTools/hr-topics/benefits/Documents/HPP-Business-Leader-Survey-Full-Report\\_FINAL.pdf](https://www.shrm.org/ResourcesAndTools/hr-topics/benefits/Documents/HPP-Business-Leader-Survey-Full-Report_FINAL.pdf) (finding that “[w]hen business leaders were asked directly about the top organizational priorities influenced by employee health, productivity and performance were most often listed” and also that over 90 percent of business leaders believed that employee health affected workplace productivity and performance).

<sup>119</sup> *Accord* Jessica L. Roberts & Elizabeth Weeks Leonard, *What Is (and Isn't) Healthism?*, 50 GA. L. REV. 1, 71 (2016) (“While an employee wellness program that requires medical testing could encourage healthy decisionmaking and perhaps facilitate choice (depending on whether it offers new options to participants that would otherwise be unavailable), it remains to be seen whether such programs actually lower costs, reduce risks, or produce better health outcomes.”); see also Al Lewis et al., *Workplace Wellness Produces No Savings*, HEALTH AFFAIRS BLOG (Nov. 25, 2014), <http://healthaffairs.org/blog/2014/11/25/workplace-wellness-produces-no-savings/> (concluding that “wellness programs produce a return-on-investment . . . of less than 1-to-1 savings cost”).

<sup>120</sup> RAND Corporation, *Do Workplace Wellness Programs Save Employers Money*, RAND Research Brief (2014), available at [http://www.rand.org/content/dam/rand/pubs/research\\_briefs/RB9700/RB9744/RAND\\_RB9744.pdf](http://www.rand.org/content/dam/rand/pubs/research_briefs/RB9700/RB9744/RAND_RB9744.pdf).

<sup>121</sup> Estimated minimum costs of a wellness program range from \$100 to \$400 per employee annually. See Donna Hughes, *The Cost of Wellness: A WELCOA Expert Interview* (2009), WELLNESS COUNCIL OF AMERICA, available at

employers admittedly introduce these programs to improve their bottom line, why are they not following up to investigate whether any program benefits exist (and if so, why these benefits exist and whether they are justified by program costs)? The scarcity of program evaluation among employers with wellness programs becomes particularly curious after considering that data collection is cheaper and easier than ever before. The recent influx of improved, lower-cost health technology—from Fitbits to blood pressure monitors to Apple Watches—have made measuring employee outcomes relatively simple and inexpensive. At the very least, the scarcity of academic research<sup>122</sup> on a topic as pervasive as workplace wellness programs becomes curious.

A major hindrance discouraging wellness program evaluation, whether by employers or academic researchers, has come from a surprising source: federal law. The Americans with Disabilities Act (ADA)—and more specifically, the regulations and guidance issued by the Equal Employment Opportunity Commission (EEOC) surrounding the Act—has rendered many employers wary of collecting and matching the necessary information to evaluate their wellness programs fully. Most notably, the agency issued ADA enforcement guidance in 2000 that severely limited the types of wellness programs that employers could implement, the collection

---

<https://www.scribd.com/document/34732477/The-Cost-of-Wellness-A-WELCOA-Expert-Interview-With-Dr-Ron-Goetzel>; see also Wellsource, *How Much Does a Good Wellness Program Cost?* (2011), *available at* [http://wellsource.com/wp-content/uploads/2015/08/How\\_Much\\_Should\\_a\\_Wellness\\_Program\\_Cost.pdf](http://wellsource.com/wp-content/uploads/2015/08/How_Much_Should_a_Wellness_Program_Cost.pdf). Some employers spend over \$500 per employee annually on wellness programs. See Austin Frakt & Aaron E. Carroll, *Do Workplace Wellness Programs Work? Usually Not*, N.Y. TIMES, Sep. 11, 2014, *available at* <http://www.nytimes.com/2014/09/12/upshot/do-workplace-wellness-programs-work-usually-not.html> (“Medium-to-large employers spent an average of \$521 per employee on wellness programs last year, double the amount they spent five years ago, according to a February report by Fidelity Investments and the National Business Group on Health.”).

<sup>122</sup> See *infra* Part III.

of data on existing programs, and the analysis that could be performed on any collected data.<sup>123</sup> The purpose of the EEOC's regulations was noble—to prevent employers from using wellness programs (and wellness program evaluation) to discriminate against disabled individuals.<sup>124</sup> Yet the consequences of these regulations on the development, assessment, and improvement of wellness programs have been injurious. Fear of liability has discouraged employers from compiling the necessary data to study wellness program effects, which, in turn, has prevented both internal analysis by employers and external analysis by researchers.

Federal regulation of wellness programs may be on the brink of a regime change, however—a regime change that this article will argue is a significant improvement. In 2015, the EEOC issued a series of new proposed rules regarding the ADA and wellness programs, designed to promote the development and expansion of workplace wellness programs. The rules were issued under the authority of the 2010 Patient Protection and Affordable Care Act (ACA), which, among other things, encouraged the implementation and expansion of wellness programs. Although finalized with little modification in May 2016, the new regulations and guidance are already in jeopardy with the incoming federal executive and legislative branches and their promised dismantling of the ACA.<sup>125</sup> This article will argue that—in spite of the impending demolition of some aspects of

---

<sup>123</sup> See 29 C.F.R. § 1630.14 *et seq.* (2000).

<sup>124</sup> See, e.g., Equal Employment Opportunity Commission, *Disability-Related Inquiries and Medical Examinations of Employees under the Americans with Disabilities Act (ADA)* (July 2000), available at <https://www.eeoc.gov/policy/docs/guidance-inquiries.html> (expressing concern with ever asking “applicants and employees to provide information concerning their physical and/or mental condition . . . [since t]his information often was used to exclude and otherwise discriminate against individuals with disabilities—particularly nonvisible disabilities” in the past).

<sup>125</sup> See, e.g., Robert Pear, Jennifer Steinhauer, & Thomas Kaplan, *G.O.P. Plans Immediate Repeal of Health Law, Then a Delay*, N.Y. TIMES, Dec. 2, 2016, at A1.

the ACA, such as the healthcare marketplace—the new workplace wellness regulations should persist, particularly since they do not mandate new costs for either employers or the government.

The 2016 workplace wellness regulations contained two major policy changes from the prior regime. First, the regulations expanded the types of incentives that employers could use to encourage meaningful participation in their wellness programs. Whereas the EEOC formerly restricted employers to the use of carrots, the EEOC now proposed allowing employers the use of sticks—and substantial sticks, at that. In fact, the rules allowed for employers to impose financial or in-kind penalties of up to 30 percent of the cost of employee-only insurance coverage. The EEOC then clarified that penalties were legally permissible in both participatory and health-contingent programs. In other words, employers could not only penalize employees for failing to participate in a wellness program, they could also penalize employees for failing to meet certain health-related benchmarks. Second, the regulations opened the door to expanding the collection and usage of data by employers with regard to wellness programs, which, in turn, opened the door to a more accurate and precise understanding on how these programs impact employee health, employee productivity, and ultimately, employers' bottom line.<sup>126</sup>

Yet as soon as these rules were proposed, they immediately met resistance and backlash from the disability community. In multiple letters submitted to the EEOC during the notice-and-comment period, representatives from a broad range of groups with an interest in the ADA criticized the proposed rules as a threat to the privacy and employment rights of disabled individuals. Even some disability scholars submitted or signed onto comments, further echoing the concerns that

---

<sup>126</sup> See *infra* Part IV.

the proposed rules would increase discrimination against the disabled. The thrust of their argument was that, under the proposed rules, employers could (1) gain access to health-related information of disabled individuals that was irrelevant to the workplace, and (2) set health-related benchmarks within their wellness programs so that they were unattainable by disabled individuals, or use the newly acquired data to otherwise discriminate against disabled individuals. In their view, any benefits from increasing the scope of permissible wellness programs came at the expense of the disabled community, and they beseeched the agency to reconsider its position.<sup>127</sup>

Despite these pleas, the EEOC did not change its position when the regulations were finalized in May 2016, and as this article will argue, should not change its position under the incoming presidential administration. While members of the disabled community have viewed the agency's position as a blow to employee rights under the ADA, this article will argue that their concerns have been overstated. The agency has already built several safeguards with regard to disabled workers into the new regulations and guidance; a few simple clarifications or additions, discussed in Part V of this article, might further ensure that these workers are protected. Moreover, critics of the new regulations have failed to recognize their potential to combat disability discrimination in the workplace. The prior regulations not only prevented program evaluation by employers, but also academic researchers, because they imposed a significant barrier to compilation of the necessary data to evaluate wellness programs rigorously. Without such evaluation, employers could persist relying on harmful assumptions and anecdotes regarding disabled employees' productivity, virtually unchecked by contrary empirical evidence.

---

<sup>127</sup> See *infra* Part V.



Better data are the key to better empirical research, which could serve as an ally for advocates from the disability community in correcting both employers' and the public's understanding regarding the effects of health and disability on employee productivity.

Using the example of obesity in the workplace, this article will demonstrate how the discouragement of program evaluation under the old regulations has harmed workers with health conditions—and how the new regulations may mitigate at least some of this harm. Part I will discuss the questions surrounding the effects of wellness programs that have persisted under the old regulations, due to lack of data. In the absence of data collection and analysis, advocates are left with sparse empirical evidence to support their position of increasing disability rights within the workplace, and arguably, without much hope of amending widespread assumptions regarding the negative effect of health on employee productivity—assumptions that may be particularly common (and harmful) for employees with visible health conditions. Part II will discuss how the EEOC's former regulations and enforcement guidance discouraged employers from collecting and analyzing data on their wellness programs. Part III illustrates the problem with the agency's former regulations and guidance by way of an example; examining the plight of obese workers, this part reveals how the very regime intended to assist workers with visible health conditions might instead be harming them. Part IV turns to the new—albeit, endangered—EEOC regulations, highlighting the significant changes for wellness program evaluation. Part V weighs the advantages of increased data collection and analysis against the privacy, equity, and venue concerns raised by disability rights advocates and scholars. In particular, this part will not only point out how the new regulations already protect against many of the concerns raised but also suggest simple additions to the EEOC's policy—perhaps in the form of additional enforcement

guidance—that might more thoroughly quell these concerns. The article concludes by arguing that regulations promoting the improvement of workplace wellness programs are in the best interest of both employers and employees.

### **Unanswered Questions about Workplace Wellness Programs**

The typical argument in favor of adopting a workplace wellness program focuses on the financial gains to employers: Wellness programs make employees healthier, which can reduce the cost of employer-provided insurance.<sup>128</sup> More fundamentally, healthier workers are better workers, which can increase employer revenue.<sup>129</sup> The take-up of this argument is quite common among private employers, with more than 67 percent of them identifying wellness programs as an effective way to improve their bottom line.<sup>130</sup> And yet, current research is, at best, inconclusive

---

<sup>128</sup> See, e.g., Jim Purcell, *Meet the Wellness Programs that Save Companies Money*, HARV. BUS. REV., Apr. 20, 2016, <http://hbr.org/2016/04/meet-the-wellness-programs-that-save-companies-money> (“[W]e must first go beyond unduly narrow interpretations of . . . [returns on investment] to understand how properly designed wellness programs can help employers lower health care costs while providing other types of cost savings and competitive advantages.”).

<sup>129</sup> See Ron Z. Goetzel et al., *Do Workplace Health Promotion (Wellness) Programs Work?*, 56 J. OCC. & ENVIRON. MED. 927, 927-34 (Sep. 2014) (“What do workplace programs aim to accomplish? If we were to gather key executives at a company who are informed about health care and ask them what they expect a workplace health promotion program to achieve, you would likely hear a range of responses, . . . [including,] ‘Workers will perform at higher levels—they will be happier, have more energy, produce better results for our company.’”); see also Suzanne Lucas, *Healthy Employees Make Happy Employees*, INC., Jan. 22, 2014, available at <http://www.inc.com/suzanne-lucas/healthy-employees-make-happy-employees.html> (arguing that “healthy employees make happy employees” and “[h]appier employees make better workers”); Ann Carrns, *Study Raises Questions for Employer Wellness Programs*, N.Y. TIMES, Jan. 7, 2014, at B3 (“[Wellness] programs have become increasingly popular, as companies aim to lower their medical costs and lift productivity by promoting healthier behavior among workers.”).

<sup>130</sup> See Kaiser Family Foundation & Health Research & Educational Trust, *Employer Health Benefits: 2013 Annual Survey* (September 2013), available at <https://kaiserfamilyfoundation.files.wordpress.com/2013/08/8465-employer-health-benefits-20131.pdf> (finding that 67 percent of employers identified wellness programs as a “very effective” or “somewhat effective” way of decreasing insurance costs and that wellness programs were the most commonly identified method of effective cost reduction among employers). Although this survey assesses the motivations of private employers only, presumably public employers are similarly motivated, given the widespread adoption of wellness programs in both the private and public sectors. See SOEREN MATTKE ET AL., WORKPLACE WELLNESS PROGRAMS STUDY: FINAL REPORT 124, RAND HEALTH, [http://www.rand.org/content/dam/rand/pubs/research\\_reports/RR200/RR254/RAND\\_RR254.pdf](http://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR254/RAND_RR254.pdf) (“This practice is not confined to the private sector; a recent review suggested that states are beginning to offer incentives for participation or goal attainment for their workers”).

as to whether any part of this argument is actually true. First, it is not clear that wellness programs actually achieve their first objective of making employees healthier. A number of empirical studies have suggested that in spite of the effort and expense involved in implementing these programs, they do little to nothing to improve health outcomes. For example, a 2013 study of a large employer concluded that its company wellness program had no effect on employees' inpatient admissions, emergency room visits, or per-month insurance costs.<sup>131</sup> Similarly, another recent study of a school district's wellness program concluded that participants actually filed more medical claims, on average, than did nonparticipants, although the average claim cost for participants was lower than the average claim cost for nonparticipants.<sup>132</sup>

Perhaps wellness programs' shortcomings in improving worker health is most persuasively documented in a 2014 meta-analysis of prior workplace wellness studies, which found that the highest quality empirical studies estimated returns on employers' investment in these programs as very close to zero, and the few randomized control trials that existed actually estimated a negative return on employers' investment.<sup>133</sup> Indeed, as some health policy scholars have noted,

What research exists on wellness programs does not support . . . optimism. This is, in part, because most studies of wellness programs are of poor quality, using weak methods that suggest that wellness programs are associated with lower savings, but don't prove

---

<sup>131</sup> Hangsheng Liu et al., *Do Workplace Wellness Programs Reduce Medical Costs? Evidence from a Fortune 500 Company*, INQUIRY 150, 150-58 (2013).

<sup>132</sup> Ray M. Merrill & James D. LeCheminant, *Medical Cost Analysis of a School District Worksite Wellness Program*, 3 PREV. MED. REP. 159, 159-65 (June 2016).

<sup>133</sup> Siyan Baxter et al., *The Relationship Between Return on Investment and Quality of Study Methodology in Workplace Health Promotion Programs*, 28 AM. J. HEALTH PROM. 347, 347-63 (July/August 2014).

causation. Or they consider only short-term effects that aren't likely to be sustained.

Many such studies are written by the wellness industry itself. More rigorous studies tend to find that wellness programs don't save money and, with few exceptions, do not appreciably improve health. This is often because additional health screenings built into the programs encourage overuse of unnecessary care, pushing spending higher without improving health.<sup>134</sup>

Moreover, even assuming that some wellness programs can improve employee health, it remains clear that not all wellness programs improve employee health—particularly given the widespread variation in scope, structure, and content of programs included within the broad umbrella term of workplace wellness. A 2014 study of the PepsiCo workplace wellness program, for instance, compared the health effects of the program's two components, a lifestyle management component (encouraging all workers to make healthier lifestyle choices in terms of weight, fitness, nutrition, stress, and smoking) and a disease management component (assisting workers already diagnosed with a chronic disease to manage their condition). Even though the study found a net decrease in health care costs as a result of the disease management component, no such savings was found with respect to the lifestyle management component. The study concluded, consequently, that even though “[w]orkplace wellness programs have the potential to

---

<sup>134</sup> Austin Frakt & Aaron E. Carroll, *Do Workplace Wellness Programs Work? Usually Not*, N.Y. TIMES, Sep. 11, 2014, available at <http://www.nytimes.com/2014/09/12/upshot/do-workplace-wellness-programs-work-usually-not.html>. Accord Mitesh Patel et al., *Premium-Based Financial Incentives Did Not Promote Workplace Weight-Loss in a 2013-2015 Study*, 35 HEALTH AFF. 71, 71-79 (Jan. 2016) (finding the threat of increased health insurance premiums did not incentivize overweight subjects to lose weight); Jill R. Horwitz et al., *Wellness Incentives in the Workplace: Cost Savings through Cost Shifting to Unhealthy Workers*, 32 HEALTH AFF. 468, 468-76 (Mar. 2013) (finding “little evidence that such programs can easily save costs through health improvement without being discriminatory”).

reduce health risks and to delay or avoid the onset of chronic diseases . . . employers and policy makers should not take for granted” that all programs will actually do so.<sup>135</sup> Other health policy researchers have echoed this conclusion, noting that “[e]mployers may misunderstand the research if they think that just any wellness program, by itself, is the surest route to reducing overall health care spending.”<sup>136</sup>

Along these lines, the findings in two of the highest quality existing studies have reinforced health researchers’ skepticism over the ability of wellness programs, at least in some forms, to achieve their health aims. A 2013 study by health economists John Cawley and Joshua Price analyzed data from a third-party company that administers wellness programs for employers.<sup>137</sup> After examining weight loss outcomes in four different types of voluntary wellness programs,<sup>138</sup> Cawley and Price found that only programs that incorporated financial penalties for failure to lose weight actually led to weight loss. Even then, the average weight loss among overweight and obese individuals in penalty programs was less than ten pounds, and all programs had high attrition rates. The authors concluded by emphasizing the greater need for research, even within

---

<sup>135</sup> John P. Caloyeras et al., *Managing Manifest Diseases, But Not Health Risks, Saved PepsiCo Money Over Seven Years*, 33 HEALTH AFF. 124, 124-31 (Jan. 2014).

<sup>136</sup> Austin Frakt & Aaron E. Carroll, *Do Workplace Wellness Programs Work? Usually Not*, N.Y. TIMES, Sep. 11, 2014, available at <http://www.nytimes.com/2014/09/12/upshot/do-workplace-wellness-programs-work-usually-not.html>.

<sup>137</sup> John Cawley & Joshua A. Price, *A Case Study of a Workplace Wellness Program that Offers Financial Incentives for Weight Loss*, 32 J. HEALTH ECON. 794, 794-803 (2013).

<sup>138</sup> Under the statutory text of the ADA, all workplace wellness programs must be voluntary; mandatory wellness programs violate the Act’s confidentiality provisions. Although this prohibition of mandatory wellness programs has remained constant over the course of the ADA’s history, what has changed is the EEOC’s definition of voluntary. Specifically, as this article will discuss in Part II, the EEOC’s new regulations broaden the definition of voluntary, thus rendering fewer wellness programs an impermissible mandatory program. *See generally* Americans with Disabilities Act of 1990, *Pub. L. 101-336*, § 102(d)(4)(b)-(c).

the successful program, on penalty amounts, peer effects, and long-term weight regain.<sup>139</sup> In 2014, another group of health researchers conducted a broad review of prior evidence on wellness programs. Noting the large number of poorly designed programs—and the prevalence of unrealistic expectations surrounding them—the researchers emphasized a greater need for program evaluation and evidence-based design of workplace programs. Moreover, even in the best wellness programs, the researchers cautioned, “Program success depends on the goals of the program, program design and implementation, and importantly how the program is evaluated. If the only expectation is that the sponsoring organization will ‘make money’ (i.e., achieve a financial gain) . . . , then implementing a best practice health promotion program may not be worth the effort.”<sup>140</sup>

The absence of a clear understanding about the types of wellness programs that improve employee health generally—or even specific health metrics such as body mass index, blood pressure, and blood glucose—undercuts the popular belief that introducing a wellness program will necessarily be financially worthwhile for an employer. Even in the case of insurance companies that reduce premiums automatically for employers who implement a wellness program of any kind (regardless of its scope, structure, content, or efficacy), it remains unclear whether premium savings outweigh implementation costs in the long run. Given recent evidence that wellness program participants file more claims with health insurance companies than do

---

<sup>139</sup> See John Cawley & Joshua A. Price, *A Case Study of a Workplace Wellness Program that Offers Financial Incentives for Weight Loss*, 32 J. HEALTH ECON. 794, 794-803 (2013).

<sup>140</sup> See Ron Z. Goetzel et al., *Do Workplace Health Promotion (Wellness) Programs Work?*, 56 J. OCC. & ENVIRON. MED. 927, 933 (Sep. 2014).

nonparticipants,<sup>141</sup> any initial premium savings that employers experience after program implementation may erode over time. Moreover, since the few studies that exist on wellness programs cast doubt on their ability to improve specific, measureable health metrics, their ability to improve overall employee health—let alone improve worker productivity—is questionable. Indeed, even though researchers have conducted some limited investigations of the link between wellness programs and measureable health metrics, entirely absent from the existing literature is an investigation of the link between wellness programs and employee productivity. Both the absence of research on the wellness program-productivity connection and the insufficient research on the wellness program-health connection are largely due to lack of available data. And the lack of available data has been driven, at least in part, by legal policy—in particular, by legal policy surrounding the ADA.

### **Who’s Afraid of the ADA?**

Over the past fifteen years, the EEOC’s stringent position towards collecting, storing, and evaluating wellness program data has understandably frightened many employers away from attempting it.<sup>142</sup> Even though a significant number of employers now have wellness programs for

---

<sup>141</sup> See Ray M. Merrill & James D. LeCheminant, *Medical Cost Analysis of a School District Worksite Wellness Program*, 3 PREV. MED. REP. 159, 159-65 (June 2016).

<sup>142</sup> Employers advised by legal counsel have certainly been discouraged from collecting and analyzing wellness data; legal defense firm websites are replete with articles that caution employers about wellness program data. See, e.g., Ilyse Schuman, Russell Chapman, & Michelle Thomas, *EEOC Issues Long-Awaited Proposed Rule on Employer Wellness Programs* (May 2015), LITTLER MENDELSON, <https://www.littler.com/publication-press/publication/eec-issues-long-awaited-proposed-rule-employer-wellness-programs> (“The lack of guidance by the agency, in conjunction with the EEOC enforcement activity, has created a quagmire for employers seeking to enhance the use of effective wellness programs . . . .”); Keith A. Markel & Wendy M. Fiel, *The Importance of Keeping Employer-Sponsored Wellness Programs ‘Voluntary’* (October 2014), INSIDE COUNSEL, <http://www.insidecounsel.com/2014/10/10/the-importance-of-keeping-employer-sponsored-welln> (“The importance of implementing wellness programs that are ‘voluntary’ [under the ADA] cannot be overemphasized.”); Kevin Kelly, *Legal Risks Behind Workplace Wellness Programs* (May 2013), LAW360,

their employees, the agency has historically made its position known through enforcement guidance. In fact, the EEOC has never issued any formal regulations specific to wellness programs until July 2016. Until this year, the regulations governing wellness programs and confidentiality under the ADA have been the ones issued back in 1991, only a year after the passage of the ADA. Under these former regulations, the direction offered to employers with regard to wellness programs was quite limited:

(d) Other acceptable examinations and inquiries. A covered entity may conduct voluntary medical examinations and activities, including voluntary medical histories, which are part of an employee health program available to employees at the work site.

(1) Information obtained under paragraph (d) of this section regarding the medical condition or history of any employee shall be collected and maintained on separate forms and in separate medical files and be treated as a confidential medical record . . . .<sup>143</sup>

In reality, the above language in the 1991 regulations with regard to wellness programs is identical to the vague language already present in the statute.<sup>144</sup> When the ADA passed in 1990 and when the regulations were issued in 1991, the popularity of wellness programs was a mere fraction of what it is today<sup>145</sup>—perhaps rendering this lack of detail on wellness programs an

---

<http://www.law360.com/articles/441487/legal-risks-behind-workplace-wellness-programs> (arguing that “employers should exercise caution in implementing wellness programs” to avoid violating the ADA and other federal statutes).

<sup>143</sup> 56 FR 35726-01 (July 26, 1991).

<sup>144</sup> See Americans with Disabilities Act of 1990, *Pub. L. 101-336*, § 102(d)(4)(b)-(c).

<sup>145</sup> The growth in the workplace wellness industry has become particularly explosive over the past decade, due to the attention devoted to wellness by President Obama’s administration and the promotion of wellness within the ACA.



understandable omission from the statutory and regulatory texts. Yet in the absence of regulatory or statutory guidance, determining the precise meaning of terms like “voluntary” and “separate” has proven particularly unwieldy for employers trying to establish, let alone evaluate, a wellness program.

By the turn of the new millennium, wellness programs were already becoming much more visible in the workplace,<sup>146</sup> which led the EEOC to take clarifying actions. Although the EEOC did not issue a formal regulation with regard to workplace wellness programs, the agency did issue enforcement guidance in July 2000. Intended to clarify the agency’s stance on health-related inquiries by employers both within and outside the scope of a wellness program, the 2000 guidance clarified the agency’s view of much of the previously ambiguous text within the 1991 regulation. With respect to wellness programs, the agency noted:

The ADA allows employers to conduct voluntary medical examinations and activities, including voluntary medical histories, which are part of an employee health program

---

*Accord Judy Peres, Workplace Wellness Programs Popular, But Do They Improve Health?*, CHI. TRIBUNE, Dec. 13, 2014, <http://www.chicagotribune.com/news/ct-workplace-wellness-met-20141212-story.html> (“Workplace wellness programs, one of the pillars of the national health care overhaul, are increasingly popular among employers who believe they can help control costs and workers who see them as a perk.”); Al Lewis, *Will Workplace Wellness Actually Make Workers Healthy?* (Feb. 2013), FAST COMPANY, <http://www.fastcoexist.com/3026401/will-workplace-wellness-actually-make-workers-healthy> (“Over the last decade—and especially [since the enshrinement of wellness](#) in the Affordable Care Act—American workers have found themselves subjected in ever-increasing numbers to workplace wellness programs. While originally these programs were confined to large organizations, that market is now saturated.”)

<sup>146</sup> See e.g., Shelley K. Schwartz, *Wellness Programs Evolve*, CNN, May 11, 1999, [http://money.cnn.com/1999/05/11/life/q\\_health/](http://money.cnn.com/1999/05/11/life/q_health/) (discussing the “hundreds of large- and mid-sized corporations adding outreach components to their workplace wellness initiatives”); Miriam Sims, *Wellness Programs Must Expand Focus*, DENVER BUS. J., Jan. 17, 1999, <http://www.bizjournals.com/denver/stories/1999/01/18/smallb5.html> (attributing the increasing popularity of wellness programs to increasing health care costs for employers in the 1980s and 1990s).

without having to show that they are job-related and consistent with business necessity, as long as any medical records acquired as part of the wellness program are kept confidential and separate from personnel records. These programs often include blood pressure screening, cholesterol testing, glaucoma testing, and cancer detection screening. Employees may be asked disability-related questions and may be given medical examinations pursuant to such voluntary wellness programs. A wellness program is “voluntary” as long as an employer neither requires participation nor penalizes employees who do not participate.<sup>147</sup>

Of course, the above enforcement guidance never enjoyed the same level of judicial deference as either the 1990 statute or the 1991 regulation.<sup>148</sup> Thus, employers were not necessarily obliged to follow the EEOC’s guidance, although risk-averse employers were certainly well advised to follow it,<sup>149</sup> particularly after a 2014 flurry of EEOC enforcement actions filed against employers

---

<sup>147</sup> Equal Employment Opportunity Commission, *Disability-Related Inquiries and Medical Examinations of Employees under the Americans with Disabilities Act (ADA)* (July 2000), available at <https://www.eeoc.gov/policy/docs/guidance-inquiries.html>.

<sup>148</sup> Courts grant a high level of deference to EEOC regulations issued as a result of the formal notice-and-comment rulemaking process. *See Chevron v. Natural Resources Defense Council*, 467 U.S. 837, 843-44 (1984) (deferring to a federal agency’s interpretation of a statute when “Congress has not directly addressed the precise question at issue” and “the agency’s answer is based on a permissible construction of the statute”). But courts, at best, grant a lower level of deference to EEOC enforcement guidance, which does not undergo the notice-and-comment rulemaking process. *See EEOC v. Arabian American Oil Co.*, 499 U.S. 244, 256-58 (1991) (finding EEOC guidelines not entitled to *Chevron* deference).

<sup>149</sup> *Accord* Keith A. Markel & Wendy M. Fiel, *The Importance of Keeping Employer-Sponsored Wellness Programs ‘Voluntary’* (October 2014), INSIDE COUNSEL, <http://www.insidecounsel.com/2014/10/10/the-importance-of-keeping-employer-sponsored-welln> (warning employers to “[m]ake sure wellness programs are legally fit under the ADA in order to avoid litigation”); Kevin Kelly, *Legal Risks Behind Workplace Wellness Programs* (May 2013), LAW360, <http://www.law360.com/articles/441487/legal-risks-behind-workplace-wellness-programs> (discussing the perils of navigating the “voluntary” standard for workplace wellness programs at risk of violating the ADA).

whose wellness programs did not comply with the agency’s 2000 guidance.”<sup>150</sup> Even for more risk-seeking employers, the guidance created concern and caution regarding the meaning and force of the guidance in the implementation of wellness programs, especially prior to the passage of the ACA.<sup>151</sup> For all employers who followed the guidance, the implications of the agency’s 2000 enforcement guidance were quite restrictive, particularly with respect to two aspects in the design and evaluation of a successful wellness programs.

First, the guidance’s definition of voluntary substantially limited compliant employers’ ability to incentivize workers to participate in wellness programs. The enforcement guidance’s bar on employers’ requiring participation in a wellness program was clearly in line with the ADA’s statutory text—intended to prohibit employers from forcing disabled employees to undergo health evaluations that were not directly related to the job, under the guise of a wellness program.<sup>152</sup> The enforcement guidance’s bar on penalties, however, went far beyond the ADA’s statutory text. According to the EEOC, employers were not allowed to penalize employees in any manner for failing to participate in a workplace wellness program. Certainly, at some point, a

---

<sup>150</sup> In 2014, the EEOC filed three actions against employers with wellness programs who did not comply with the agency’s former guidance. All three actions were ultimately unsuccessful for the EEOC, which may have suggested to the agency the need for revised regulations and guidance on wellness programs. *See, e.g.*, *EEOC v. Flambeau, Inc.*, 2015 U.S. Dist. LEXIS 173482 (W.D. Wis. Dec. 30, 2015) (alleging ADA violations); *EEOC v. Honeywell International Inc.*, 2015 U.S. Dist. \_\_\_\_ (D. Minn. Oct. 27, 2014) (alleging ADA and GINA violations); Lewis Krauskopf & Mica Rosenberg, *U.S. Judge Denies EEOC Bid to Stop Honeywell Wellness Penalty*, REUTERS, Nov. 3, 2014, <http://www.reuters.com/article/honeywell-intl-eeoc-idUSL1N0ST26K20141103> (reporting that a federal judge decided against the EEOC in the Honeywell case).

<sup>151</sup> Accord Ilyse Schuman, Russell Chapman, & Michelle Thomas, *EEOC Issues Long-Awaited Proposed Rule on Employer Wellness Programs* (May 2015), LITTLER MENDELSON, <https://www.littler.com/publication-press/publication/eeoc-issues-long-awaited-proposed-rule-employer-wellness-programs> (describing the “quagmire” that the EEOC’s 2000 enforcement guidance had created for employers with wellness programs). The quagmire created by the 2000 enforcement guidance became even more confusing in 2012 when the Eleventh Circuit held a financial penalty for nonparticipation legal under the ADA’s “safe harbor” provision, which allows for the administration of a “bona fide benefit plan that are based on underwriting risks, classifying risks, or administering such risks.” *See id.*; *Seff v. Broward County*, 691 F.3d 1221 (11th Cir. 2012); Americans with Disabilities Act of 1990, *Pub. L. 101-336*, 42 U.S.C. § 12201(c)(2).

<sup>152</sup> *See* Americans with Disabilities Act of 1990, *Pub. L. 101-336*, 42 U.S.C. § 12102(d).

penalty for nonparticipation might become so onerous as to become, in essence, a requirement to participate in a wellness program. But the 2000 enforcement guidance banned all penalties, no matter how small. At the same time, the enforcement guidance said nothing about using positive incentives to reward participation. If read literally, the EEOC's guidance seemed to allow incentives for participation so large that they might arguably force wellness program participation.<sup>153</sup> The agency's position of allowing giant carrots, but prohibiting even small sticks, appeared inconsistent.

More importantly, the agency's position may have dampened compliant employers' ability to develop a wellness program that actually improves employee health. Behavioral economists have demonstrated in a variety of contexts that averting loss provides greater motivation for behavior modification than does acquiring equivalent gains. The phenomenon of loss aversion, which is one of the principal results of prospect theory,<sup>154</sup> has been demonstrated to hold in the weight-loss and wellness context. For example, recall the 2013 Cawley and Price study finding better weight-loss outcomes for individuals faced with financial penalties than for individuals faced with equivalent financial rewards.<sup>155</sup> By prohibiting the use of even small penalties for nonparticipation in a wellness program, the agency's guidance arguably tied employers' hands in crafting an effective workplace wellness program.

---

<sup>153</sup> The U.S. Department of Labor's 2006 Health Insurance Portability and Accountability Act of 1996 (HIPPA) regulations formerly placed a limit on positive incentives at 20 percent of the total cost of plan coverage, but neither the EEOC's ADA regulations nor the 2000 ADA guidance placed any limit on positive incentives. For a discussion of the Labor Department's 2006 regulations and the changes made to these regulations as a result of the ACA, see 77 FR 70620, 70621 (Nov. 26, 2012).

<sup>154</sup> Prospect theory famously originated with Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision under Risk*, 47 *ECONOMETRICA* 263, 263-291 (Mar. 1979).

<sup>155</sup> See, e.g., John Cawley & Joshua A. Price, *A Case Study of a Workplace Wellness Program That Offers Financial Incentives for Weight Loss*, 32 *J. HEALTH ECON.* 794, 794-803 (2013).

Second, the EEOC’s enforcement guidance prohibited complaint employers from investigating how their workplace wellness program was influencing (or not influencing) workplace productivity. By requiring that “any medical records acquired as part of the wellness program [be] kept . . . separate from personnel records,”<sup>156</sup> the 2000 enforcement guidance excluded researchers, third-party wellness program administrators, and employers from asking the wellness-productivity question using data from the field. This ban on matching personnel records to wellness program records was intended to prevent employers from discovering a disabling (or potentially disabling) health condition through a wellness program and subsequently using this information to discriminate in the workplace. While well intentioned, the agency’s ban perpetuated the dearth of hard empirical evidence regarding whether wellness programs affect workplace productivity, and if so, how they affect workplace productivity. By prohibiting even researchers from compiling and analyzing the necessary data for program evaluation, anecdotes and assumptions about these programs’ efficacy have taken hold.<sup>157</sup>

Anecdotes and assumptions, if erroneous, are harmful to both employers and employees. Most obviously, if employers are introducing and supporting workplace wellness program under the

---

<sup>156</sup> Equal Employment Opportunity Commission, *Disability-Related Inquiries and Medical Examinations of Employees under the Americans with Disabilities Act (ADA)* (July 2000), available at <https://www.eeoc.gov/policy/docs/guidance-inquiries.html>.

<sup>157</sup> *Accord* RAND Corporation, *Do Workplace Wellness Programs Save Employers Money*, RAND Research Brief (2014), available at [http://www.rand.org/content/dam/rand/pubs/research\\_briefs/RB9700/RB9744/RAND\\_RB9744.pdf](http://www.rand.org/content/dam/rand/pubs/research_briefs/RB9700/RB9744/RAND_RB9744.pdf) (“The press and trade publications strongly endorse workplace wellness programs as a good investment for employers, and even the normally skeptical academic world has joined the bandwagon”); see also Jessica L. Roberts & Elizabeth Weeks Leonard, *What Is (and Isn’t) Healthism?*, 50 GA. L. REV. 1, 71 (2016) (arguing that the ACA merely “capitalized on [wellness programs’] increasing popularity, analogous to putting a car on a train that was already running down the rails.”)

mistaken belief that worker productivity will improve, then employers are wasting their money. Programs that were sold to employers to improve their bottom line may, instead, be hurting their bottom line. Yet the positive anecdotes and assumptions that persist around these programs in the absence of hard empirical evidence may be most harmful to employees, and particularly employees with visible health conditions. The popular—but empirically uncorroborated—story that wellness programs improve health, and healthier workers are more productive workers, relies on the assumption that better health improves productivity. The hidden implication of such an assumption is that workers with a health condition must, by nature, be less productive.

But that assumption may not always be true—and, in fact, may be wholly false for some health conditions and some occupations. In the absence of available data to investigate the wellness-productivity question in the workplace, employers can persist unchecked in subscribing to popular assumptions and, as a result, persist in regarding individuals with visible health conditions as less productive than their counterparts without visible health conditions. But such unsubstantiated regard of individuals with visible health conditions is precisely what the ADA was intended to protect against: prohibiting workplace discrimination against individuals who are “regarded as having . . . an impairment.”<sup>158</sup> In other words, by blocking researchers and employers from gaining a more precise, data-grounded understanding of the effects of health conditions, health metric improvement, and wellness on workplace productivity, the EEOC’s 2000 enforcement guidance has blocked disability advocates’ opportunity to craft a countervailing (or at least a more nuanced) narrative regarding the effects of health on workplace productivity. In turn, harmful and potentially false assumptions regarding workers with visible

---

<sup>158</sup> See Americans with Disabilities Act of 1990, *Pub. L. 101-336*, 42 U.S.C. § 12102(1)(c).

health conditions have festered among employers and the general public. An example of a common visible health condition, obesity, will illustrate this concern in the next part.

### **Unintended Consequences of the Old Regime: The Example of Obesity**

The enforcement guidance issued by the EEOC in 2000 has imposed a significant roadblock to compiling the necessary data for evaluating wellness programs rigorously, even as the collection of data has become simpler and cheaper with the influx of improved, lower-cost health technology over the last several years. As noted in the previous part, the result has been a glut of workplace wellness programs with very little evidence supporting their effectiveness and even less evidence supporting the mechanisms behind any effectiveness. Clearly, allowing health data collection and matching to personnel records would be beneficial to employers, as it would allow them to evaluate their wellness programs empirically, modify any aspects that needed improvement, and maximize the returns from their programs. Yet the benefits of allowing health data collection and matching to personnel records may be less obvious for employees. Through the example of obesity, this part will explain the benefits that may accrue to employees—and particularly employees with visible health conditions—from allowing employers to collect and match wellness program data.

Obesity has become a major concern for employers.<sup>159</sup> Well known is the so-called “obesity epidemic,” which refers to a tripling in U.S. obesity rates over the last two decades.<sup>160</sup> According to recent data, more than one-third of the U.S. adult population is obese, with the nationwide obesity rate currently at 34.9%.<sup>161</sup> Because obesity is correlated with a host of other health conditions—such as type 2 diabetes, coronary heart disease, high blood pressure, high cholesterol, and musculoskeletal problems<sup>162</sup>—employers worry about its effect on business costs. Employer cost concerns take two forms. First, employers who provide health insurance worry about obese employees driving up premiums, due to obesity’s association with higher health care costs.<sup>163</sup> Second, employers (regardless of whether they provide insurance) often hold negative assumptions about the link between obesity and productivity.<sup>164</sup>

---

<sup>159</sup> See, e.g., Bruce Y. Lee, *Obesity is Everyone’s Business*, FORBES, Sep. 1, 2015, <http://www.forbes.com/sites/brucelee/2015/09/01/obesity-is-everyones-business/#71cc2ab243fa> (arguing that obesity creates serious issues for employers); Meredith Melnick, *Study: Obese Workers Cost Employers \$73 Billion Per Year*, TIME, Oct. 11, 2010, <http://healthland.time.com/2010/10/11/study-obese-workers-cost-employers-73-billion-per-year/> (describing a study funded by the lap-band device company showing that obesity is costly in the workplace).

<sup>160</sup> See, e.g., Nat’l Ctr. for Chronic Disease Prevention & Health Promotion, *The Obesity Epidemic*, CTRS. FOR DISEASE CONTROL & PREVENTION (July 22, 2011), <http://www.cdc.gov/cdctv/diseaseandconditions/lifestyle/obesity-epidemic.html>.

<sup>161</sup> See Cynthia L. Ogden et al., *Prevalence of Childhood and Adult Obesity in the United States, 2011-2012*, 311 J. AM. MED. ASS’N 806, 806 (2014).

<sup>162</sup> See, e.g., NAT’L INST. OF HEALTH, *MANAGING OVERWEIGHT & OBESITY IN ADULTS: SYSTEMATIC EVIDENCE REVIEW FROM THE OBESITY EXPERT PANEL 1–4* (2013), <http://www.nhlbi.nih.gov/sites/www.nhlbi.nih.gov/files/obesity-evidence-review.pdf>.

<sup>163</sup> For an exploration of the link between obesity and the cost of employee benefits, see Jay Bhattacharya and M. Kate Bundorf, *The Incidence of the Healthcare Costs of Obesity*, 28 J. HEALTH ECON. 649, 649-658 (2009).

<sup>164</sup> See, e.g., Karen Insley, *What is Obesity Causing YOU as an Employer?*, LINKEDIN (Dec. 9, 2015), <https://www.linkedin.com/pulse/what-obesity-costing-you-employer-karen-insley> (“Direct costs [from obese employees] to you the employer can mean that staff productivity decreases”); National Business Group on Health, *Weight Management* (Aug. 14, 2015), [https://www.businessgrouphealth.org/resources/topics/weight\\_management\\_2.cfm](https://www.businessgrouphealth.org/resources/topics/weight_management_2.cfm) (“Reducing obesity and promoting healthy weight may . . . positively impact both productivity and personal well-being.”).



With respect to the first concern, employers—and particularly small employers—may be right to be concerned about obese employees driving up insurance premiums. Evidence does exist to suggest that obesity is associated with higher health care costs, not necessarily due to obesity itself, but due to its comorbidities. According to one study, obese individuals incur over \$700 more per year in medical expenditures than do normal-weight individuals.<sup>165</sup> For large employers, one obese worker will not likely drive up health insurance premiums because of risk pooling (although many obese workers could drive up large employers’ premiums).<sup>166</sup> For small employers, however, one obese worker can drive up health insurance premiums much more easily because of a smaller risk pool.<sup>167</sup> For all employers who provide health insurance, whether large or small, the Health Insurance Portability and Accountability Act of 1996 (HIPAA) prohibits increasing an employee’s contribution based on a visible health condition.<sup>168</sup> As a result, employers will end up bearing at least some of any obesity-related increases in premium costs.

---

<sup>165</sup> Eric A. Finkelstein, Ian C. Flebelkorn, & Guijing Wang, *National Medical Spending Attributable to Overweight and Obesity: How Much, and Who’s Paying?*, W3 HEALTH AFF. 219, 219–226 (2003).

<sup>166</sup> Accord Jay Bhattacharya and M. Kate Bundorf, *The Incidence of the Healthcare Costs of Obesity*, 28 J. HEALTH ECON. 649, 649-658 (2009) (“[A]s the firm size grows large, the marginal costs to any particular worker of higher expected medical costs tend toward zero. An implication of this is that, even if pooling exists at the level of the firm, we may observe wage offsets associated with obesity driven by limitations in pooling among small firms”).

<sup>167</sup> *See id.*

<sup>168</sup> *See* 29 U.S.C. § 1182(b) (2012) (“A group health plan . . . may not require any individual . . . to pay a premium or contribution which is greater than such premium or contribution for a similarly situated individual enrolled in the plan on the basis of any health status-related factor.”); *see also* U.S. Department of Labor, Employee Benefits Security Administration, *The HIPAA Nondiscrimination Requirements* (2016), [https://www.dol.gov/ebsa/faqs/faq\\_hipaa\\_ND.html](https://www.dol.gov/ebsa/faqs/faq_hipaa_ND.html) (“Under HIPAA, an individual cannot be denied eligibility for benefits or charged more for coverage because of any health factor.”).

Nevertheless, popular assumptions that obesity reduces a worker's productivity may be unsubstantiated. Because of the lack of data on the relationship between health and productivity—described in Part II and directly resulting from the 2000 EEOC Enforcement Guidance—very little is documented empirically regarding the relationship between obesity and productivity. Moreover, what evidence does exist casts doubt—or at least nuance—on popular assumptions. One 2008 study documents a correlation (but not a causal relationship) between obesity and workplace absenteeism.<sup>169</sup> But other studies using publically available labor market data call common beliefs about the negative effects of obesity on productivity into question. Although the studies using publically available data have repeatedly demonstrated that obese workers earn less and are less likely to be employed than non-obese workers,<sup>170</sup> lower earnings and employment effects are not necessarily the result of obese workers' lower productivity; they might also stem from weight-based discrimination in the workplace. Along these lines, a 2004 empirical study suggests that a nontrivial portion of the so-called obesity wage and employment penalty may stem from discrimination—not productivity—effects, particularly for obese women.<sup>171</sup>

---

<sup>169</sup> See John Cawley et al., *Occupation-Specific Absenteeism Costs Associated with Obesity and Morbid Obesity*, 49 J. OCC. & ENVIRON. MED. 1317, 1317-24 (Jan. 2008).

<sup>170</sup> See, e.g., Jose A. Pagan & Alberto Davila, *Obesity, Occupational Attainment, and Earnings*, 8 SOC. SCI. Q. 756, 756-70 (1997); Susan Averett & Sanders Korenman, *The Economic Reality of the Beauty Myth*, 31 J. HUM. RESOURCES 304, 304-30 (1996); Steven L. Gortmaker et al., *Social and Economic Consequences of Overweight in Adolescence and Young Adulthood*, 329 NEW ENG. J. MED. 1008, 1008-12 (Sept. 30, 1993).

<sup>171</sup> See John Cawley, *The Impact of Obesity on Wages*, 39 J. HUM. RESOURCES 451, 451-74 (2004) (using the National Longitudinal Survey of Youth to show that even after controlling for education, skill level, and family characteristics, a wage penalty persisted for obese women).

A recent study from 2016 goes more directly to the obesity-productivity question, using publicly available labor market data, by comparing the occupational characteristics of obese and non-obese workers. The study theorizes that if the obesity penalty were largely driven by productivity effects, then the obesity penalty should be greatest for workers in physically demanding occupations since obesity is correlated with the development of musculoskeletal conditions. In other words, if obesity's negative effect on productivity were driving the obesity penalty, then physically demanding occupations would employ fewer obese workers, and the obese workers in such occupations would earn less than non-obese workers. On the other hand, if the obesity penalty were mostly driven by discrimination effects, then the obesity penalty should be greatest for workers in occupations that require interaction with the public, in which appearance is likely to be most salient for employers and their customers.<sup>172</sup> Finding that obese workers are actually more likely than non-obese workers to be employed in physically demanding jobs—and obese workers in physically demanding jobs are paid the same as non-obese workers—the study's empirical results contradict popular assumptions about obesity's negative effects on productivity. Moreover, the study highlights that a substantial portion of the obesity penalty can be explained by how poorly obese individuals (and particularly obese women) fare in public interaction jobs, indicating the prevalence of weight-based discrimination in the workplace.<sup>173</sup>

---

<sup>172</sup> See Jennifer Bennett Shinall, *Distaste or Disability? Evaluating the Legal Framework for Protecting Obese Workers*, 37 BERKELEY J. EMP. & LAB. L. 101, 101-42 (2016) (using data from the Current Population Survey, Eating and Health Module, and the Occupational Information Network to compare the occupational characteristics of obese and non-obese workers).

<sup>173</sup> See *id.*

The 2016 study calls into question many of the common assumptions about the relationship between obesity and workplace productivity. Moreover, both the 2004 and 2016 studies suggest the prevalence of systematically negative employer attitudes towards obese workers. Such negative attitudes have also been documented in numerous psychological studies, which consistently find automatic stereotyping of obese individuals as lacking self-discipline, lazy, less conscientious, less competent, sloppy, and more likely to have a personal problem.<sup>174</sup> Many of these documented negative attitudes towards obese workers are directly related to perceptions of workplace productivity.

With the pervasiveness of negative, yet unsubstantiated, perceptions of obese workers' lack of productivity, the EEOC's 2000 enforcement guidance has done nothing to improve or correct these inaccurate perceptions among employers or the general public. By barring employers, and even researchers, from collecting and analyzing wellness program data, the guidance has

---

<sup>174</sup> See, e.g., Tanya Berry & John C. Spence, *Automatic Activation of Exercise and Sedentary Stereotypes*, 80 RES. Q. EXERCISE & SPORT 633, 633-40 (Sep. 2009) (finding subjects associated words such as "unmotivated, lethargic, unfit, lazy, inactive, sluggish, idle, weak, sickly, [and] loaf" with pictures of overweight individuals); Mark V. Roehling, *Weight-Based Discrimination in Employment: Psychological and Legal Aspects*, 52 PERSONNEL PSYCHOL. 969, 969-1016 (Dec. 1999) (concluding from a review of prior psychology studies that obese individuals are stereotyped as lacking personal traits required for productivity in the workplace); R. Pingitore et al., *Bias Against Overweight Job Applicants in a Simulated Employment Interview*, 79 J. APPLIED PSYCHOL. 909, 909-17 (1994) (finding subjects were less likely to hire individuals in a padded suit to make them look heavier); C. R. Jasper & M. L. Klassen, *Perceptions of Salespersons' Appearance and Evaluation of Job Performance*, 71 PERCEPTUAL & MOTOR SKILLS 563, 563-66 (1990) (demonstrating that subjects rated obese applicants for a sales position as less desirable candidates than non-obese applicants); Esther D. Rothblum et al., *Stereotypes of Obese Female Job Applicants*, 7 INT'L J. EATING DISORDERS 277, 277-83 (1988) (finding subjects were more likely to characterize obese job applicants as lacking self-discipline, supervisory potential, professional appearance, and personal hygiene); W. H. Decker, *Attributions Based on Managers' Self-Presentation, Sex, and Weight*, 71 PSYCHOL. REP. 175, 175-81 (1987) (finding subjects rated normal-weight managers as more likely to be good supervisors than overweight managers); J. C. Larkin & H. A. Pines, *No Fat Persons Need Apply: Experimental Studies of the Overweight Stereotype and Hiring Preference*, 6 SOC. WORK OCCS. 312, 312-27 (1979) (finding in simulated employment interviews that subjects were less likely to describe overweight applicants as neat, productive, ambitious, disciplined, or determined).

prevented a rigorous investigation into the validity of stereotypes surrounding obese individuals in the workplace. As such, it is hardly surprising that these stereotypes persist—not only around obese workers’ ability to do their jobs, but also around their ability to succeed in wellness programs. A 2008 survey of employers conducted by researchers at George Washington University and the University of Chicago found that 93 percent agreed with the statement that obesity was “the result of poor lifestyle choices”; 83 percent believed that obesity was “the result of poor willpower.”<sup>175</sup> These results are in line with other U.S. public opinion surveys, which consistently assign culpability to obese individuals for their weight, blaming a lack of self-control.<sup>176</sup>

Researchers, however, take virtually the opposite view of obese individuals’ ability to control their weight. Contrary to popular belief, once an individual has become obese, losing weight is not necessarily a simple formula of reducing calories in while increasing calories out, or of eating less and exercising more.<sup>177</sup> Rather, neurological research has revealed how, upon weight loss, the body unceasingly fights to regain the weight, regardless of whether an individual’s

---

<sup>175</sup> STOP Obesity Alliance, *Employer/Employee Survey on Obesity in the Workplace* (May 2008), <http://stopobesityalliance.org/research-and-policy/research-center/survey-results/>. See also John R. Gabel, et al., *Obesity and The Workplace: Current Programs and Attitudes among Employers and Employees*, 28 HEALTH AFF. 46, 46-56 (2009); STOP Obesity Alliance, *Fast Facts: Employer and Employee Attitudes Toward Obesity* (2016), <http://stopobesityalliance.org/wp-content/themes/stopobesityalliance/pdfs/Fast%20Facts%20Employer%20Employee%20Attitudes.pdf>.

<sup>176</sup> See, e.g., Jayson L. Lusk & Brenna Ellison, *Who Is To Blame for the Rise in Obesity?*, 68 APPETITE 14, 17-18 (2013) (finding in a U.S. public opinion survey that 80 percent of respondents blamed individuals as primarily responsible for the nationwide rise in obesity rates); J. Eric Oliver & Taeku Lee, *Public Opinion and the Politics of Obesity in America*, 30 J. HEALTH POL. POL’Y & L. 923, 933 (2005) (finding in a U.S. public opinion survey that 65% of Americans believed that obese people lacked personal willpower, and 62% of respondents thought that obesity was solely the result of an individual’s choice to consume unhealthy food).

<sup>177</sup> For a description of the research on the difficulty of losing weight once an individual becomes obese, see Jennifer Bennett Shinall, *Unfulfilled Promises: Discrimination and the Denial of Essential Health Benefits under the Affordable Care Act*, DEPAUL L. REV. (forthcoming, 2016).

initial weight was above or below normal.<sup>178</sup> As a result, approximately 90 to 95 percent of individuals who successfully lose weight on a diet will regain the weight within several years.<sup>179</sup> Indeed, a 2014 article in the Journal of the American Medical Association concludes after evaluating the collective medical research that “[a]ttempts to lower body weight without addressing the biological drivers of weight gain, including the quality of the diet, will inevitably fail for most individuals.”<sup>180</sup> Even the Handbook of Obesity, a research guide written by leading scientists and practitioners, intended to provide “up-to-date coverage of the range of subjects that make up the field of obesity research[,]”<sup>181</sup> famously concludes that the long-run results of traditional diets that encourage restricting calories and increasing exercise are “poor and not long-lasting.”<sup>182</sup>

All this to say, to the extent employers’ wellness program values and rewards weight loss (which many do), legal policy should not discourage, but encourage data collection on actual outcomes. Medical research suggests that employee weight loss from wellness programs is likely to be modest at best and not necessarily sustainable in the long run. Yet the employer survey evidence discussed above reveals that employers’ expectations are often unrealistic and based in anecdote;

---

<sup>178</sup> For an accessible summary of research identifying neurological pathways that contribute to energy homeostasis—meaning in this case, the return to original, starting weight., see Roger D. Cone, *The Central Melanocortin System and Energy Homeostasis*, 10 TRENDS ENDOCRINOLOGY & METABOLISM 211, 211-216 (Aug. 1999).

<sup>179</sup> Susan C. Wooley & David M. Garner, *Controversies in Management: Dietary Treatments for Obesity Are Ineffective*, 309 BMJ 655, 655 (1994).

<sup>180</sup> David S. Ludwig & Mark I. Friedman, *Increasing Adiposity: Consequence or Cause of Overeating*, 311 J. AM. MED. ASS’N 2167, 2167 (2014)

<sup>181</sup> See GEORGE A. BRAY, ED., HANDBOOK OF OBESITY (1998).

<sup>182</sup> See Luc F. Van Gaal, *Dietary Treatment of Obesity*, 875-76, in GEORGE A. BRAY, ED., HANDBOOK OF OBESITY (1998). (“Losing weight is relatively easy, but the maintenance of weight loss may be more distressing . . .”).

moreover, in the event that employees fail to meet designated weight goals, employers may resort to blame and characterize it as a personal failing on the employee's part. If instead, employers were allowed to view the aggregate data and take a bird's eye view of employee weight-loss outcomes, employers might gain a better understanding of the difficulties of the weight loss process. Comparing an obese worker's weight loss outcomes, for example, to program-wide weight loss outcomes may cause the employer to reconsider conclusions that an obese worker is lazy or lacking in willpower. Furthermore, to the extent that the wellness program contains any weight-loss goals, viewing the data in context might spur employers to revise program goals—for instance, by substituting relative health metric targets for absolute targets—such that they are more realistic for individual employees.

Admittedly, some employer opinions may not be directly swayed by hard empirical evidence on their wellness program outcomes. Nonetheless, by preventing researchers from working with employers to access, collect, and analyze workplace wellness program and productivity data, former EEOC policy has stood in the way of disability advocates' ability to sway anyone's opinion, including the opinion of the general public. For this reason, legal policy should not discourage, but encourage data collection on the relationship between employees' weight-loss outcomes and their productivity at work—and more generally, on the relationship between all wellness program health outcomes and productivity at work. Although intended to protect workers with health conditions like obesity, the former EEOC policy has arguably backfired. Certainly obese workers have not been helped by the former EEOC policy that prevents employers from analyzing their health metrics and productivity metrics together. In the classic case of unintended consequences, obese workers—and arguably, all workers with visible health

conditions—have been harmed by the very legal regime designed to protect them. By impeding both researchers’ and employers’ ability to understand the true relationship between health and productivity, the 2000 EEOC enforcement guidance has helped to perpetuate health-status based discrimination in the workplace, rather than to combat it. Although recently promulgated EEOC regulations, described in the next part, can go a long way in remedying the unintended consequences of the old regime, the new regulations are already endangered.

### **The Promise of a New Regulatory Regime**

More than a decade after releasing its enforcement guidance on wellness programs under the ADA, the EEOC at last issued a much needed update on employer wellness programs. Rethinking and clarifying the previous regime, the agency proposed new rules on wellness programs—which were issued under the authority of the ACA—on April 20, 2015.<sup>183</sup> The ACA contained multiple provisions encouraging the development and expansion of workplace wellness programs, under the premise that these programs can meaningfully improve participants’ health metrics and reduce participants’ need for health care<sup>184</sup> (although how accurate that premise is remains questionable). As such, the EEOC asserted in the interpretive guidance accompanying the proposed rules that “it ha[d] a responsibility to interpret the ADA in a manner that reflects both the ADA’s goal of limiting employer access to medical information

---

<sup>183</sup> See generally 80 FR 21659-01 (Apr. 20, 2015).

<sup>184</sup> See, e.g., Patient Protection and Affordable Care Act, PL 111-148, §§ 4301-4306 (March 23, 2010) (providing support for workplace wellness programs); 42 U.S.C. § 280I (2012) (providing for “technical assistance” for the “utilization of evidence-based prevention and health promotion approaches in the workplace”);

42 U.S.C. § 300gg-4(m) (2012) (requiring a report on wellness programs by the Department of Labor).



and HIPAA's and the Affordable Care Act's provisions promoting wellness programs.”<sup>185</sup> The agency then opened the floor to a two-month notice-and-comment period, and approximately one year later, issued substantially similar final regulations on May 17, 2016. The new regulations went into effect two months later, on July 18, 2016.<sup>186</sup>

These regulations, this article will argue, have the potential to advance researchers', employers', and the public's understanding of the relationship between wellness programs and health, the relationship between health and productivity, and ultimately, the efficacy of workplace wellness programs. Specifically, the new regulations lay the groundwork for the building blocks necessary for remedying much of the harm from the prior regulations in two important ways. First, the new regulations vastly expand the definition of a voluntary wellness program. While the new regulations continue to embrace the plain-meaning definition of voluntary—prohibiting employers either from requiring wellness program participation<sup>187</sup> or from taking an adverse employment action against an employee for non-participation<sup>188</sup>—the new regulations remedy the ambiguity and inconsistency within the prior rules and 2000 enforcement guidance. No longer are employers allowed to use giant carrots, but prohibited from using tiny sticks, to encourage good-faith participation in a wellness program. Rather, both penalties and rewards are equally allowable now, within defined bounds:

---

<sup>185</sup> 80 FR 21659-01, 21662 (Apr. 20, 2015).

<sup>186</sup> *See* 81 FR 31126 (May 17, 2016).

<sup>187</sup> *See* 29 C.F.R. § 1630.14(2)(i) (requiring that wellness programs that collect health metrics do “not require employees to participate”).

<sup>188</sup> *See* 29 C.F.R. § 1630.14(2)(iii) (requiring that wellness programs that collect health metrics do “not take any adverse employment action or retaliate against, interfere with, coerce, intimidate, or threaten employees”).

The use of incentives (financial or in-kind) in an employee wellness program, whether in the form of a reward or penalty, will not render the program involuntary if the maximum allowable incentive available under the program (whether the program is a participatory program or a health-contingent program, or some combination of the two . . .) does not exceed . . . [t]hirty percent of the total cost of self-only coverage (including both the employee's and employer's contribution) of the group health plan.<sup>189</sup>

Considering the inadequacies of the prior regulations and enforcement guidance, this clearer and more consistent definition of voluntary gives employers more flexibility and additional tools to develop a successful wellness program. The new regulations now allow employers to use significant financial incentives, either positive or negative, of up to 30 percent of the cost of individual health plan coverage to encourage healthy behavioral modifications.<sup>190</sup> To the extent that the disappointing results in the few prior studies of wellness programs have been driven by limitations in incentive magnitude, the new regulations allow employers to test whether larger financial incentives lead to larger improvement in health metrics. Moreover, by allowing employers to use financial penalties, the EEOC has eliminated its arguably incongruous preference for equivalent financial rewards. More importantly, the agency has allowed wellness program administrators to take advantage of the behavioral phenomenon of loss aversion—that

---

<sup>189</sup> 29 C.F.R. § 1630.14(3). This section goes onto clarify how to calculate 30 percent of self-only coverage if the employer offers a group health plan but the employee is not enrolled, if the employee offers multiple group health plans, and if the employer does not offer a group health plan.

<sup>190</sup> The regulations make an exception for tobacco cessation programs that are part of a wellness program. In tobacco programs, employers may use financial incentives, either positive or negative, of up to 50 percent of the cost of individual health plan coverage. *See* 81 FR 31126, 31136 (May 17, 2016).

is, utilizing the fact that averting financial loss provides greater motivation for behavioral modification than does acquiring an equivalent financial reward<sup>191</sup>—to design programs that more effectively improve health metrics. Given the generally disappointing evidence regarding wellness programs’ current impact on health metrics (discussed in Part I), and the fact that financial penalties have already been shown to be more effective than rewards in one weight-loss context,<sup>192</sup> this regulatory advancement represents an important step towards dispensing with wellness programs that sound good in theory, but offer little to no benefit in reality.

The second, more critical way in which the new regulations can potentially remedy harm from prior EEOC regulations and guidance is through the omission of language about keeping “any medical records acquired as part of the wellness program . . . confidential and separate from personnel records.”<sup>193</sup> The removal of this language from agency guidance at last opens the door to both employers and researchers studying the relationship (if any) between workplace wellness programs, employee health metrics, and employee productivity. The new regulations are still adamant about keeping employee medical information confidential,<sup>194</sup> as required by the statutory language of the ADA,<sup>195</sup> and prohibit the collection of employee health metrics without

---

<sup>191</sup> See generally Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision under Risk*, 47 *ECONOMETRICA* 263, 263-291 (Mar. 1979).

<sup>192</sup> See John Cawley & Joshua A. Price, *A Case Study of a Workplace Wellness Program that Offers Financial Incentives for Weight Loss*, 32 *J. HEALTH ECON.* 794, 794-803 (2013).

<sup>193</sup> Equal Employment Opportunity Commission, *Disability-Related Inquiries and Medical Examinations of Employees under the Americans with Disabilities Act (ADA)* (July 2000), available at <https://www.eeoc.gov/policy/docs/guidance-inquiries.html>.

<sup>194</sup> See, e.g., 29 C.F.R. § 1630.14(d)(4)(iv) (“A covered entity shall not require an employee . . . to waive any confidentiality protections in this part as a condition for participating in a wellness program or for earning any incentive.”)

<sup>195</sup> See 42 U.S.C. § 12112(d) (“[I]nformation obtained regarding the medical condition or history of the applicant is collected and maintained on separate forms and in separate medical files and is treated as a confidential medical

intention of improving employee health, wellness programs, or both.<sup>196</sup> But the new regulations specifically permit the “measurement, test, screening, or collection of health-related information without providing results, follow-up information, or advice . . . [if it] is used to design a program that addresses at least a subset of the [health] conditions identified.”<sup>197</sup>

In fact, the EEOC’s interpretive guidance encourages the development of wellness programs that actually improve both employee health and productivity, acknowledging that efficacy was a principal goal of the ACA.<sup>198</sup> In the final regulations, the agency added the requirement that “[a]n employee health program, including any disability-related inquiries or medical examinations that are part of such program, must be reasonably designed to promote health or prevent disease.”<sup>199</sup> This additional requirement not only serves as a further protection for employee health metrics—since it authorizes employers to collect data only if it furthers the goal of improving employee health—but also incentivizes employers to ensure that their program is reasonably likely to be effective. To that end, the interpretive guidance defines reasonable design as “hav[ing] a reasonable chance of improving the health of, or preventing disease in, participating employees, and . . . not overly burdensome, a subterfuge for violating the ADA or other laws prohibiting employment discrimination, or highly suspect in the method chosen to

---

record.”); 81 FR 31126, 31142 (May 17, 2016) (“Employers and wellness program providers must take steps to protect the confidentiality of employee medical information provided as part of an employee health program.”).

<sup>196</sup> See 29 C.F.R. § 1630.14(d)(1) (“A program consisting of a measurement, test, screening, or collection of health-related information without providing results, follow-up information, or advice designed to improve the health of participating employees is not reasonably designed to promote health or prevent disease.”)

<sup>197</sup> 29 C.F.R. § 1630.14(d)(1).

<sup>198</sup> See 81 FR 31126, 31131 (May 17, 2016) (acknowledging the importance of “effectiveness of wellness programs that the Affordable Care Act clearly intends to promote” in developing the agency’s final regulations).

<sup>199</sup> 29 C.F.R. § 1630.14(d)(1).

promote health or prevent disease.”<sup>200</sup> The safest, most accurate way for an employer to know that the company wellness program has a reasonable chance of improving health is to analyze data collected as part of the program, respond to the empirical results, and modify the program accordingly. Thus, the agency’s addition of the reasonable chance of improving health requirement and removal of the ban on matching health to personnel records work in tandem towards the development of wellness programs that are capable of—in the long run, through trial and error—improving employee health metrics, raising workplace productivity, and advancing employer understanding of the effects of health conditions on workplace productivity.

### **Countering Resistance: Concerns against the New Regime**

This article has argued, with some force, that the new ADA regulations and interpretive guidance on wellness programs are a positive development for all affected parties—including employees with visible health conditions. But not everyone has supported the EEOC’s recent position change on wellness programs. Indeed, even though this article has argued that the new regulations have opened the door to the collection and analysis of data that can be helpful in the long run for disability law advocates, practitioners, and scholars, these groups largely voiced strong opposition to the new regime during the 2015 notice-and-comment period. Concerns regarding the new regulations have largely centered on three issues: privacy, equity, and improper venue. This part will address each concern in turn, ultimately concluding that most of these concerns have either been addressed by the additional language in the final regulations or can be addressed by some additional suggested agency guidance.

---

<sup>200</sup> 81 FR 31126, 31140 (May 17, 2016).

## A. Privacy

Perhaps the most forceful backlash against the EEOC's new position towards employer wellness programs has focused on the potential of the new regulations to invade employee privacy, and in particular, disabled employee privacy. Over 300 separate entities—ranging from insurance companies and large employers to advocacy groups and individuals affected by disability—submitted letters to the EEOC during the notice-and-comment period.<sup>201</sup> The disability rights advocates and scholars who submitted comments in 2015 largely took a dismal view of the proposed rule, out of concern that it would undermine the confidentiality protections guaranteed by the ADA to disabled employees.<sup>202</sup> For instance, one advocacy group that counts several well-known legal scholars among its ranks, argued in its comment that the proposed rule was “inconsistent” with the ADA because it “allow[ed] employers to use steep financial penalties in wellness programs to force workers to disclose sensitive medical information to their employers.”<sup>203</sup> The new proposed rule, thus, ignored the fact that “[h]istorically, many employers

---

<sup>201</sup> See Comments (2015), Amendments to Regulations under the Americans with Disabilities Act, Regulations.gov, <https://www.regulations.gov/docketBrowser?rpp=50&so=DESC&sb=commentDueDate&po=0&dct=PS&D=EEOC-2015-0006>.

<sup>202</sup> See, e.g., Consortium for Citizens with Disabilities, *Comments on Proposed Rule, Amendments to Regulations under the Americans With Disabilities Act, RIN 3046-AB01* (June 19, 2015), available at <https://www.regulations.gov/docketBrowser?rpp=50&so=DESC&sb=commentDueDate&po=0&dct=PS&D=EEOC-2015-0006> (“We are surprised to see that, as the ADA approaches its 25th anniversary, the EEOC is proposing a rule that would significantly diminish workers’ rights to keep disability-related information unrelated to their job performance out of the hands of their employers and protect themselves from discrimination.”); American Diabetes Association, *Comments on Proposed Rule, Amendments to Regulations under the Americans With Disabilities Act, RIN 3046-AB01* (June 19, 2015), available at <https://www.regulations.gov/docketBrowser?rpp=50&so=DESC&sb=commentDueDate&po=0&dct=PS&D=EEOC-2015-0006> (“In order to preserve the intent of laws such as the Americans with Disabilities Act, the Commission’s final rule must prohibit the use of outcomes-based programs that base financial rewards or penalties on biometric screening outcomes standards that are coextensive with or directly related to a disability”).

<sup>203</sup> Bazelon Center for Mental Health Law, *Comments on Proposed Rule, Amendments to Regulations under the Americans With Disabilities Act, RIN 3046-AB01* (June 19, 2015), available at <https://www.regulations.gov/docketBrowser?rpp=50&so=DESC&sb=commentDueDate&po=0&dct=PS&D=EEOC>

asked applicants and employees to provide [similar medical] information . . . [and] used it to exclude and otherwise discriminate against individuals with disabilities.”<sup>204</sup> The comment concluded that the proposed rule was inconsistent with the statutory language of the ADA because wellness-program-related medical inquiries were “not job-related and consistent with business necessity.”<sup>205</sup>

These concerns, while well intentioned, fail to acknowledge both the relatively limited group of individuals to whom the concern applies and the large numbers of individuals with health conditions that the new regulations stand to help. The concern that employer collection of wellness program health data will induce additional disability bias in the workplace is only valid for those individuals who (1) have a nonvisible health condition that (2) has remained unknown to the employer outside the context of the wellness program. The concern is not valid for individuals whose health condition has already been plainly visible to the employer (such as the millions of obese workers described in Part III) or whose nonvisible health condition has been previously identified to the employer because of the need for a reasonable accommodation in the workplace. For individuals whose health conditions would have already been identified by the employer in the absence of a wellness program, the new regulations should, if anything, be more beneficial than detrimental in the workplace. As argued in Parts III and IV, even in the post-ADA regime, employers have continued to make harmful assumptions about the effect of health conditions on workplace productivity. Researchers and advocates have been limited in their

---

-2015-0006. Trustees of the Bazelon Center include Samuel Bagenstos, Martha Minnow, Elyn Saks, and Stephen Morse.

<sup>204</sup> *Id.*

<sup>205</sup> *Id.*

ability to produce counterevidence to such assumptions under the data ban imposed by the old regulatory regime. The new regulations make possible the production of counterevidence—or at least, more nuanced evidence. A more precise, data-driven understanding of the complex relationship between health and workplace productivity cannot harm, and may help, the fortunes of workers with visible health conditions.

The concerns that the new ADA wellness program regulations are inconsistent with the statutory text are similarly unfounded. The argument that wellness-program-related medical inquiries are “not job-related and consistent with business necessity” ignores the very reason that so many employers have put these programs in place: to improve their bottom line.<sup>206</sup> Whether to reduce health insurance premium costs, improve employee productivity, or both, employers’ objective with respect to these programs is not to delve into the most private aspects of workers’ personal lives, but rather to decrease expenditures and increase output. Under the old EEOC wellness regime, compliant employers were unable to evaluate their wellness programs, leading to a proliferation of wellness programs with questionable efficacy. To the extent that data collection and analysis assists employers in designing better wellness programs that actually decrease expenditures and increase output, the EEOC’s new wellness regime is job-related and consistent with business necessity.

---

<sup>206</sup> See, e.g., Kaiser Family Foundation & Health Research & Educational Trust, Employer Health Benefits: 2013 Annual Survey (September 2013), available at <https://kaiserfamilyfoundation.files.wordpress.com/2013/08/8465-employer-health-benefits-20131.pdf> (finding that 67 percent of employers identified wellness programs as a “very effective” or “somewhat effective” way of decreasing insurance costs and that wellness programs were the most commonly identified method of effective cost reduction among employers).



Finally, in response to the concerns regarding privacy raised by disability advocates during the notice-and-comment period, the EEOC augmented the final regulations and interpretive guidance to ensure they will indeed be a net positive for workers with health conditions. The agency's requirement that wellness program inquiries into health metrics are "reasonably designed to promote health or prevent disease"<sup>207</sup> bars employers from going on a fishing expedition for employee medical data that is not colorably job-related or consistent with business necessity. Moreover, the agency went to great lengths in the regulations and interpretive guidance to discuss how wellness program health data should be handled, particularly to the extent that they will be matched to personnel records. The agency strongly encourages employers to use de-identified, aggregate data, whenever possible, by requiring in the regulation,

Except . . . as is necessary to administer the health plan, information obtained [from a wellness program] . . . regarding the medical information or history of any individual may only be provided to an ADA covered entity in aggregate terms that do not disclose, or are not reasonably likely to disclose, the identity of any employee.<sup>208</sup>

In the interpretive guidance, the agency clarifies that using health data to improve the wellness program's efficacy is a legitimate administrative use.<sup>209</sup>

More critically, to the extent that non-aggregate data is required for program analysis, the agency guidance prohibits anyone with the authority to make personnel decisions or the authority to take an adverse employment action from viewing identified data.

---

<sup>207</sup> 29 C.F.R. § 1630.14(d)(1).

<sup>208</sup> 29 C.F.R. § 1630.14(d)(4)(iii).

<sup>209</sup> See, e.g., 81 FR 31126, 31139 (May 17, 2016) (allowing employers to use wellness program data collection to "to design a program that addresses at least a subset of the [health] conditions identified.")

Individuals who handle medical information that is part of an employee health program should not be responsible for making decisions related to employment, such as hiring, termination, or discipline. . . . Employers that administer their own wellness programs need adequate firewalls in place to prevent unintended disclosure.<sup>210</sup>

Because building such firewalls can often be difficult, the agency goes on to encourage the use of third-party vendors or researchers in handling such sensitive data.

Use of third-party vendors or researchers that maintain strict confidentiality and data security procedures should reduce the risk that medical information will be disclosed to individuals who make personnel decisions, particularly for employers whose organizational structure makes it difficult to provide adequate safeguards. If an employer uses a third party to analyze the data, it should be familiar with the third party's privacy policies for ensuring the confidentiality of medical information, particularly if that third party is a vendor.<sup>211</sup> The EEOC's requirement that employers use either third parties or firewalls when handling identified employee wellness program data protects against disability advocates' fears that employers will use these programs to trawl for personal health information and, in turn, use it against employees in the workplace. It also opens the door to employers partnering with academic researchers to analyze their wellness program data in a way that has not been possible before.

---

<sup>210</sup> 81 FR 31126, 31142 (May 17, 2016).

<sup>211</sup> *Id.*

## B. Equity

Also identified in many comments on the EEOC's proposed rule was the concern that health problems disproportionately affect historically disadvantaged groups, and in particular, racial and ethnic minority groups.<sup>212</sup> To the extent the new regulations allow for financial penalties in wellness programs—and in particular, financial penalties in health-contingent wellness programs—the regulations might allow programs to have a disparate impact on protected classes. Unfortunately, it is true that many health conditions are correlated with minority status.<sup>213</sup> For example, obesity rates among African-Americans and Hispanics are disproportionately higher than obesity rates among whites.<sup>214</sup> If wellness programs are allowed to penalize workers financially for failure to achieve certain health targets, then minority groups face a risk of being disproportionately fined. This issue becomes particularly problematic since members of minority

---

<sup>212</sup> See, e.g., Sondra Solovay et al., *Response to Request for Comments on Proposed Amendments to Regulations and Interpretive Guidance Implementing Title I of the Americans with Disabilities Act (ADA), Regarding Employer Wellness Programs* (June 19, 2015), available at <https://www.regulations.gov/docketBrowser?rpp=50&so=DESC&sb=commentDueDate&po=0&dct=PS&D=EEOC-2015-0006> (“We further recognize that weight stigma and discrimination disproportionately impact many minority groups including protected racial and ethnic minorities as well as lesbians.”).

<sup>213</sup> Accord James P. Smith & Raynard Kington, *Demographic and Economic Correlates of Health in Old Age*, 34 *DEMOGRAPHY* 159, 159 (1997) (discussing the “quantitatively large socioeconomic status-health gradient”); Risa Lavizzo-Mourey and David Williams, *Being Black Is Bad for Your Health*, *U.S. NEWS & WORLD REP.*, Apr. 14, 2016, <http://www.usnews.com/opinion/blogs/policy-dose/articles/2016-04-14/theres-a-huge-health-equity-gap-between-whites-and-minorities>.

<sup>214</sup> See, e.g., Jennifer Bennett Shinall, *What Happens When the Definition of Disability Changes? The Case of Obesity*, 5(2) *IZA J. LAB. ECON.* 1, 1-30 (2016) (demonstrating that African-American and Hispanic men and women have higher obesity rates than other men and women); see also Centers for Disease Control and Prevention, *Adult Obesity Facts* (Sep. 21, 2015), <https://www.cdc.gov/obesity/data/adult.html> (“Non-Hispanic blacks have the highest age-adjusted rates of obesity (47.8%) followed by Hispanics (42.5%), non-Hispanic whites (32.6%), and non-Hispanic Asians (10.8%)”).

groups are more likely to be living at or near the poverty level,<sup>215</sup> thus rendering any financial penalty especially harmful.

Without question, allowing health-contingent wellness programs to inflict a financial penalty of up to 30 percent of the cost of individual health care coverage is potentially devastating to an impoverished worker. Although the 30 percent allowance is specifically permitted within the text of the ACA,<sup>216</sup> additional protections are arguably needed to prevent disparate impact discrimination against minorities.<sup>217</sup> The EEOC's final regulation and interpretive guidance do provide three additional protections against wellness programs that target (intentionally or unintentionally) the poor and minorities, although some of these protections need to be clarified further in future agency guidance.

First, even though a financial penalty of up to 30 percent of the cost of individual health care coverage is permissible in a wellness program, the final EEOC regulations limit the maximum permissible penalty by requiring that the penalty be calculated based on low-cost health plans. For workers with employer-provided health insurance, the 30 percent is calculated based on the employer's lowest-cost plan, not the plan in which the worker is actually enrolled.<sup>218</sup> If the

---

<sup>215</sup> See, e.g., Economic Policy Institute, *The State of Working America, Poverty* (2016), <http://stateofworkingamerica.org/fact-sheets/poverty/> (“Among racial and ethnic groups, African Americans had the highest poverty rate, 27.4 percent, followed by Hispanics at 26.6 percent and whites at 9.9 percent.”).

<sup>216</sup> See 42 U.S.C. § 300gg-4(a)(3) (“The reward for the wellness program, together with the reward for other wellness programs with respect to the plan that requires satisfaction of a standard related to a health status factor, shall not exceed 30 percent of the cost of employee-only coverage under the plan”).

<sup>217</sup> Disparate impact discrimination on the basis of race and ethnicity is prohibited by Title VII of the 1964 Civil Rights Act. See 42 U.S.C. § 2000e-2(k) (2012).

<sup>218</sup> See 29 C.F.R. § 1630.14(d)(3)(iii) (permitting a maximum penalty of “[t]hirty percent of the total cost of the lowest cost self-only coverage under a major medical group health plan where the covered entity offers more than

employer does not offer a health plan, the 30 percent is calculated based on the “second lowest cost Silver Plan for a 40–year-old non-smoker on the state or federal health care Exchange in the location that the covered entity identifies as its principal place of business.”<sup>219</sup> Moreover, the agency pointed out in its interpretive guidance that Treasury Department rules not only incorporate an individual’s income but also include an assumption that individuals have “fail[ed] to satisfy the requirements of a wellness program” in regulating health plans’ affordability (and health plans’ ultimate cost).<sup>220</sup> Consequently, since the baseline cost of individual health care already incorporates an affordability standard, the EEOC asserted that an additional affordability standard would be redundant. Of course, if the healthcare marketplace is abolished and a benchmark Silver Plan ceases to exist, an additional (and perhaps more straightforward) affordability standard will be required from the EEOC.

Second, as an additional protection against low-income, historically disadvantaged individuals bearing the majority of wellness program penalties, the final regulations specifically prohibit wellness programs that “exist[] mainly to shift costs from the covered entity to targeted employees based on their health.”<sup>221</sup> This aspect of the regulations puts employers on notice to be vigilant regarding who is bearing the brunt of a wellness program’s financial penalties, and how much of that burden has shifted as a result of wellness program penalties. Still, agency

---

one group health plan but participation in the wellness program is offered to employees whether or not they are enrolled in a particular plan”).

<sup>219</sup> 29 C.F.R. § 1630.14(d)(3)(iv).

<sup>220</sup> 81 FR 31126, 31142 (May 17, 2016).

<sup>221</sup> 29 C.F.R. § 1630.14(d)(1).

benchmarks with respect to how much burden shifting is too much would go far in clarifying this notice provision for employers.

Third, and most importantly, the interpretive guidance warns against “imposing a penalty solely on an employee’s failure to achieve a particular health outcome (such as failing to attain a certain weight or cholesterol level).”<sup>222</sup> This statement by the agency appears to indicate that health-contingent programs may only require relative health metric improvements of participants, not absolute (and sometimes, individually unachievable) standards. In other words, for a person who is obese, losing 5 percent of body mass might be realistically attainable, but becoming normal weight may not be realistically attainable. This statement in the guidance potentially appears to soften the worrisome blow faced by minorities from these new regulations, but without question, the agency needs to issue additional guidance to clarify what types health metric improvements employers may reasonably require of participants in a health-contingent wellness program.

### **C. Improper venue**

A final concern with allowing employers to analyze wellness program data is improper venue—that is, medical researchers, and not employers, are best suited to assess the effects of wellness programs on health metrics and productivity.<sup>223</sup> Researchers, so the argument goes, are better equipped to deal with collecting and securing sensitive health data, and they lack the potential

---

<sup>222</sup> 81 FR 31126, 31133 (May 17, 2016).

<sup>223</sup> For an example of a comment that raised this issue during the notice-and-comment period, see National Women’s Law Center, *National Women’s Law Center Supporter Comments on RIN 3046-AB01* (June 19, 2015), available at <https://www.regulations.gov/docketBrowser?rpp=50&so=DESC&sb=commentDueDate&po=0&dct=PS&D=EEOC-2015-0006> (“Some HRAs ask employees if they are pregnant or planning on becoming pregnant. Employers have no business being the ones who do this. I see nothing wrong in wellness programs, but info should be limited to medical personnel.”).

ulterior motives of employers to take adverse employment actions on the basis of poor health. It is certainly true that researchers have more experience and institutional oversight in dealing with confidential information than do most employers. Nonetheless, employers are still in the best position to collect employee health metric data, and recall from Part V(A) that the new regulations actually encourage the analysis of such data by third parties.

Employer wellness program data from the field—unlike researcher-generated data—has the potential to avoid many of the sample selection bias and external validity concerns prevalent in medical wellness research. Medical research studies generally randomize treatment, but they cannot always recruit a subject pool representative of the population in a clinical setting. To the extent that a subject pool is unbalanced or favors a particular group (for example, individuals affected by poverty, a particular racial or ethnic group), the study results may not be externally valid beyond that group.<sup>224</sup> Data from employer wellness programs may recruit a more representative subject pool because a very large portion of the population works<sup>225</sup>—especially since employers are now allowed to use larger financial incentives to encourage universal wellness program participation. Indeed, the greater ability of employers to induce wellness

---

<sup>224</sup> For a discussion of this and other external validity issues with randomized clinical research trials, see Peter M. Rothwell, *External Validity of Randomised Controlled Trials: “To whom do the results of this trial apply?”*, 365 LANCET 82, 82-93 (Jan. 1, 2005).

<sup>225</sup> According to a 2015 Report from the Bureau of Labor Statistics, 62.7 percent of the civilian noninstitutional population is in the labor force, and 59.3 percent of the civilian noninstitutional population is currently employed. See Bureau of Labor Statistics, *Employment Status of the Civilian Noninstitutional Population by Age, Sex, and Race, 2015 Annual Averages* (2016), CURRENT POPULATION SURVEY, available at <http://www.bls.gov/cps/cpsaat03.pdf>.

program participation under the new regulations should significantly improve the external validity of employer-generated data.<sup>226</sup>

Furthermore, even a clinical research study that has a completely representative subject pool is not in a position to assess the effects of health metrics on productivity in a realistic manner.

Medical researchers can ask individuals of varying health statuses to complete certain tasks and then compare outcomes and timeliness of task performance. Within medical research on obesity, for example, a few studies have investigated differences in obese and non-obese subjects' ability to complete tasks, such walk one-fourth mile, walk up ten stairs without resting, kneel, lift ten pounds, walk between rooms on the same floor, and stand from an armless chair.<sup>227</sup> Although lifting ten pounds may be relevant to some jobs, it is not relevant to all jobs. Even for jobs in which it is relevant, a person's ability to complete such tasks one time in a clinical setting may not reveal much about a person's ability to complete such tasks repeatedly on the job. Only employers can collect data on productivity in a realistic setting: the workplace. And because the recent EEOC regulations have opened the door to the matching of this productivity data to wellness records, employer-generated data can provide unique insight into the effects of various health conditions on actual job performance in a more authentic setting than clinical trials.

---

<sup>226</sup> For a discussion of the importance of recruiting as many employees as possible to improve the external validity of workplace wellness data, see Gemma C. Ryde et al., *Recruitment Rates in Workplace Physical Activity Interventions*, 27 *Am. J. Health Promotion* e101, e101-e112 (May 2013).

<sup>227</sup> See, e.g., Andrea L. Hergenroeder, Jennifer S. Brach, Amy D. Otto, Patrick J. Sparto, & John M. Jakicic, *The Influence of Body Mass Index on Self-Report and Performance-Based Measures of Physical Function in Adult Women*, 22 *CARDIOPULMONARY PHYSICAL THERAPY J.* 11, 11-20 (2011); Dawn E. Alley & Virginia W. Chang, *The Changing Relationship of Obesity and Disability*, 298 *J. AM. MED. ASS'N* 2020, 2020-27 (2007); Kenneth F. Ferraro, Ya-Ping Su, Randall J. Gretebeck, David R. Black, & Stephen F. Badylak, *Body Mass Index and Disability in Adulthood: A 20-Year Panel Study*, 92 *AM. J. PUB. HEALTH* 834, 834-40 (2002).



Finally, employer-generated data from the field provides an exceptional opportunity to understand another important, yet understudied, aspect of wellness programs: peer effects. Assessing peer effects in a clinical setting is challenging—at best, researchers can group together recruited subjects into teams, yet team members may have nothing in common and no reason to see each other outside the context of the clinical trial. The resulting peer group, therefore, is somewhat forced and unrealistic. Coworkers, however, form a more natural peer group that is present in an individual’s life almost every day of the week. Prior research has demonstrated that coworker influence may lead to positive peer spillovers outside the wellness setting. One study, for example, found that after incentivizing some employees to attend a retirement plan benefit fair, not only did the incentivized employees enroll in a retirement plan at higher rates, but so did their coworkers within the same department. (Presumably, the study authors concluded, the nonincentivized coworkers had accompanied the incentivized employees to the benefit fair.)<sup>228</sup> Positive coworker spillovers could plausibly play a similar role in the wellness setting; the recent EEOC regulations open the door to employers collecting the necessary data and researchers studying this potentially important method of improving health.

## **Conclusion**

Disability rights advocates and scholars have reacted harshly to the EEOC’s 2016 regulations and guidance on workplace wellness programs, decrying its potential to increase health-status-based discrimination in the workplace.<sup>229</sup> More access to employee health data—so their argument went—would enable employers to identify more disabled workers, and in turn, to take

---

<sup>228</sup> See Esther Duflo & Emmanuel Saez, *The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment*, 118 Q. J. ECON. 815, 815-42 (2003).

<sup>229</sup> See *supra* notes 81, 91, 102 and accompanying text.

more adverse employment actions against them. But this argument ignores the fact that the new regulations have opened the door to better data for researchers, which can (and likely will) produce data that can bolster advocacy efforts on behalf of the disabled community. It further ignores the fact that employers are already making erroneous, harmful assumptions about the abilities of workers with visible health conditions. The present lack of data that can credibly illuminate the relationship, if any, between wellness programs, health, and workplace productivity has allowed such assumptions by employers and the general public to persist unchecked. Data can help researchers, policy makers, and employers understand that a health condition may not necessarily render a worker less productive. Data can also help researchers, policy makers, and employers understand how best to place and accommodate workers with a health condition.

Despite the fact that the ADA has existed for a more than a quarter of a century, workers with health conditions have continued to encounter labor market discrimination based on harmful employer stereotypes.<sup>230</sup> The prior EEOC regulations and guidance, which discouraged employers from collecting health data, also apparently failed to teach employers that “physical or mental disabilities in no way diminish a person's right to fully participate in all aspects of society.”<sup>231</sup> Multiple surveys documenting continued and widespread employer beliefs about the

---

<sup>230</sup> See, e.g., Jennifer Bennett Shinall, *What Happens When the Definition of Disability Changes? The Case of Obesity*, 5(2) IZA J. LAB. ECON. 1, 1-30 (2016) (showing that obese workers have still continued to experience worse employment outcomes even after federal courts began extending ADA coverage to obesity); Daron Acemoglu and Joshua D. Angrist, *Consequences of Employment Protection? The Case of the Americans with Disabilities Act*, 109 J. POL. ECON. 915, 915-57 (2001) (demonstrating that labor market outcomes for disabled workers actually declined after the passage of the ADA); Marjorie L. Baldwin & William G. Johnson, *Labor Market Discrimination against Men with Disabilities in the Year of the ADA*, 66 S. ECON. J. 548, 548-66 (2000) (demonstrating a correlation between wages and stigma associated with the underlying disability).

<sup>231</sup> Americans with Disabilities Act of 1990, *Pub. L. 101-336*, § 101(a)(1).

negative effects of health conditions on employee productivity demonstrate as such.<sup>232</sup> The time has come to take a new approach—and open the door to the possibilities of data and empirical analysis, giving them a fair chance to correct the injurious assumptions still prevalent about health conditions in the workplace.

Even just a handful of employers taking advantage of the new regulations and sharing their data with academic researchers would represent a marked improvement over the prior regime.

Nevertheless, this recent, positive shift in EEOC policy has become endangered within its first few months of life. The new regulations were issued under the authority granted to the agency by the ACA, and incoming federal elected officials have made widely known their intent to undermine enforcement of—and if possible, repeal—the statute.<sup>233</sup> Undoubtedly, the healthcare marketplace provisions within the ACA will be these officials' first target, but whether the entire Act is in jeopardy remains unclear. Already some signals suggest that not all incoming officials support a repeal of the entire Act,<sup>234</sup> and given that the wellness program provisions do not pose additional costs on employers or the government, these provisions are more likely than others to remain untouched by Congress. And as this article has argued, these provisions—and the resulting regulations—should remain untouched by incoming federal officials. Reverting back to

---

<sup>232</sup> See *supra* notes 11-12 and accompanying text.

<sup>233</sup> See, e.g., Robert Pear & Thomas Kaplan, *G.O.P. Plans to Replace Health Care Law with 'Universal Access,'* N.Y. TIMES, Dec. 15, 2016, at A20 (“House Republicans . . . [are] preparing for a rapid legislative strike on the law next month”).

<sup>234</sup> See, e.g., Richard P. Asensio, *What Trumpcare Might Actually Look Like*, FORTUNE, Dec. 18, 2016, <http://fortune.com/2016/12/18/obamacare-trumpcare-health-reform-republicans/> (“[Trump] has indicated his desire that most of Obamacare be repealed, while retaining popular provisions such as allowing children to stay on their parents' plans until they're 26 and forcing insurers to cover people with pre-existing conditions.”).

the prior regime will do nothing to advance public understanding regarding wellness, health, and productivity, and more importantly, will do nothing to assist the disabled community.

**Part III: Standards and Measurements in Business,  
Organizations, and Law Firms**

**Chapter 8: An Innovative Approach to Hiring Lawyers: One Firm’s New Program Reflects Its Firm Values and Eliminates Implicit Bias /**  
*Lisa A. Brown, Professional Development Partner, Schiff Hardin LLP*<sup>235</sup>

**I. Introduction: The “Traditional” Hiring Process and the Reasons for Abandoning It**

In 2011, Schiff Hardin LLP, an AmLaw 200 firm headquartered in Chicago, determined that the traditional law firm interviewing process was not bringing the firm the talented, diverse lawyers it sought. Large firms like Schiff Hardin had for years used virtually the same cookie-cutter interview process for entry-level associate hiring. It consisted of an on-campus screening interview, followed by a callback interview that consisted of four 30-minute one-on-one interviews and a lunch.

We knew our hiring goals and values were different from other firms’. Shouldn’t our recruiting process also be different from other firms’? This question made us take a step back and ask ourselves what we were looking for. We found several answers.

Most fundamentally, we wanted new lawyers who valued what we value: collaborating with colleagues, focusing on our clients’ needs, communicating clearly in writing and orally, taking ownership of developing their careers, and learning and seeking out new and interesting work challenges from day one.

In addition, Schiff Hardin had long had a unique associate development model, and we sought a recruiting process that would complement it. We do not hire new associates into practice groups

---

<sup>235</sup> A version of this essay was originally published in IILP Review 2017: The State of Diversity and Inclusion in the Legal Profession. It is re-printed with permission from the Institute for Inclusion in the Legal Profession.

but rather let them spend up to a year exploring different areas. We focus on associate training and have a full-time legal writing coach in-house who hosts workshops and works one-on-one with our newest associates (and with more experienced lawyers). Further, associates get early experience because most Schiff Hardin teams include only one partner and one associate who work closely together. Finally, associates develop and advance at their own pace. Our competency structure is flexible and does not limit associates to lock-step advancement with their class. Instead, associates progress and are advanced based on their individual merit.

We determined that law students are much more than their grades and academic qualifications. And we found that grades and academic successes alone were not strong predictors of success at Schiff Hardin. We were further concerned that the traditional interviewing process could be implicitly biased against diverse candidates. Finally, we needed information that would show us whether candidates had the attributes to succeed (and be happy) practicing law at Schiff Hardin. We wanted a more complete understanding of our candidates and a process that was objective and effective.

Our research further showed that one of the most frequent reasons younger associates did not succeed at Schiff Hardin was because of their written communication skills. For that reason, we looked for an early way to analyze candidates' writing, both for screening purposes but also to determine how we might help someone with writing challenges.

## **II. How the New Hiring Program Works**

The new hiring program has several different parts. In addition, we retained some of our old system, including one-on-one interviews, a review of law school writing samples, and a lunch with associates.

First, we expanded the pool of candidates we consider. We felt comfortable interviewing at more law schools and more job fairs because of our new callback process. Between 2009 and 2014 we more than doubled the number of law schools we visited, including an HBCU, and committed to interviewing at several job fairs that focused on diverse candidates, including the Cook County Minority Job Fair, the Southeastern Minority Job Fair, and the National LGBT Bar Association Lavender Law Career Fair. We also committed to interviewing candidates with a greater range of grades, eschewing a threshold grade or class rank requirement.

Second, we created a new callback interview format. During callbacks, candidates interview with a group of three to four partners (the “panel interview”) for an hour. The partners take turns asking behavioral interview questions designed to gather more information about the candidates, including their work, academic, extracurricular, community, and other individual life experiences. The questions explore candidates’ experiences solving real-world problems, working with and leading teams, learning new skills, resolving conflicts, and building successful relationships. Those are all traits associated with long-term success at Schiff Hardin. The format is substantive and interactive; the tone is rigorous and dynamic.



We also tried to eliminate any implicit bias in the interview process by making the panel interviews more objective and by making the panel interviewers more accountable. Interviewers do not receive candidates' law school transcripts. We include at least one racially diverse, female, or LGBT partner on each interview panel and put all interviewers through the same rigorous training program. Further, the "structured" aspect of the panel means that interview scores do not depend on personal connections or the idiosyncratic leanings of particular interviewers. The panel follows a standardized behavioral interview format. We cover the same topics with every candidate, and ask virtually the same questions, digging deeper with customized follow-up questions.

The evaluation process also ensures that the four interviewers "own" their evaluation more than they do in a one-on-one interview. After conducting a panel interview, the panel members discuss the candidate's responses and work together to reach a consensus evaluation of the candidate. With this process, panel members cannot rely on "gut feel" but must instead articulate and defend their evaluations on the basis of whether the candidate has demonstrated specific traits and characteristics. Panel members then broker consensus as a group. In addition, because they've spent an hour with the candidate – rather than the typical 30 minutes -- and more time discussing the interview, they are more invested in the process and in each individual candidate they interview. Finally, the process eliminates another possible source of implicit bias: the ill-prepared or poor interviewer. This type of interviewer fails to gather relevant information from the candidate and instead falls back on "gut feel" or conventional measures of achievement, such as grades.

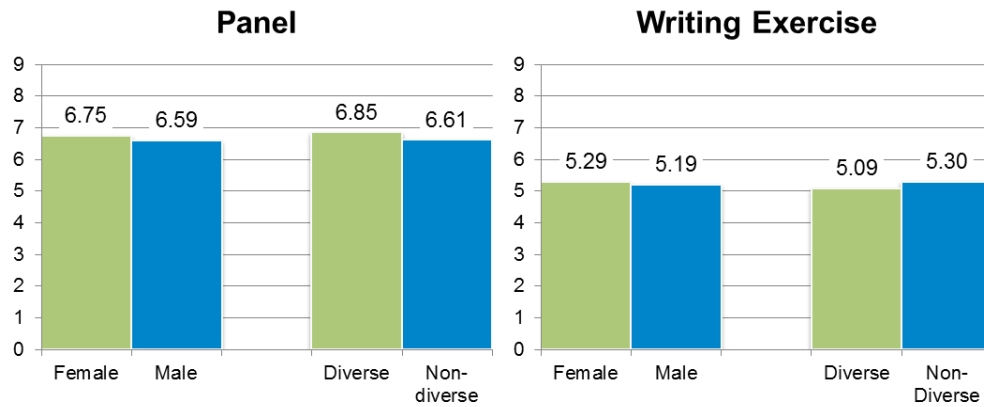
Finally, we added a writing exercise to our callback interview, which the candidate completes while at the firm. We provide a personalized letter to the candidate describing a brief client problem. The problem is discrete and can be addressed by the candidate in the time allotted without any specific knowledge of the subject matter. We ask the candidate to draft a response. This exercise does not resemble any law school assignment that we know about and therefore does not favor candidates who have performed well in legal writing class. Rather, it measures analytical and communication abilities that all lawyers must have: how to read and digest a legal issue and explain it to a lay person who is experiencing a problem. Our evaluation of these exercises is completely blind – the evaluator does not know the race, gender, law school, or any other characteristics of the candidate. Our evaluation focuses both on the tone of the work – especially the candidate’s ability to convey empathy and relate to the writer – as well as the substantive content and writing style.

No one part of our callback process is dispositive. The hiring committee considers all aspects of the interview – the panel interview, the writing exercise, the one-on-one interviews, the lunch interview, as well as the candidate’s paper record.

### **III. The Results**

We now have data that shows that the panel interviews and writing exercise mitigate implicit bias. Women and racially diverse candidates both perform well. Further, the data shows that the new system does not favor students from elite law schools or students with any particular pre-law school work experience.

## Difference in Means Analysis of Panel Interview and Writing Exercise

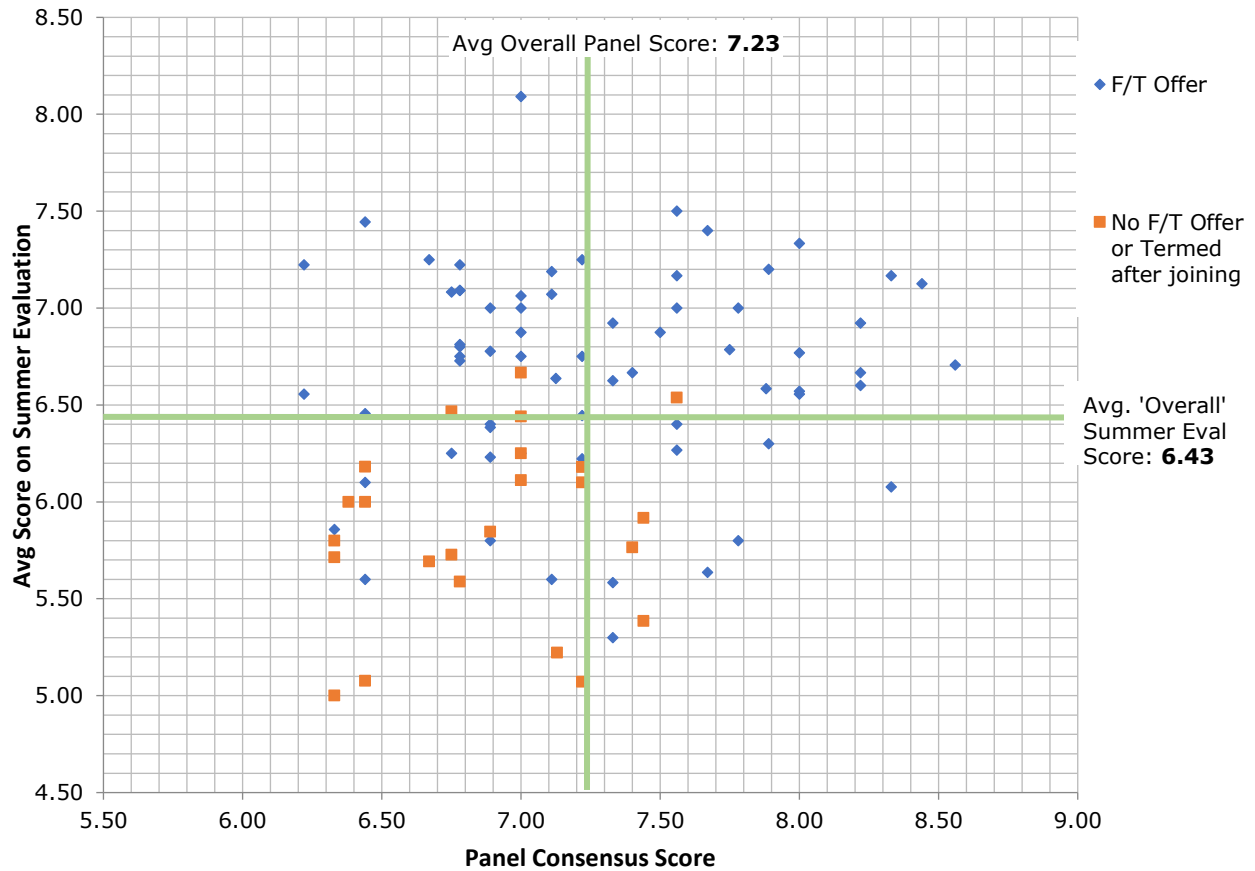


### **AVERAGE PANEL INTERVIEW and WRITING EXERCISE SCORES:**

n = 515 law students interviewed between 2011-2015. Mean comparisons show no statistically significant difference in scores between men, women, racially diverse, or non-racially diverse candidates on the writing exercise. Mean comparisons show that women and racially diverse candidates perform slightly better than their counterparts in the panel interview.

Also, the data shows that high scores on the panel interview and the writing exercise are powerful selection tools. Associates who receive permanent offers and stay at Schiff Hardin for more than one year tend to have performed better in the panel and on the writing exercise:

## Predictive Power of Panel Interview on Summer Associate Performance



In addition to quantitative results, we also have qualitative results. Each year, we engage an outside consultant to gather candidate feedback to ensure that the *experience* of the panel interview is not felt differently by different groups. It isn't. Diverse and non-diverse candidates report that they like having the opportunity to share more of their stories and life experiences than they do in traditional law firm interviews. They also report that the panel interview feels "fairer" than other law firm interviews. They know we are covering the same topics and asking every candidate the same questions, and so their success is not tied to first-impression bias or "hitting it off" personally with the interviewer. Further, during the panel interview, they get a glimpse into the firm's culture – including the relationship among partners and the investment in

associates. Many also like having the ability to show through their writing exercise that they are ready to communicate with clients when they start practicing law.

Finally, our national diversity rankings have also improved since we implemented our new recruiting process. Vault's annual list of the 25 Best Law Firms for Diversity are based on law firms' own associates' rankings of how their firm does at fostering a diverse workforce. Before 2011, Schiff Hardin had not made the rankings. In 2016 we ranked #2, #7, and #9 nationally for best law firms for women, racial minorities, and overall diversity.

#### **IV. Conclusion**

When Schiff Hardin's hiring process shifted from the traditional format of hiring entry level associates, we did not know what effect these changes would have on our recruiting efforts. We have been pleasantly surprised. Since we started interviewing at a larger number of law schools and job fairs, and using the panel interview and writing exercise, more students are signing up for interviews and a higher percentage is accepting offers of summer employment. We have also seen additional diverse candidates, including more women, more LGBT candidates, and more racial minorities, and have been more successful at hiring them.

While we continue to review this process and analyze the results, early signs are promising. We hope the process helps the firm develop and retain all our associates. And there has been one unexpected benefit: the new process differentiates Schiff Hardin and is seen by the marketplace – by candidates, recruiters, and law schools – as part of what makes the firm unique.

**Chapter 9: Excerpt from *Diversity in Action: A Manual for Diversity Professionals in Law – The Measure of Success or What Matters Is Measured and What Is Measured Matters* / Presented by Theresa D. Cropper, Chief Diversity Officer at Perkins Coie LLP, co-authored by Anna L. Brown**

**OVERVIEW**

A metric system is a key component of a diversity initiative. A few questions to ask include:

What is the importance of measuring the composition, succession, distribution, projections, and realities of your organization? Why would or should you collect, analyze, distribute, and review data? Data will give you historical and present-day snapshots of your organization. It is an important tool in making projections and plays a critical role in the diversity development of an organization.

To understand data, you must first understand your organization's systems, structures, titles, and processes, which will differ from organization to organization. You must have a working knowledge of your organization's hiring, evaluation, compensation, promotion, and leadership selection systems. Further, you must understand the published and political hierarchies. You must know all the titles and the trajectories of each and should appreciate the finance, human resources, and data collection processes. Finally, you must know the questions to ask to obtain the information you need to make accurate measurements and assessments.

Data is most important if it is studied and used to education organizational leaders on the diversity realities of the organization. In this section, we will examine the mechanics of data, what to collect, how, and why, the questions to ask, the value of the data, the publication and use of the data, and examples of how data can play a critical role in diversity development.

## **THE MECHANICS OF DATA**

The mechanics of data examines how, what, and why data is gathered. The following are some questions to consider regarding the mechanics of your data:

- What data do you need most? What information is on the top ten request lists?
- How many lawyers are in your organization?
- What is the gender composition?
- What ethnicities are represented?
- How large is your LGBT community?
- Do you track lawyers with disabilities? If so, what is the data?
- What geographic areas are lawyers located?
- What practices do they represent?
- What levels of seniority do they have?
- Who are the leaders in the firm?
- Who are the members of the important committees and decision-making committees?
- Who are we hiring?
- Who is leaving?
- Who is staying? Who is promoted?
- Who gets bonuses? Who doesn't?
- Who is reaching the annual billable house requirement, if any?
- Who isn't?
- Who receives the highest performance reviews? Who receives the lowest?
- Who participates in pitches and is identified in RFPs?

How do you collect this data? In addition to working with your strategic partners in finance and human resources, you must understand how your organization collects data generally and where and why the data is compiled and stored. It is important to establish a fundamental knowledge of a shared language with your data partners. You need to understand if a particular data pool is head count, full time equivalents (FTEs), or timekeepers. You must be aware of the differences and the impact analysis. For example, if you have a significant number of women who are not FTEs, that data pool will not include them, but they will be included in the head count data pool. Timekeepers may include paralegals or other employees who are not lawyers. Finance and human resource teams may operate with different data pools. Is the financial information compatible with human resource information? Is there one system of data that keeps performance, hours, and demographic data such as gender, LGBT and ethnicity? How often is such information updated in the system? What data pools and systems does the marketing department use, if any, to track RFPs and pitches?

It is important to regularly collect, review, and publish data both internally and externally. Many external publications, bar associations, and/or clients or reporting authorities request data on a regular basis. How should data be managed internally? How often should you review data and what data should be reviewed on a regular basis? Who reviews it?

### **Monthly Data Recommendations**

The following data should be collected and reviewed monthly:



- Overall numbers: total population, with all categories (to include gender, ethnicity, LGBT, disability, and veteran status)
- Overall numbers by location with the same categories
- Overall numbers by practice area or department with the same categories
- Further breakdown of ethnic categories (not included in the EEOC categories) such as South Asian and Middle Eastern
- Comprehensive list of all lawyers in all categories, including practice area, location, and tenure with organization
- Comprehensive list of all hires and all departures
- Billable hours pace or other performance-based assessment of junior to senior lawyers and associates

### **Quarterly Data Recommendations**

The following data should be collected and reviewed quarterly:

- Pitch and RFP data reports for all categories (to include gender, ethnicity, LGBT, and disability status)
- Billable hours or performance on top clients and top matters by lawyer demographics
- Hires
- Departures

### **Annual Data Recommendations**

The following data should be collected and reviewed annually:

- Leadership composition by demographics

- Identify demographics of practice group leaders, department leaders, office leaders, and executive and management committees
- Committee composition by demographics
- The composition of impact/decision-making committees
- Evaluations or quantitative assessments of trajectories (on track for partnership, directorships, bonuses, performance bonuses)
- Executive/partner compensation
- Salary review and assessment of any disparities
- Attrition rates for the practice and by department, location, demographics (gender, ethnicity, LBGT, and disability), and tenures
- Retention rates for the practice and by department, location, demographics, and tenures
- Exit interview data, if available

Why do you need this data? It will provide you with the tools to assess status and progress, to understand challenges and opportunities, and to evaluate and develop solutions. Studying this data will provide the answers to critical questions that are necessary and inevitable when evaluating, developing, or assessing diversity efforts in any organization.

### **ADDITIONAL QUESTIONS TO CONSIDER**

- What percentage of executives or partners are women? What percentage of junior lawyers or associates are women? What percentage of leaders in general are women? Is the organization's overall percentage of women the same as the percentage of women in leadership, executive/partner ranks, or at the higher compensation levels?

- Ask the same questions regarding ethnicity. What is the data for each specific ethnicity?
- What are those percentages for disabled and LGBT employees?
- Is the organization led and managed by the same demographic percentages represented in the overall population?
- Who is working for the top clients and on high-profile matters? What are the demographics of the client leads and project leads for those high-profile matters? Who are the top billers for such clients and who puts in top hours for highly visible key matters?
- Who is being identified for client contact with pitches?
- Who is being identified in RFPs?
- Does the organization track the status of RFPs and audit whether those identified in the RFP are actually assigned the work on the matter when the business is obtained or assigned?
- Who is being identified for key teams? How is that work distributed and monitored?
- What is the demographic composition of the higher compensation levels?
- Who is performing on pace? Who is not? What are their practice areas? Where are they located?
- Is there any correlation between the demographics of departures and practice areas and/or locations?
- Is there any correlation between the demographics of promotions and practice areas and/or locations?
- Are there any concerns raised by the evaluation process?

- Are there any concerns regarding retention at particular tenure points? Are lawyers departing at a particular tenure period?
- Are there representational issues in any practice area, department or location? Perhaps there are no women in executive roles or with partner status. Perhaps there are practice areas, departments, or locations with a particular ethnicity that is not represented or a “lonely only.”
- What are departing lawyers citing as reasons for leaving? Is there any data suggesting that a disproportionate amount of departures reside in a particular practice area, department, or location?

## **THE VALUE OF DATA**

The value of comprehensive data is significant. It is most valuable to the organization and to advancing diversity and inclusion goals when the diversity practitioner studies the data and makes an analysis. After the questions have been asked and the data collected, the study and analysis must begin. The diversity professional must review historical data and regularly review new data. If data is collected regularly, a comparative analysis should be part of the periodic process. Study and review the month-to-month data to discern if any trends are developing. Quarterly data should be reviewed on a rolling year-to-year basis, comparing quarters to quarters (e.g., compare Q1 2012 to Q1 2011). Annual data should be assessed each year and then in three-, four, and five-year increments. The data must be studied and reviewed and then put in the context of the organization’s diversity efforts. The data also should be evaluated for trends, peaks, valleys, and a steady pace, up or down. It should be viewed as a tool for education; it is the foundation for new policies and program and it provides direction for future diversity efforts.

The data establishes a baseline for assessment. This assessment will tell the story of successes and challenges. It has often been said that “what gets measured, gets done.” The data will reveal where work is needed and where systems have been effective and ineffective. It will show the status of the organization’s overall diversity representation, what can be expected, and if there is more work to be done in certain areas, such as recruitment efforts. The data will tell you where to concentrate certain efforts and will illuminate if certain groups are not being included and where. It will identify patterns and trends with respect to who is leaving and from which departments or locations. This will help you discern if your data represents a recruitment issue or a retention issue. In other words, are you bringing the attorneys in, but not keeping them, or are the attorneys staying, but you are not recruiting with high levels of diversity?

The evaluation process data will help you understand the performances of your lawyers by demographic. It will show who is performing well, who is receiving, merit bonuses, if applicable, and who is receiving production bonuses, from the perspective of gender, race, LGBT status, and other diversity areas. The data will also identify those who are underperforming and their practice areas and offices. A careful study will show whether there are any trends with performance that correlate with tenure, office, practice areas, demographics, and any combination thereof.

An analysis of the data pipeline will show leadership and promotion projections. It will reveal whether there are people from underrepresented groups who are in a position to ascent to leadership or key committees. Further, it will assist you in identifying opportunities for

placement key committees and in leadership development programs. The data will establish trends and will show who has served on committees and in leadership roles, and who is available. If there is no apparent pool, then the data will reveal exactly where to start building the pipeline and how long it will take through organic growth, without lateral hire intervention. The data also will reveal deficiencies in particular tenures statuses, whether midlevel, senior, or partner/executive, and the impact of those deficiencies on diversity efforts. The data will help you to examine the nature and timing of the pipeline, which can be critical in shaping future decisions.

Another key area in which the data is critical is with respect to RFPs, client pitches, and access to key clients and matters. The data will identify patterns and trends with respect to who has had opportunities to work on RFPs, pitches, and key client or high-profile matters. The data will provide information on the number of RFPs or pitches and who has been identified in these efforts. It will also tell you about demographic representation among the lawyers identified in the RFP and whether there is a pattern of repeat names, that is, whether the same people are typically identified for RFPs. A system of data collection will tell you if those identified in the RFP actually received the work when the firm or department was successful, and who is most active in leading RFPs by attorney, practice group, department, and office. This is important information when raising awareness and fashioning remedies to the challenges identified through the study of such data.

Compensation data is often extremely sensitive both in terms of confidentiality and from a risk management perspective. Notwithstanding, this data can be aggregated directly and is critical in

identifying the presence of disparities and determining whether there may be systemic issues in promotion, access to key client and high-profile opportunities, or other systems that can affect compensation. (For example, in a law firm, how a client matter is opened; how credit is distributed among partners/lawyers, etc.) Further, the study of compensation can be at the junior/associate or the executive/partner level. Discoveries regarding compensation require high-level examination, as these are highly confidential and important areas.

Issues regarding practice areas and office locations can also be examined. Such an examination may include the overall compensation, the leadership, the clients, the demographics of the different levels of tenure, the pipeline, the nature of the evaluations, and the levels of compensation. It will be a story that has both obvious and surprise chapters that will be important to understand and appreciate when implementing programs in the diversity efforts. The study and analysis of this data is a task best done with the leaders of relevant areas. This will yield mutual appreciation as well as a leader's perspective with historical context. It will also help create an understanding if there is a need for careful monitoring or intervention – such as a “lonely only” or a potential lateral who would address a particular deficiency (e.g., if there are no female partners, but there is a strong lateral partner candidate who is female). The data will also tell you who is leaving and the groups and offices where there is greater or lesser attorney retention. This could be important information when responding to exit interview information. You can learn why people leave, but you also need to understand why people stay.

## **THE PUBLICATION AND USE OF DATA**

Let's start with who looks at this data and why. Should all data be made available to anyone internally? How should the dissemination of data be managed? Why is this important? The publication and distribution of data internally must be strategic and done in a manner that ensures there is always integrity with public data. How should the dissemination of data be managed? One person should be the final repository of all data and provide the final draft of the data for approval by the diversity professional. Any and all requests for data, whether a casual request by a practice area leader or a response to a client survey, must go through one central person. This ensures that dissemination of the data is done within the context of all requests; this person knows the difference between requests and the nature of the data pool involved in the response. Using one central person will ensure there is a consistent "apples to apples" comparative analysis or that explanations are provided when different pools of data are used and appear to be inconsistent. For example, the number of lawyers may appear different when using FTE instead of head count, or if the data is limited to U.S. offices rather than being global. Having one person who serves as the focal point for distributing data supports the integrity of the system, diminishes the opportunity for disparity, and increases the ability to quickly resolve misunderstandings.

Who looks at the data and why? There are public sources of information about law firms and, in certain instances, law departments, and certain data is easily retrievable online. There can and should be quick and easy access to this information, for example, through the National Association for Law Placement (NALP), American Lawyer Media, Vault, and so on. Information that is not public should be distributed judiciously and strategically. Data should be shared with



the executive leadership of the firm. Quarterly reports to the firm or department manager can be helpful in the diversity educational process and to practice areas and office or department leaders. Further, data should be shared with the leadership of key committees for understanding, awareness, and discussion. By collectively studying the data with these key individuals and groups, the diversity professional can develop better relationships and build support for diversity initiatives within the firm.

Should all data be made available to anyone internally? Is it prudent to provide all the data that is available to the diversity professional to the diversity committee, department, or firm? If the data needs an explanation, then it is probably best to share when it is specifically requested; however, most data needs an explanation. Discussion is always needed when questions arise, for example, as already discussed, regarding the distinction between head count and FTE. Sharing the data that is already public means that everyone has the same data and there is a shared knowledge of the intricacies of such data. Sharing other data that is not public and may be unsolicited can lead to confusion and additional discussion to clarify or explain the data.

Why is it important to always have confidence in the integrity of your data? The diversity professional should have the last look at the data before it is published. This is to ensure that the data is responsive, self-explanatory, and is consistent with and/or distinguishable from any other data that is public. Regardless of whether you like the numbers or not, they must be above reproach and the organization must be able to rely on the integrity of its diversity data.

## **HOW DATA CAN PLAY A CRITICAL ROLE**

Data is a critical tool in your diversity work. It can be the foundation of relationships with firm or department leaders, the impetus for a change in policy or strategic direction, or the story that provides epiphanies on a systemic bottleneck or ineffectiveness. Data is an important element in developing the strategic plan and picture. It will provide the measurement by which you can evaluate programs and processes – a key factor in determining success. Data can also be distracting, confusing, and alienating. It is a powerful tool that should be used with strong analysis and a sensitive perspective. The diversity professional must be careful to use data effectively and without indictment, even if the data is not flattering or provides challenges to a particular area of the organization. An astute diversity professional can create important educational campaigns based on the needs identified in the data. For example, data can raise awareness and strengthen discussions regarding a lack of gender or ethnic diversity in a particular office or location. Data helps to distinguish between a perceived deficit or concern and a deficit or concern that is statistically determined.

## **WHAT YOU SHOULD KNOW**

- Always, always ensure that your data has integrity. Check and double check. Errors are unacceptable.
- Try to keep an open mind about what you may learn. Expect to be surprised. If you think you know what the data will tell you, then your bias is already set.
- Run your numbers before you are expected to measure them. If it is annual data, check the trends quarterly. If it is quarterly data, check the monthly information. Do not be caught off guard because you were not looking for trends.

- Ask a lot of different question about the data to take a comprehensive look at it. Analyze it from others' perspectives. Defend the data. Attack the data. Look at the data from many angles. What does it tell you? What does it tell HR? Recruiting? Etc.
- Ask if you are tracking enough data or too much. Can you explain and rationalize all the data reports your department generates? If not, then do so.
- If the data is not helpful or educational, you must question why it is collected.

**Chapter 10: Diversity and Inclusion Standard and Practice in the 21st Century / *Effenus Henderson, Co-Founder of the Institute for Sustainable Diversity and Inclusion; President and CEO at HenderWorks Inc.***

CEO Henderworks, Inc.

Co-Director, Institute for Sustainable Diversity and Inclusion

Co-Chair, SHRM Diversity and Inclusion Standards Taskforce

Convener, Workgroup 8 – Diversity and Inclusion, Technical Committee 260, ISO

May 30, 2017

**Abstract**

Standards for assessing the effectiveness of diversity and inclusion efforts in organizations play a very important role in a time of significant transition and disruptive change. Organizational leaders must understand the impact of shifting demographics on our workforces, our government, educational systems, customers and communities. Standards provide a very useful tool for gauging progress and setting expectations for performance. As the importance of diverse workplaces and inclusive leadership and governance grows, and as the world becomes much more demographically diverse, leaders must understand how to engage and collaborate with diverse employees and civil society members. Inclusive leadership behavior, institutional policies, practices and are increasing in important in today's workplace. These behaviors and practices are being aggressively examined by diverse activists, investors and regulatory agencies. This is being done through a variety of group based identities (such as race, ethnicity, gender, age, sexual orientation, disability and cognitive differences). Standards provide a common language, a minimum set of acceptable criteria for examining elements of an organization's diversity and inclusion framework. Such efforts help to provide a consistent set of metrics to examine the effectiveness of an organization's efforts. Standards are not a panacea for diversity and inclusion organizational development and change efforts, but one of many tools that should be employed to address these important factors.

**Keywords:** Diversity, inclusion, standards, compliance, affirmative action, micro-aggression, human resource management, inclusive behavior, ISO, ANSI, ASTM, SHRM.

Diversity and Inclusion Standards and Practice in the 21st Century

As a corporate executive with significant experience in the human resources profession covering work and relationships in over 18 countries and five continents, I am convinced that the development and establishment of diversity and inclusion or D&I standards, both domestically in the United States and internationally, can play a pivotal role in the development and sustainability of inclusive organizations and institutional practices.

Historically, diversity and inclusion strategies have been driven by governmental compliance, risk avoidance and framed within a legalistic oriented mindset. Leaders focus on ways to avoid discrimination lawsuits, costly settlements and the impact of far reaching conciliation agreements. While such concerns are very important foundational areas to pay attention to, they are just one of several areas that should be monitored closely. Organizational policies, systems and practices should be assessed and monitored for disparate impact on minority and female employees as well as other demographic groups within the organization. Leadership behavior should reinforce the organization's commitment to diversity, and be aligned with respective and inclusive organizational values.

Based on my experience as a practitioner and consultant in diversity and inclusion, I believe there are nine important drivers that should be considered and their impact integrated into the work of human resource professionals, diversity practitioners and organizational leaders.

Leaders should consider the impact of these drivers on governmental, business and organizational strategies and outcomes. Assessing workforce trends in these areas can be useful in evaluating the organization's results at the individual, departmental and organizational levels.

Understanding these nine trends and building action plans to address gaps can enable organizations to be more competitive, agile, and able to respond to changing workplace, regulatory and market demands. The nine major trends:

1. The exponential growth in the numbers of people of color and immigrants will result in unprecedented demographic change in the US and global population, especially in urban centers.
2. Innovation and creativity to vitalize new and emerging markets and economies will be enhanced by growing cognitive diversity (cultural, thinking and linguistic differences)
3. Organizations will be faced with increased civic activism by a restless middle class and the economically disadvantaged who feel increasingly marginalized and adversely impacted by the changing business landscape.
4. Civil society is losing confidence and trust in many governmental leaders based on their disrespectful behavior and conduct, corruption and lack of integrity.
5. The evolving networked and shared economy is creating new and evolving employment value propositions impacting employee engagement and retention.
6. This new shared economy is changing, disrupting and restructuring global supply chains
7. Workplace habits and attitudes of millennials are creating a collision of values between older Boomers and newer entrants into the workplace.
8. Social media has changed the way information is shared and communicated with immediate impact on personal and organizational reputation and brand management.
9. These social media changes are impacting leadership, marketing and communications strategies.

Diversity and inclusion efforts have evolved over the past four decades to consider most of these trends. Early efforts were focused on regulatory and compliance requirements resulting from various equal opportunity laws and regulations requiring organizations to take good faith efforts to improve the diversity of their workforces. The United States Department of Labor's Office of Federal Contract Compliance Programs monitors such efforts at federal contractors by reviewing their affirmative action programs. When violations are found in their affirmative action programs contractors are required to make adjustments in their processes, systems and practices so that they are more inclusive.

As diversity and inclusion practices have matured, organizational strategies have shifted from compliance related remedies to plans and initiatives that seek to "monetize the value of diversity and inclusion practices." Emerging practices extend beyond federal contractors and are being incorporated by many progressive employers. These evolving practices are more strategic and aimed at satisfying stakeholders who are becoming more diverse. These stakeholders include employees, customers, suppliers, investors, community leaders, and government regulators. These stakeholders are assessing and monitoring the practices of organizations. With this increased focus on best practices, employers are focusing on relevant metrics and benchmarks to be more effective. Organizations such as Diversity, Inc., the Conference Board, Working Mother Media, and others are reviewing company practices and ranking companies based on their diversity and inclusion practices. However, the methodology and focus areas are often varied among these rankings and few standardized tools exist to assess effectiveness.

In 2006, Julie O'Mara and Alan Richter helped lead the development of some early benchmarks for diversity which were published by the Tennessee Valley Authority. This work was later



revised and updated to become what is known as the Global Diversity and Inclusion Benchmarks,<sup>236</sup> which provide a comprehensive framework for evaluating diversity efforts within organizations. The most recent edition, published in 2016 and which includes input from 95 Expert Panelists with deep expertise in D&I, underscores the highly researched nature of these benchmarks.

Since this early work in 2006, more and more organizations are convinced of the need for standards in the diversity and inclusion area. In 2009, the Society for Human Resource Management (SHRM) convened a group of diversity thought leaders to assess if additional work should be done. One recommendation from the group was to form a national taskforce under the auspices of SHRM. In partnership with the American National Standards Institute (ANSI), SHRM decided to sponsor the development of a set of national “minimum standards” in diversity and inclusion with emphasis on competencies for top diversity leaders, minimum requirements for diversity programs, and metrics to assess the effectiveness of organizational D&I practice.

### **The National and International Debate About Standards**

Rosemary Hays-Thomas and Marc Bendick, Jr, in an article entitled Professionalizing Diversity and Inclusion Practice: Should Voluntary Standards be the Chicken or the Egg, which was published in the *Industrial and Organizational Psychology: Perspectives on Science and Practice* (11/2012), commented: “workplace diversity and inclusion practices today are based primarily on unevaluated experience and intuition rather than empirical evidence. Would voluntary professional practice standards in this field effectively raise the level of current and future

---

<sup>236</sup> <http://diversitycollegium.org/downloadgdib.php>

practice? They identified four barriers that must be overcome: “(1) limited evaluation of D&I practices (addressable through research); (2) limited conceptual and analytical skills among D&I practitioners (addressable through education and training); (3) employer concerns about self-incrimination (addressable through policy advocacy); and (4) limited employer commitment to D & I itself (addressable through conceptual rigor).”<sup>237</sup>

Rosemary and Marc offered very important considerations for those considering the development and implementation of D&I standards. Their insights were very important as we addressed and crafted standards. The foundational work by O’Mara and Richter and 95 Expert Panelists also helped shape the work of the SHRM Diversity and Inclusion standards work. It was also important in the development of global standards as well.

### **What is a Standard?**

A standard is a document that provides minimum requirements, specifications, guidelines or characteristics that can be used consistently to ensure that materials, products, processes and services are fit for their purpose.

“In order to evaluate performance you need standards – every performance auditor understands that. Aside from performance evaluation (which is justification enough) setting HR standards improves the quality and consistency of HR practice, clarifies the value created by HR management, focuses on better practice with timely revisions to practitioner guidance, clarifies definitions and evaluation metrics and measures, identifies

---

<sup>237</sup> Hays-Thomas, R. and Bendick, M. (2013), Professionalizing Diversity and Inclusion Practice: Should Voluntary Standards Be the Chicken or the Egg? *Ind Organ Psychol*, 6: 193–205

sources of evidence for performance evaluation and establishes a sound platform for big data analytics.”<sup>238</sup>

Standards must be framed in a language that business leaders understand. We must be sensitive to the questions raised by corporate leaders and make sure that we are addressing their issues and concerns. The following questions are often asked by corporate leaders and must be addressed by those seeking to develop standards:

1. What are the most influential metrics being used in professional workplaces?
2. How sound are these metrics, and how are they deployed?
3. What is the potential for developing standard data collection protocols to improve quality, interpretation, and effectiveness?
4. Given the potential threat of litigation, does law impede the use of diversity metrics?
5. What is the relationship between globalization and diversity metrics?
6. How do such standards contribute to business outcomes?

And, the administrative impact to the organization must be considered as described by Hays-Thomas and Bendick. The following challenges must be addressed:

- Administrative Burden - Will standards lead to greater administrative burden?
- Risk Mitigation - Is there a risk that the required data gathering will lead to discoverability in litigation?
- Competencies of Diversity Lead Professional – Does the individual understand how D&I leads to better organizational outcomes as well as change management requirements?

---

<sup>238</sup> <http://hrstandards.com.au/photo-gallery>

- Types of Metrics - What are the most important metrics to track performance? \
- Scorecards - Where should scorecards and results be tracked?

Many companies struggle with effectively measuring the results of diversity initiatives. In part, the challenge is determining what measures will yield useful information. For others, this task is difficult because they do not yet collect the necessary data required to measure diversity.

Diversity programs, for example, are often considered to have “intangible” results, such as improved communication or improved teamwork...yet such improvements may have a significant impact on productivity, growth and profits. Clearly, metrics is the path forward for organizations to successfully track results, as well as identify diversity management concerns that need to be addressed.<sup>239</sup>

Diversity and inclusion initiatives are not ends to themselves. They can help organizations and businesses become more profitable and competitive. However, without consistent standards and metrics, diversity and business leaders struggle with their diversity programs. Many rely solely on workforce demographics that point out gaps in the number of women and minorities in the workplace. They often do not assess turnover patterns, employee engagement trends, and productivity gains (or declines). Studies and a diversity research has affirmed that when managed properly, diverse teams outperform less diverse one especially in problem solving, innovation and market development.

---

<sup>239</sup> Brenman, Marc, Diversity Metrics, Measurement and Evaluation, Workforce Diversity Network,  
[http://www.workforcediversitynetwork.com/res\\_articles\\_diversitymetricsmeasurementevaluation.aspx](http://www.workforcediversitynetwork.com/res_articles_diversitymetricsmeasurementevaluation.aspx)

## **History of SHRM's Standards Development Work**

At the beginning of the D&I Standards effort in which I was selected to help lead, SHRM was the only organization licensed to create ANSI-certified standards in HR and Diversity. The goal of the proposed standards was to define minimally effective diversity and inclusion practices for organizations across the country. The Diversity and Inclusion Standards Project was initiated in 2010 with a taskforce of over 150 participants from private industry, educational institutions, consulting thought leaders, and non-profit organizations. In late 2014, SHRM shifted its emphasis and elected to discontinue its role in developing national human resource standards. In late 2015, it decided to transition the oversight of its standards development work to the American Society for Testing and Measurement (ASTM). The following is a high-level summary of the taskforce's work prior to the transition to ASTM.

## **Brief Overview of SHRM's D&I Standards Work**

The initial strategy was grounded in a belief that a minimal set of national standards needed to be established to help enable the advancement of organizational work and capability in this area.

The belief was based on the idea that standards in D&I should be promoted to establish a baseline of professionalism and that standards should be established using nationally accepted development protocols such as those endorsed and required by ANSI (American National Standards Institute). ANSI is well known for establishing standards in many other areas in business in the United States. Implicitly, a longer-term goal aspiration for some taskforce members was to also support the development of global standards through the International Standards Organization (ISO).

At the beginning of the SHRM Diversity and Inclusion Standards effort, Cari Dominguez, former head of the EEOC and I were tapped to serve as co-chairs. Over 200 diversity practitioners, consultants, and educators volunteered to help shape the standards. The work of the D&I Standards Taskforce was divided to focus on three principal areas:

1. Top Diversity Professional: A proposed minimum effective knowledge, skills and awareness of an organization's top diversity and/or inclusion professional that leads an effective diversity and inclusion program
2. Diversity Programs: A set of minimum effective features of a diversity and/or inclusion program including core elements (essential and optional), minimum standards for each element, and best practices
3. Measures and Metrics: A proposed minimum effective panel (collection) of diversity and/or inclusion metrics and measures that are periodically gathered in an effective diversity and inclusion program

Task force leaders were selected for each of the three areas and volunteers were assigned to each of these three sub-team areas. Drawing from literature, best practices, and consulting expertise, each sub-team recommended standards for their respective areas. The Taskforce was successful in developing draft recommendations in the Top Diversity Professional and Diversity Program areas and shared the recommendations with SHRM. Work was halted when SHRM decided to transition its human resource standards work to the American Society for Testing and Measurement. As of this writing, taskforce members are awaiting direction from ASTM and SHRM on when and how to proceed with the work that has been done by the taskforce.

## **National and International Standard Developing Organizations**

Since the beginning of the SHRM effort in human resources standards development, several organizations have stepped up their interest and involvement in the development of human resource standards.

American National Standards Institute (ANSI) oversees the creation, promulgation and use of thousands of norms and guidelines that directly impact businesses in nearly every sector, including but not limited to acoustical devices, construction equipment dairy and livestock production, energy distribution. ANSI is also actively engaged in accrediting programs that assess conformance to standards – including globally-recognized cross-sector programs such as the ISO 9000 (quality) and ISO 14000 (environmental) management systems. Thanks to SHRM’s initial move into standards, ANSI has begun publishing standards in the human resource disciplines.<sup>240</sup>

### **ANSI believes that standards “build trust.”**

Standards Build Trust. Reliability and trust are fundamental components of any process, business, or service. Behind the scenes, standards and conformity assessment ensure this reliability and trust. In short, standards make everyday life work. Everything from tech gadgets and the products we rely on in our offices and homes, to services that fuel the global economy and ensure health and safety, relies upon standards and conformance to ensure safety,

---

<sup>240</sup> [www.ansi.org](http://www.ansi.org); <https://www.nist.gov/news-events/news/2016/09/world-standards-week-2016>

dependability, and interoperability. A product or service conforming to an international standard represents a trusted symbol of quality, safety, and compatibility. Most of all, standards—developed through a process that is balanced, open, and transparent—help engender trust among the people, businesses, and governments that have placed their confidence in the standardization process.

The American Society for Testing and Measurement (ASTM International) is one of the largest voluntary standards developing organizations in the world. ASTM International is a not-for-profit organization that provides a forum for the development and publication of international voluntary consensus standards for materials, products, systems and services. The organization's volunteer members represent producers, users, consumers, government, and academia from more than 140 countries and they develop technical documents that are the basis of manufacturing, management, procurement, codes and regulations for dozens of industry sectors.

ASTM volunteer members belong to one or more standards-writing committees, each of which covers a subject area such as steel, petroleum, medical devices, consumer products, nanotechnology and additive manufacturing.

Since late 2014, SHRM has been in discussions with ASTM to become its preferred agent for the development of human resource standards. SHRM has suspended its involvement in standards development. Once SHRM finalizes the hand off of HR standards work to ASTM, members of the SHRM diversity taskforce hope that the work can be reinitiated and the standards development process completed.



International Organization for Standardization (ISO) is an independent, non-governmental membership organization and the world's largest developer of voluntary standards.

ISO is made up of representatives from 162 member countries who are the national standards bodies around the world, with a Central Secretariat that is based in Geneva, Switzerland.

The Technical Advisory Group (TAG 260) is providing direction on human resource standards and practices. In a recent development, ISO has established a new working group to develop global standards in diversity and inclusion, which I will lead as convener and Lorelei Carobolante will serve as co-convener and project leader.

There is growing interest in the development of diversity and inclusion standards around the world. In a recent discussion with experts from many different countries, they shared their views on the benefits of minimum standards in the diversity and inclusion practice. They shared the following benefits:

- Minimum standards reflect the customer base and therefore have positive effects on customer relations
- Diversity and inclusion helps organizations become more agile and adaptive to disruptive change
- Diverse workforce helps foster better ideas to strengthen positive business results (decision-making)
- Creativity and innovation (collaborative synergy) is enhanced when more cognitive diversity is present.

- Employer branding is improved when diverse stakeholders perceive the company as inclusive
- Employee retention is improved when all employees feel valued, appreciated and their differences respected.
- Inclusive institutional systems and practices increase the likelihood that all employees will be productive and their potential realized.
- Community citizens and activists expect organizations to reflect their communities. When the demographics within organizations reflects the demographics of the community and customers it serves, better guidance on how to reduce tension and violence in the community is achieved.
- Diverse talent prefer work for companies and organizations that are more respectful and inclusive.

Brian Levine, PhD, a consultant with Mercer Consulting, shared his thoughts in a presentation entitled, “Diversity Metrics to Tell a Story.” While Brian shares a concern that there are some problems with standard metrics, he believes a focus on “descriptive and predictive metrics that go beyond representation are important.<sup>241</sup> When referring to metrics he cites foundational metrics such as dashboards which monitor count rates (turnover, hiring, promotions, spans of control, etc.) and strategic metrics which take a deeper dive (proven inferences about cause-and-effect relationships such as “why do people quit? Do lateral moves make employees more promotable? What is the impact of span on employee engagement? These can be useful in

---

<sup>241</sup> Levine, Brian, Diversity Metrics to Tell a Story, Mercer Consulting, 2016 Conference of the Research Group on Legal Diversity, American Bar Foundation, Northwestern University, Chicago, Illinois, May 6, 2016.

strategy making, forecasting/leading indicators and problem solving. Dr. Levine believes that metrics that should be tracked and ultimately reported should vary by organization but that there are certain absolute or generic measures that most organizations should track and report.

However, effective diversity and human capital management is more about fit than universal truths, benchmarks or so-called “best practices. The story should drive the metrics, and is, itself, more important than the metrics.”

### **Growing Linkage to Sustainability, Human Rights and Global Diversity**

Several organizations including the United Nations, Reuters, the United States Government, and a growing number of foreign governments are incorporating diversity and inclusion standards and metrics in their assessment of workforce and workplace practices. For instance, Reuters recently introduced a new Diversity Index aimed at evaluating diversity and inclusion performance in major organizations. A quote from Reuters’s website explains the organization’s index:

Designed on the hypothesis that companies tracking, reporting and achieving on measures of diversity, inclusion and people development will offer better performance over time than those achieving lower scores, or not tracking these measures, Reuter’s D&I Index contains the information to help investors identify long-term opportunities and risks their investments.<sup>242</sup>

---

<sup>242</sup> <http://financial.thomsonreuters.com/en/products/data-analytics/market-data/indices/diversity-index.html>

Additionally, the United Nations 2030 Sustainable Development Agenda includes standards relating to gender equality, inclusive work practices and equality. Quoting from the Agenda, “the 2030 Agenda seeks to realize the human rights of all and to achieve gender equality and the empowerment of all women and girls. They are integrated and indivisible and balance the three dimensions of sustainable development: the economic, social and environmental.”<sup>243</sup>

### **My Work as A Practitioner – The Hendex Analytics Model – Foundational Standards<sup>244</sup>**

In my work, as both a human resources professional and chief diversity officer for a major international company, there are some standard categories that I have found are important in assessing diversity, workforce representation, and culture change. I have labeled this model the “Hendex Analytical Model.” My findings were drawn from research about how major regulatory agencies such as the U.S. Department of Labor’s Office of Federal Contract Compliance Programs and the Equal Employment Opportunity Commission analyze the practices of companies and organizations to determine if they are making “good faith” efforts to achieve workforce diversity and inclusion. The DOL is requiring more concrete proof of an organization’s “action oriented” efforts.<sup>245</sup> Analytical data, both quantitative and qualitative, are important in reviewing progress. I recommend, at a minimum, that the following areas be used as a part of an internal assessment process:

---

<sup>243</sup> <https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>. Retrieved on \_\_\_\_\_

<sup>244</sup> Henderson, Effenus (2016). The Hendex Analytical Framework for Diversity and Inclusion

<sup>245</sup> [http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&sid=3b71cb5b215c393fe910604d33c9fed1&rgn=div5&view=text&mp;node=41:1.2.3.1.2&idno=41#se41.1.60\\_62\\_117](http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&sid=3b71cb5b215c393fe910604d33c9fed1&rgn=div5&view=text&mp;node=41:1.2.3.1.2&idno=41#se41.1.60_62_117)

1. **Workforce Utilization.** Assess “Placements against opportunities” in jobs and categories that are significantly underrepresented with respect to women or minorities. When an opportunity to recruit talent for categories where underrepresentation exists, target selection and recruitment strategies and practices to increase the pool of diverse talent. If hiring rates are below suggested levels, it will take longer to close the gaps in workforce representation. This will mean more aggressive strategies will need to be put in place to close gaps.
2. **Workforce Trends.** Monitor year-over-year workforce change in the same categories. Are the numbers in those same categories remaining the same or improving? If not, it may be an indicator of higher turnover for the women and/ or minorities in those roles, suggesting the need for some intervention strategy. The problem may be related to the work climate, onboarding processes or retention factors (such as a poor supervisor or leader). Examine employee survey data and ensure that data about diversity and inclusion is a part of the assessment tool. This may mean incorporating specific questions relating to the workplace culture and climate for diversity. It can also mean asking for demographic data from respondents so that data can be disaggregated by identity groups (women, minority groups, age, tenure, etc.). This allows the organizations to identify differences in satisfaction and potential problem areas by identity group and allow for more targeted interventions.
3. **Diversity Index.** Analyze the index of the overall workforce diversity in the salaried ranks as compared to those in the production or hourly ranks. Are the percentages proportional for each of the underrepresented groups in both the hourly and salaried ranks? If a minority group is well represented in the hourly ranks but not at professional

and managerial levels, it might be a symptom of institutional bias. Bias can occur when hiring, promotion, development and other practices create a disparate impact on certain groups. This might result in practices in which minorities and women are disproportionately located in staff roles, hourly jobs or in areas where progression to higher level positions are more difficult.

4. **Diversity of Diversity.** Evaluate the diversity of diversity to avoid the “favored minority” risk. If a group is only represented in staff or hourly positions but not in the higher paid and salaried ranks, while another minority group is represented in those positions, it may be an indication of rising dissatisfaction with the less favored minority group. Many high-tech firms have higher percentages of Asian, Indian and immigrant employees in their technical and professional positions as well as the workforce overall. Many point to these statistics as proof that they have diverse workforces. However, the two largest minority groups in the US – African Americans and Hispanics may be “missing in action” and not well represented in these positions in relation to availability. The reason many advocates such as the Reverend Jessie Jackson and Reverend Al Sharpton and others are focusing on the industry is that utilization numbers for African Americans and Hispanic workers are demonstrably below expected levels. This is further exacerbated by the fact that many of the firms are in urban centers where these groups are sizable in numbers and growing. In addition, these demographic groups are a growing percentage of their customer base.
5. **Rate of Rise.** Examine upward mobility of underrepresented women and minorities into higher-level management and leadership roles (jobs in the top three levels of management). Look at the composition of the top level of management year over year,

how has the diversity at these levels changed? Has it improved? Are both women and people of color well represented? How are institutional processes and practices administered to ensure that women and minorities are moving up the corporate ladder? How are underrepresented employees reflected on task forces, in succession planning reviews, and management development programs and processes?

6. **Resource Allocation.** “Show me the money!” Assess spending levels and resources (FTEs) devoted to creating and sustaining relations with important organizations and associations that have significant pools of this diverse talent that could be tapped. Relationships are important in building trust and the perception that the organization is a “great place to work.”
7. **Evidence of Leadership Commitment and Communications.** Encourage explicit engagement and demonstrated commitment of the CEO and Board of Directors (by setting the example in its composition) and by its requirement for annual reviews of the organizational diversity and inclusion strategies, including as an element of its people review processes but as importantly in its business planning processes. Share and communicate leadership expectations.
8. **Diversity of Candidate Pools.** Require diverse candidate slates and diverse selection teams (sometimes referred to as the “Rooney Rule”)<sup>246</sup> in filling positions where there is substantial underutilization of women and minorities. If there is not diversity in the candidate slate, send the hiring manager back to the drawing board.

---

<sup>246</sup> Douglas C. Proxmire, Coaching Diversity: The Rooney Rule, Its Application and Ideas for Expansion, The American Constitution Society for Law and Policy, December 2008

9. **Bias in Institutional System, Policies and Practices.** Periodically assess institutional systems, policies and practices to insure they are inclusive (such as recruitment, succession, development, pay equity, selection, etc.). Voluntarily share workforce composition data internally and externally on an annual basis. Monitor pay equity issues and address intentionally when gaps are found.
10. **Education and Training.** Require participation in mandatory training every two years on topics such as respectful behavior, everyday bias, harassment prevention, and inclusive leadership behavior by all employees including targeted training for leaders and supervisors. Require that senior leaders model participation requirements, thereby setting the tone and expectation.
11. **Diversity Councils and Employee Resource Groups.** Support the establishment and sponsorship of employee resource groups (ERGs) whose charter is to support building a more diverse and inclusive culture. ERGs are very important in providing feedback on culture and climate and they serve as valuable forums for employee development and growth, volunteerism, and brand advocacy.
12. **Employee Issue Resolution Process.** Provide a rigorous internal issue resolution process as a means for employees to raise issues and concerns about the work place. Make it easy for employees to raise issues, present innovative ideas and challenge inappropriate behavior. Ensure that retaliation by supervisors and employees is not tolerated.



## Conclusion

In my work, as both a human resources professional and chief diversity officer for a major international company, I am convinced that minimum standards can help to build more sustainable and effective systems, practices, behaviors and policies to support a fairer and more equitable and inclusive workplace. I offer the following principles to consider:

Diversity and Inclusion standards help to:<sup>247</sup>

Reinforce compliance with legal and regulatory requirements. Being compliant with relevant legislation is the absolute minimum standard that organizations must achieve. The Standards reinforce compliance but also provide a rationale as to why organizations should do more than the minimum.

Support active and positive approaches to working with diverse employees. Rather than being reactive to continuing problems, standards provide managers with information to establish processes and practices that encourage and support inclusive workplaces.

Contribute to a fair and equitable work environment. Diversity and inclusion standards help in examining, developing and sustaining institutional processes that are consistent and free of bias, so that expectations in performance and relationships essential to an effective workplace are established and promoted.

Align D&I strategies with organizational vision, mission and values that support organizational excellence in social responsibility, governance and accountability. Diversity and inclusion strategies and practices do not function in isolation, but are part of the organization's overall

---

<sup>247</sup> Adapted from information: <http://www.hrcouncil.ca/resource-centre/hr-standards/about.cfm>

approach to interacting with its key stakeholders (employees, customers, suppliers, investors, regulators, and the communities in which it operates). Diversity and inclusion standards must be integrated with other standards, such as good governance, business conduct, social responsibility and financial accountability.

Foster individual learning, respectful and inclusive behavior that leads to organizational improvement. Diversity and inclusion standards should support organizations in identifying areas for improvement and to make a clear link between D&I and organizational results.

Provide tools that will build more inclusive systems and workplaces. By implementing minimum standards, organizations can make a reasonable commitment to D&I excellence and allocate resources to ensure capacity is built and sustained and in the areas of greatest organizational need.

## **Chapter 11: Diversity & Inclusion - A Service Academy Approach to Assessments / *Patricia L. Williams, Ph.D., PHR, Chief Diversity Officer at the U.S. Naval Academy***

### **Introduction**

The United States Naval Academy (USNA) is the second oldest of the five United States Military Service Academies, with an overarching goal that has remained the same: to meet the demand signal of the Navy and Marine Corps with approximately one third of the total officer accessions in the Naval Service each year.

The Naval Academy is a national treasure and unique institution of higher learning whose sole purpose is to develop ethical leaders of character and consequence who are prepared to lead Sailors and Marines, and into harm's way as necessary. This includes teaching the fundamental elements of leadership, and understanding myriad leader interactions, as well as Experiential Leader Development—where students engage in the practical application of leadership principles that are woven into daily, real-life examples and scenarios. As one of the world's finest leadership laboratories, the Naval Academy has produced Navy and Marine Corps leaders for 172 years. Its unique and far-reaching mission is:

“To develop Midshipmen morally, mentally, and physically and to imbue them with the highest ideals of duty, honor, and loyalty in order to graduate leaders who are dedicated to a career of naval service and have potential for future development in mind and character to assume the highest responsibilities of command, citizenship, and government.” (USNA Strategic Plan 2020).

The four-year undergraduate program is tuition free, and is for the Naval Service, graduating both Navy and Marine Corps officers. It is a federally funded institution, supported by a robust alumni association and foundation. Most students incur a five-year service obligation; however, some students incur additional service obligations after follow-on education and training. For example, those who attend follow-on training to become naval aviators incur additional obligations. Our students are considered active duty military members during their time at the Service Academy.

Students matriculate into the Naval Academy from every state and territory of the United States, and from foreign countries to learn and begin their careers of service to others. The students are very diverse with rich cultural heritage. In fact, in recent years, the academy has admitted the most diverse classes in the institution's history to include women, as well as racial and ethnic minorities. The Naval Academy offers 25 academic majors, which allows for the exploration of particular fields of study, including the more recent additions of Nuclear Engineering, Cyber Operations, and Operations Research.

We attribute much of the Naval Academy's high-volume application success to ongoing efforts to reach out to all of America so that all of America will be aware of this unique opportunity, and ideally be inspired to apply for admissions and matriculation into the U.S. Naval Academy. And of course, at no point do we sacrifice quality for quantity. It is truly a whole person admissions assessment.

The overall Brigade of Midshipmen includes a handful of International students, including foreign exchange students. At the Naval Academy, we seek to create an environment that fosters both ownership and membership within the overarching campus community; thereby, enabling the notion of an all-important sense of belonging, which results in greater retention and lower attrition rates among students. Some of our retention success pillars include:

- A strong sense of campus community
- Hands-on academic support with active monitoring, intervention, and resources
- Personal development, counseling, and mentoring
- Post-graduation counseling and career planning

We provide a cadre of qualified professionals to supervise, and we provide high-level, successful student support services and programs. Some of the various support services offered include:

- Chaplain Center
- Midshipmen Development Center (MDC)
- Brigade of Midshipmen Medical Care
- Center for Academic Excellence
- Personal Touch Academic Advising
- Sexual Assault Prevention and Response
- Physical Education Department/Physical Readiness Test Remediation
- Sports Injury and Physical Therapy Office
- Plebe (freshman)-Sponsor Program of local host families for midshipmen
- Character Development and Training
- Stockdale Ethics Center

## **The USNA Journey “Not College”**

Albeit a four-year bachelor degree granting institution, the Naval Academy is not the average college, here we refer to our students as midshipmen, and the student body as the Brigade of Midshipmen. From day one, begin the military service training in conjunction with the academic education.

“Not College”—Naval Academy students are on the cusp of important and impactful service around the world and within the continental United States in an increasingly complex and ever-changing environment. Yet, in many ways, Naval Academy students are indeed your regular college students with a few exceptions. For instance, we hire all of our graduates for a minimum five-year military service obligation where given the fluidity of the 21<sup>st</sup> century; a life of service presents recent graduates with many challenges like globalization, an increasingly high-tech, diverse workforce, continued proliferation of the Internet, and the volatility wrought by terrorism and other asymmetric threats.

Akin to their peers at other academic institutions, our students are concerned about their education and their future. They simply include the additional concern of a life of service and sacrifice in defense of freedom and democracy. Moreover, theirs’ is an academically rigorous 47-month course of study where all students, regardless of major, earn a Bachelor of Science degree because of the core engineering/technical curriculum. As with many like-minded institutions, ours is a robust Science, Technology, Engineering, and Mathematical (STEM) program, where we routinely graduate more than 65% of the students in STEM disciplines. The

STEM program is complete with wide-ranging STEM outreach to local and national communities designed to engage and influence students and teachers; including STEM Educator training, student workshops; robotics tournaments and Rubik's Cube Relay Teams; wind tunnels; operating underwater remotely operated vehicles, building bridges, designing underwater gliders and aluminum foil boats, the making of ice cream.

Our incoming student's journey begins each summer as Plebes after Induction Day. Plebe Summer is the summer training program required of all incoming freshmen (Plebes) to the United States Naval Academy. The program lasts for approximately six weeks and consists of rigorous physical and mental training. The purpose of Plebe Summer is to turn civilian students into midshipmen. It is positive pressure with a purpose. Albeit a military college environment, our sense of community mechanisms include things like Plebe Summer and Plebe Year with such key factors as the freshmen experience via an enhanced/extended freshman orientation; indoctrination and introduction to the Naval Academy mission; development of a new identity; and common experiences as first year students. Further, in terms of organizational/company structure, the student body is divided into 30 Companies of approximately 150 students each.

Beginning in the sophomore year, students begin taking on more and more leadership roles with a significant emphasis on leader development and teamwork. In terms of unique USNA identity, our students are in uniform, and often quickly adapt to the new midshipman identity. There is an underlying emphasis on responsibility to the institution that ultimately results in a life-long ethos of camaraderie.

Naval Academy campus community outcomes include that certain social support (Esprit de Corps), as well as peer and mentor support; individual accountability; and self-advocacy/self-empowerment, all of which fosters a greater sense of belonging and campus identity. There are feelings of loyalty, enthusiasm, and devotion to fellow students and to the profession of arms at large, or the capacity of a group's members to maintain belief in an institution or goal; particularly in the face of opposition or hardship that makes USNA different from the typical college experience.

When it comes to midshipman personal development and counseling, the more than 600 Naval Academy faculty and staff serve as gatekeepers in their ability to recognize distressed students and persuade them to use the helping resources available. Often, students undergo difficulties related to the transition and adjustment to USNA. In other cases, students may experience a less predictable crisis, which challenges their capacity for effectively coping with this college environment. Such crises include the death of family members or friends, dissolution of important relationships, or significant shifts in plans for their future. A dedicated team of experienced active-duty Navy Chaplains, Religious Program Specialists, civilian employees, and volunteers accomplish the daily work of delivering religious ministry at the Academy. Moreover, a wide variety of training, educational, and clinical services are provided to the midshipmen and staff to support the mission of the Naval Academy and to respond to the individual needs and goals of our diverse midshipman student population.

Life Skills for Leaders are educational courses on a variety of topics designed to help midshipmen navigate many of the life challenges they are likely to face at the Naval Academy



and as future leaders in the Navy and Marine Corps. Courses are open to all midshipmen, faculty, and staff, and are taught by a variety of instructors from the Midshipmen Development Center, the Chaplains, and the Academic Center for Excellence. Many faculty and staff also serve as hands-on mentors, offering a wide range of ongoing support.

The Naval Academy follows the Department of the Navy (DON) guidance in that a diverse and inclusive workforce has never been more important to the Department of the Navy's success. Given that a more diverse force makes us stronger, more effective, and more innovative, recruiting, retaining, and promoting top performers is a readiness imperative (Mabus, 2016).

### **Navy Alignment**

Our Alignment with the Department of the Navy's Diversity Policies & Programs is vital to our program's success. In fact, as members of the DON D&I Council, after more than two years' worth of diligent collaboration, we developed a now DON published DON D&I Strategic Roadmap, which serves as a guide for diversity and inclusion throughout the Department of the Navy, to include the U.S. Naval Academy. The D&I Roadmap focuses on three strategic imperatives:

- Strategic Imperative 1: Recruit and access from a diverse group of applicants to secure a high-performing and innovative workforce that reflects all segments of society.
- Strategic Imperative 2: Cultivate an inclusive culture that accelerates opportunities to empower each individual's maximum impact, encourages innovation and collaboration,

enhances developmental opportunities, and retains the best talent to enable uniformed and civilian personnel to contribute to their full potential.

- Strategic Imperative 3: Develop strategies to equip leaders with the ability to effectively manage diversity, be accountable, measure results, and refine approaches to engender a sustainable culture of inclusion. (DON D&I Roadmap, 2017)

### **Diversity and Inclusive Excellence**

Inherent in the Navy and Naval Academy core values of Honor, Courage, and Commitment is the fact that the Academy embraces diversity, and cultivates an engaged, inclusive, and innovative work environment, primarily led by our Office of Diversity, Inclusion, and Equal Opportunity (ODIEO). Within ODIEO, Diversity and Inclusion (D&I) is at the forefront of all that we do, and is one of the ways in which we incorporate diversity and inclusion into the mission of the Naval Academy. The Diversity Office supports faculty, staff, and students via D&I management across a broad range of efforts and initiatives.

In order to maintain our warfighting edge, it is essential that our people be diverse in experiences, backgrounds, and ideas; personally and professionally ready; and proficient in the operation of their weapons and systems. Rest assured, IT IS about Inclusive Excellence! D&I is a strategic imperative that is critical to mission readiness and success.

The business case for diversity has been made in both public and private sectors. Diverse experiences and differing thoughts lead to:

- Better decision-making and problem-solving capabilities

- Effective use of the workforce's talents
- Greater cultural competence to collaborate with colleagues and enhance mission performance
- More creativity and innovation
- Promoting Navy as an employer of choice in an increasingly diverse nation

The mission of the Naval Academy Office of Diversity, Inclusion, & Equal Opportunity is:

“To support, foster, and leverage the unique and diverse talents of faculty, staff, and future Navy and Marine Corps officers through an inclusive Naval Academy campus and community environment free from discrimination or harassment of any kind” (ODIEO).

The Naval Academy ODIEO was established in 2008. And, of the five U.S. Military Service Academies, the Naval Academy Chief Diversity Officer is the only active duty / military officer position. Fellow Service Academy counterparts are civilian Chief Diversity Officers, which does provide a certain measure of continuity in this key and essential senior leadership position. More recently, the ODIEO increased from two military to three military and two civilian personnel with the realignment of the military Equal Opportunity and Command Climate Specialist, a fellow military Enlisted position; along with two civilians, who are responsible for managing the Naval Academy Equal Employment Opportunity program for civilian employees. While our Equal Opportunity Specialists necessarily focus on all things compliance to include campus-wide climate surveys, they also provide insights and support within the D&I realm and body of work.

In an increasingly interdependent world, as we meet 21<sup>st</sup> global warfare demands, we must understand and appreciate the differences and similarities between people and leaders within the workplace.

We must create a climate of sufficient trust, collaboration, and coordination so we may capitalize on those things we have in common, while addressing our natural, unconscious biases. One might ask: what exactly does this mean? In essence, at its most basic core is this notion of basic dignity and respect for all, the effective use of personnel talent, reduced conflict between individuals, enhanced work relationships; shared organizational vision, greater innovation and flexibility; improved collaboration and productivity, fewer microinequities and/or microaggressions, promote and sustain a culture of inclusive excellence, and more conscious learners, leaders, and critical thinkers.

Within the Diversity Office, we manage diversity and foster inclusive excellence amongst organizational stakeholders where all individuals are valued, engaged, treated fairly and respectfully, and have equitable access to opportunities and resources. This includes providing key support of the various Departments across the Academy to include the Offices of Academics, Admissions, Athletics, and Assessments, as well as the Dean of Students.

We are steered by our Guiding Principle that the Naval Academy is a diversity friendly, healthy, and productive educational institution. We achieve this in large part with active and engaged support of Brigade Clubs and Organizations. For example, the Office of Diversity, Inclusion, and Equal Opportunity, in coordination with midshipmen Extra-Curricular Activities (ECA)

organizations and affinity groups, conducts in-reach to the Brigade of Midshipmen, and hosts as well as sponsors myriad educational, informational, and social events, to include cultural heritage celebrations throughout the year. The wide range of clubs available to students, depending on their interests and talents are listed here:

<b><u>Cultural</u></b>
Midshipman Black Studies Club
Filipino – American Club
French Club
German Club
International Club
Korean Midshipmen Club
Latin American Studies
Spectrum (LGBT) Club

<b><u>Music</u></b>
Church Choirs
Drum and Bugle Corps
Glee Club
Gospel Choir
Masqueraders (theater)
Pipes and Drums
String Ensemble

<b><u>Service</u></b>
Lucky Bag (Yearbook)
Midshipmen Action Group
Mids for Kids
Nat. Eagle Scout Association
Trident Newspaper
Tutoring

Every year, ODIEO sponsors the Academy’s Heritage Months celebrations. There are six history and heritage months celebrated at USNA: Black History Month in February, Women’s History Month in March, Asian American Pacific Islander Heritage Month celebrated in April, LGBT Pride month in June, Hispanic Heritage Month from September – October, and Native American Heritage Month in November. Students, staff, faculty, friends and family have excellent opportunities to come together to render appropriate recognition in honor of our rich heritage and tradition. In addition to Naval Academy efforts, we actively support our sister service academies via visits geared toward shared best practices and benchmarking. In addition, we also support fellow service academy students; for example, in November 2016, we escorted and housed in our dormitory a group of U.S. Coast Guard Academy (USCGA) Cadets and their culture and heritage club: USCGA Genesis Council on a visit to the Smithsonian’s new National Museum of African American History and Culture.

The Office of Diversity, Inclusion, and Equal Opportunity also supports midshipman attendance at myriad conferences and outreach events where our students mentor, are mentored, as well as engage in critical personal and professional development. Across the Naval Academy Campus, ODIEO serves as a resource for collaboration, mentoring and campus-wide resource for faculty, staff, and students. And, not only does this facilitate a greater sense of belonging, the USNA Diversity Office's efforts positively influence admissions applications, admissions outreach, retention, and inclusive excellence.

### **Assessments**

As the Chair of the Equity Study and Assessment Committee and the Proportional Outcomes study, we have worked closely with the Academy Effectiveness Board (AEB) that serves as the Naval Academy's Institutional Assessments arm, and is responsible for developing and maintaining an effective process for conducting ongoing assessments, which monitors and reports on the institution's overall effectiveness in fulfilling its mission. This institutional-wide focus requires the AEB to assess all three mission areas (moral, mental, and physical), as well as mission-supported functions related to the four-year leadership immersion program.

The AEB's Institutional Effectiveness Assessment Report for AY2015-16, as reviewed and approved by the Superintendent serves as an overview assessment report on institutional effectiveness. A number of key leaders in this capacity are responsible for developing and maintaining momentum in revising and moving forward with the Naval Academy assessment process focused on improving the institution's overall effectiveness.

More specially, as leader of the ESAC sub-committee, chartered to assess the equity of access and achievement of various midshipmen demographic sub-groups from multiple perspectives, we worked across campus Departments to ultimately recommend policies and procedures to address any identified disparities in outcomes. ESAC members include representation from key stakeholders across the campus. In such efforts, it is instructional to participate in benchmarking, and share best practices from sister academy's Chief Diversity Officers and other D&I Practitioners across industries. The Institute for Federal Leadership in Diversity and Inclusion was quite instrumental as a resource in the mastery of the D&I core concepts of strategic leadership competency; strategic communication competency; business acumen competency; building coalitions competency; change management competency; and D&I measurement competency. The completion of classes for the six competencies resulted in a Certificate of Mastery in D&I within the federal government for both the Chief Diversity Officer and the Assistant Chief Diversity Officer, who given the benefits of the mastery of such skills and competencies, can continue to foster inclusive excellence and garner greater support and understanding of Diversity and Inclusion concepts and best practices across the campus.

In order to address the question: at the end of the day, how do you know you've been successful, we focus on Institutional Effectiveness via the Academy Effectiveness Board and its Proportional Outcomes study. In the role of ESAC Chairman, and assessment efforts, in addition to the aforementioned, we also collect and analyze data to identify causal factors for attrition rates among women and minorities, to assess parity with majority groups. Further, we conduct disaggregated data analysis for each race and ethnic category to better understand attrition patterns, and to make recommendation, as necessary to college leadership.

With acceptable individual assessments currently in place, along with regular AEB Institutional Assessment, our efforts directly relate to the USNA Strategic Plan 2020's Strategic Imperatives:

1. Recruit, admit, and graduate a talented and diverse Brigade of Midshipmen.
2. Graduate officers whose attributes, educational and experiential preparation meet Navy and Marine Corps' current and future requirements.
3. Attract, develop & retain faculty, staff & coaches exemplifying the highest standards.
4. Align Midshipmen moral, mental, and physical experiences to prepare them for future service.
5. Integrate ethical leadership & character development efforts across all programs.
6. Leverage internal/external collaboration to engage Midshipmen in relevant learning opportunities that develop a broad range of competencies,
7. Establish and maintain state-of-the-art facilities.
8. Apply exemplary business and assessment practices.
9. Develop strategic relationships with alumni, friends, and national institutions.
10. Maintain institutional flexibility & achieve a margin of excellence with the Alumni Foundation (USNA Strategic Plan 2020)

As with any institution of higher learning, we continue institutional assessments based on the Academy's Strategic Objectives and Imperatives, and in concert with the overarching, robust Strategic Imperatives published by the Department of the Navy. The idea of self-assessment/measurement is to help identify and correct, if necessary, any trends in diversity management; thereby fostering a greater sense of inclusivity.



Federal data show the USNA student-faculty ratio is 7 to 1 at West Point, 8 to 1 at Colorado Springs and 9 to 1 at Annapolis. In terms of school rankings, the U.S. News & World Report's Best Colleges Rankings routinely ranks the U.S. Naval Academy as the number one high school counselor rankings; the number five undergraduate engineering program; and the number nine best liberal Arts College in America. The Class of 2016 graduated with an 89.5% four-year graduation rate, and some notable

Naval Academy Graduates include:

1 President of the United States	3 Cabinet Members
19 Ambassadors	24 Members of Congress
5 State Governors	5 Secretaries of the Navy
1 Secretary of the Air Force	5 Chairmen of the Joint Chiefs of Staff
29 Chiefs of Naval Operations	4 Vice Chairmen of the Joint Chiefs of Staff
9 Commandants of the Marine Corps	2 Nobel Prize Winners
73 Medal of Honor Recipients	53 Astronauts
49 Rhodes Scholars	28 Marshall Scholars
122 Olmsted Scholars	37 Fitzgerald Scholars

Additionally, demographic data alludes to the ongoing success of Naval Academy outreach efforts. For instance, the classes of 2020 and 2019 were the most diverse in the school's 172-year history. The student body included students from every state in the Nation, Guam, Puerto Rico, the Virgin Islands, the District of Columbia, as well as a number of international students from

countries like Bulgaria, Cameroon, Honduras, Mexico, Peru, Singapore, et al. The student body includes a 34% minority population, to include 25% females.

## **Outreach**

Some of the most rewarding work accomplished by the Office of Diversity, Inclusion, and Equal Opportunity is the Outreach work to our nation's youth, as well as community influencers in business and in education. The Naval Academy Diversity Office works in conjunction with the Office of Admissions to conduct numerous campus tours and events throughout the year, and especially over the summer, which also includes a USNA Admissions brief. We are committed to reaching out to all of America so everyone is aware of the unique opportunity that is the U.S. Naval Academy. There are two things about which I am most passionate: education and opportunity. And, we have a unique opportunity to provide educational opportunities to our nation's youth, who are our future leaders.

In our outreach efforts, we work closely with the United States Naval Academy STEM Center and some of the top Faculty in the STEM field. They were excited to assist the Diversity Office with students from the School for Legal Studies in Brooklyn, NY for a day of STEM activities where middle and high school students participated in hands-on activities to include one of the many vehicles in the halls of the Rickover Hall lab decks. In addition, in 2015, over 30 Houston, TX high school students and their chaperone traveled to Annapolis for a tour of the Naval Academy, and were also treated to our world class Admissions brief. Their personalized tour included the engineering labs, supersonic wind tunnel, the wave tank within the hydromechanics laboratory, and a wide-range of other state of the art engineering facilities.

Interestingly, in the summer of 2016 we offered the first-ever Summer Heroes Youth Program (SHYP) Pilot program in support of at-risk, low-income, inner-city, under-served middle school students from the Baltimore, Maryland area. Run by the Naval Academy Diversity Office and the USNA STEM Center, the two-week day camp not only served to open the eyes of about 45 middle school youth to the amazing world of STEM, camp experiences also served to promote positive leadership attributes in the 30 Midshipmen sophomore and juniors responsible for the middle school student's education and practical exercises. The three Baltimore area school teacher, chaperones, and principals reported much satisfaction from their and their student's participation. The Pilot program promoted positive development in urban-based, at-risk middle school students, between the ages of 10 and 14 years old. It also promoted positive leadership attributes in the midshipmen involved. The SHYP Pilot time frame was in the very beginning of the summer, and was a unique and valued professional development opportunity for the midshipmen involved who spent a month working this program, with the first two weeks focused on program development and leadership/STEM training, while the last two weeks was the actual camp and student interaction, learning, and engagement.

Yet another of our proud accomplishments was the 2016 year-long celebration of the 40<sup>th</sup> Anniversary of Women's admittance to the Naval Academy, and all of the nation's military service academies. The year-long celebration kicked-off on March 1<sup>st</sup> 2016 with the beginning of Women's History Month with one of the highlights on display at the Naval Academy Museum in the form of a USNA Museum Exhibit that opened on Induction-Day of 2016, and is entitled "Ability, Not Gender" and will be on display for approximately one year. In addition to a couple

of amazing advertisement larger than life banners displayed around the campus that reflect women leadership within the Brigade of Midshipmen, we opened an amazing U.S. Naval Institute History Conference in September 2016. The 2016 Naval History Conference at Alumni Hall at the U.S. Naval Academy was an ideal venue and vehicle in which to continue our 40<sup>th</sup> Anniversary of Naval Academy Women celebrations. The history conference (The Athena Conference: Heroines of Past, Present, and Future) was an excellent addition to our celebrations as we concluded the 40<sup>th</sup> Anniversary of women being admitted to the service academies with a 40-Year Anniversary Tailgate during a Navy home football game where women Naval Academy graduates from each class were honored on the field during the half-time celebrations.

Throughout 2016, attendees from all Naval Academy classes, and most especially the very first female Class of 1980, reported much healing and reconnection over the course of the year-long 40<sup>th</sup> anniversary celebration of women's admittance to the U.S. Naval Academy.

From the first African American Naval Academy graduate in 1949, to the first women Academy graduates in 1980, to the first female Naval Academy graduates to serve on Navy submarines in 2010, to embracing transgender member military service in 2016, the Naval Academy has broken barriers and shifted the campus climate towards greater inclusive excellence.

In summary, ongoing challenges facing ODIEO and the Academy include reaching out to all of America; continuing to raise awareness of unconscious biases; and fostering greater inclusive excellence. As lauded by Verna Myers (2012), it is not enough to be asked to the dance; the

greater inclusivity milestone is being asked to dance as opposed to remaining invisible and/or marginalized.

## **Conclusion**

After four years, our students, ethical leaders of character and consequence, are deemed ready to lead Sailors and Marines in the fleet, and are graduated and commissioned as ensigns in the Navy or second lieutenants in the Marine Corps. And, as evidenced in the Class Portraits for the Class of 2019 and the Class of 2020, the Naval Academy continues to excel in its Diversity and Inclusion efforts, and its overarching purpose of producing future military leaders.

Our efforts encompasses the core themes listed below, as well as the teaching and honing of the fundamental elements of leadership at arguably the world's greatest leadership laboratory: the United States Naval Academy.

- Naval Academy Core Themes:**
- ✓ **Develop Honorable Leaders**
  - ✓ **Exemplify 21st Century Sailor and Marine**
  - ✓ **Prepare for Careers of Service**

## References

Academic Year 2015-16 Institutional Effectiveness Assessment Report Retrieved September 27, 2016 from <https://www.usna.edu/StrategicPlan/institutional-assessment.php>.

Classes of 2019 & 2020 Naval Academy Class Portraits

DON D&I Strategic Roadmap. (2017). Retrieved January 30, 2017 from <http://www.21stcentury.navy.mil/Documents/DON-Diversity-and-Inclusion-Roadmap.pdf>.

Leaders to Serve the Nation, USNA Strategic Plan 2020 Retrieved May 30, 2014 from <https://www.usna.edu/StrategicPlan>

Myers V. A. (2012). Moving diversity forward: How to go from well-meaning to well-doing, Chicago: ABA Book Publishing.

ODIEO. (2017). D&I mission statement Retrieved April 6, 2017 from <http://www.usna.edu/Diversity>

Secretary of the Navy Statement. (19 August 2016). SECNAV Mabus Diversity & Inclusion Policy.

## **Appendix of Conference Materials**

Conference Schedule—including listings of presentations and panels

[http://www.americanbarfoundation.org/uploads/cms/documents/2016\\_conference\\_schedule.pdf](http://www.americanbarfoundation.org/uploads/cms/documents/2016_conference_schedule.pdf)

Summary of Conference Sessions and Discussion

[http://www.americanbarfoundation.org/uploads/cms/documents/abflegaldiversityconference2016\\_notes.pdf](http://www.americanbarfoundation.org/uploads/cms/documents/abflegaldiversityconference2016_notes.pdf)

Compiled Power Point Presentations from Conference

[http://www.americanbarfoundation.org/uploads/cms/documents/compiled\\_2016\\_slides.pdf](http://www.americanbarfoundation.org/uploads/cms/documents/compiled_2016_slides.pdf)

A transcript of proceedings is available on request. Please contact Robert L. Nelson, American Bar Foundation, at [rnelson@abfn.org](mailto:rnelson@abfn.org).