

# HEINONLINE

Citation: 47 Stan. L. Rev. 1161 1994-1995

Content downloaded/printed from  
HeinOnline (<http://heinonline.org>)  
Mon Dec 19 14:30:24 2011

- Your use of this HeinOnline PDF indicates your acceptance of HeinOnline's Terms and Conditions of the license agreement available at <http://heinonline.org/HOL/License>
- The search text of this PDF is generated from uncorrected OCR text.
- To obtain permission to use this article beyond the scope of your HeinOnline license, please use:

[https://www.copyright.com/ccc/basicSearch.do?  
&operation=go&searchType=0  
&lastSearch=simple&all=on&titleOrStdNo=0038-9765](https://www.copyright.com/ccc/basicSearch.do?&operation=go&searchType=0&lastSearch=simple&all=on&titleOrStdNo=0038-9765)

# The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity

Linda Hamilton Krieger\*

*Title VII's disparate treatment model of discrimination is premised on the notion that intergroup bias is motivational in origin. This premise, in turn, is based on a number of assumptions regarding the nature of human inference and the respective roles played by cognition and motivation in social judgment and decisionmaking. Applying insights from cognitive psychology, Professor Krieger examines the assumptions about human inference embedded in current disparate treatment theory and questions the premise that discrimination necessarily manifests intent or motive. She suggests that a large number of biased employment decisions result not from discriminatory motivation, as current legal models presume, but from a variety of unintentional categorization-related judgment errors characterizing normal human cognitive functioning. Because of the lack of fit between the present disparate treatment model and the phenomenon it purports to represent, courts and litigants are presented with a confusing array of increasingly ill-defined and questionably premised analytical paradigms. Worse, as currently constructed, it may be exacerbating intergroup tensions and inflating both social and financial adjudication costs. Searching for solutions, Professor Krieger explores the legal and policy implications of a cognitive process approach to discrimination and equal employment opportunity and evaluates a variety of modifications to existing equal employment opportunity law.*

## INTRODUCTION

A few years ago, I had one of those experiences that slips almost unnoticed into your consciousness and then quietly wreaks havoc on your tidy way of looking at something.

---

\* Senior Research Fellow, Stanford Center on Conflict and Negotiation. B.A., Stanford University, 1975; J.D., New York University, 1978. This article has profited from the patience, faith, and efforts of many people including colleagues here at the Law School and in the Department of Psychology, research assistants, members of the *Stanford Law Review*, and the staff of Stanford's Robert Crown Law Library, whose helpfulness and cheerful efficiency repeatedly blew my mind. But I am sure they will understand if, rather than thanking each of them here, I instead take the opportunity to dedicate this work to my late father, Dr. John A. Krieger, who never agreed with my politics and would doubtless quarrel with much of what I have written here, but who nonetheless has always been the tide that brings me home.

I was working on an unremarkable Title VII<sup>1</sup> case. My client was a young Salvadoran man who had been the only nonwhite employee at a box manufacturing plant in California's Central Valley. He had been denied a promotion, then fired, and he was convinced that he had been treated less favorably than his Caucasian coworkers because of his national origin. Their transgressions had been systematically overlooked or explained away; his had consistently led to oral and written reprimands that now served to justify his termination. They received commendation for their achievements; his seemed to go unnoticed, or were attributed to the efforts of others. And then there were the subtle things: the way people looked at him, their tone of voice—telltale signs of bias that, he told me, "I can't prove, but I just know."

This was not going to be an easy case. The employer had no facially discriminatory policies,<sup>2</sup> nor any identifiable neutral policies that I might argue disproportionately disadvantaged Latinos.<sup>3</sup> No one had made any derogatory ethnic comments, so far as I could determine.<sup>4</sup> But there was a subtle, yet

1. Title VII of the Civil Rights Act of 1964, as amended, 42 U.S.C. §§ 2000e to -16 (1994), prohibits discrimination in employment on the basis of race, color, sex, national origin, and religion.

2. "Facial discrimination," the most conspicuous form of discrimination, describes adverse employment decisions made explicitly on the basis of a person's membership in a protected class. Facial discrimination almost always violates Title VII. It may be justified in only two instances. The first is by proof that restricting employment to a particular class is a bona fide occupational qualification (BFOQ). See, e.g., *UAW v. Johnson Controls, Inc.*, 499 U.S. 187, 201 (1991) (restricting the BFOQ defense to objective, verifiable requirements concerning job-related skills and aptitudes). The second is by proof that the discriminatory policy is within the context of a valid affirmative action plan. See, e.g., *Johnson v. Transportation Agency*, 480 U.S. 616, 642 (1987) (stating that such a plan is "fully consistent with Title VII, for it embodies the contribution that voluntary employer action can make in eliminating the vestiges of discrimination in the workplace").

3. Because this case implicated no such practices, Title VII's disparate impact theory—which, unlike disparate treatment theory, does not require a showing of intent to discriminate—could not be used to establish liability. One common misperception about Title VII is that a plaintiff can prevail in virtually any type of case by making an un rebutted showing of disparate impact on a group protected by Title VII. This is not true. Disparate impact theory requires that the plaintiff first identify a specific employment practice. The plaintiff must then prove that this particular practice has a statistically significant disparate impact on a qualified group protected by Title VII. *Watson v. Fort Worth Bank & Trust*, 487 U.S. 977, 994 (1988). These requirements were not eliminated by the Civil Rights Act of 1991, the relevant portions of which are now codified at 42 U.S.C. § 2000e-2(k)(1)(A) (1994). Even if this showing is made, the employer may prevail by proving that the practice is "job related for the position in question and consistent with business necessity." 42 U.S.C. § 2000e-2(k)(1)(A) (1994).

Because most individual employment decisions do not implicate identifiable practices that can be shown to have a statistically significant disparate impact on members of a protected group, very few Title VII cases are actually amenable to disparate impact treatment. According to an American Bar Foundation study, disparate impact cases comprised only 1.84% of all employment-related civil rights cases in the federal court docket between 1985 and 1987. American Bar Foundation Employment Discrimination Litigation Survey, Computer File (1990), cited in John J. Donohue III & Peter Siegelman, *The Changing Nature of Employment Discrimination Litigation*, 43 STAN. L. REV. 983, 998 n.57 (1991).

4. The lack of derogatory comments barred the use of the relatively plaintiff-friendly burdens of proof articulated in the Supreme Court's decision in *Price Waterhouse v. Hopkins*, 490 U.S. 228, 244-45 (1989), to establish liability. Since *Price Waterhouse* was decided, many circuit courts have limited its applicability to situations in which a plaintiff can prove by "direct evidence" that his or her gender, race, or ethnicity was a motivating factor in a challenged employment decision. See, e.g., *Caban-Wheeler v. Elsea*, 904 F.2d 1549, 1555 (11th Cir.), cert. denied, 498 U.S. 848 (1990); *Jackson v. Harvard Univ.*, 900 F.2d 464, 467 (1st Cir. 1990); *Gagne v. Northwestern Nat'l Ins. Co.*, 881 F.2d 309, 315 (6th Cir. 1989).

For a discussion of the confused and limited scope of the *Price Waterhouse* "motivating factor" paradigm in disparate treatment litigation, see text accompanying notes 262-282 *infra*.

discernible pattern of differential treatment emerging from the time records and personnel files obtained in discovery.

If my client were to prevail in establishing a Title VII violation, it would have to be—as in well over 90 percent of all Title VII cases<sup>5</sup>—under the “disparate treatment” theory of discrimination first established in *McDonnell Douglas Corp. v. Green*.<sup>6</sup> Under this theory, he would have to prove not only that he received less favorable treatment than his Anglo coworkers, but that his superiors purposefully, deliberately, and intentionally treated him differently because of his national origin.<sup>7</sup> To be blunt, to establish that my client had been wronged, I would have to prove that the plant manager was a racist and a liar.<sup>8</sup>

As is usual in employment discrimination cases, the challenged manager adamantly denied that my client’s national origin played any part in his decisionmaking process. He claimed instead that my client had arrived late to work too many times, had violated too many safety rules. As for the promotion, even without these performance problems, he was just “too easy-going”—not the “take charge, don’t-mess-with-me kind of guy” that a foreman has to be.

Interviewing this manager had not been pleasant. He was angry and defensive and, as I questioned him about time records indicating that two Anglo employees had been late as often as my client, he got even angrier. Finally, ignoring his attorney’s admonitions “to answer only the question asked,” he exploded. “Look, I don’t appreciate being called a bigot. Mateo’s being a Mexican [sic] didn’t make any difference to me; it’s like I didn’t even notice it.”

Later that day, I came back from the Valley, changed my clothes, grabbed the newspaper, and sat down with my three-year-old son to watch *Sesame Street*. That’s when it happened. I looked up from my newspaper and saw the television screen divided into four sections. In each section was a child in a raincoat—three of them red, one yellow—and Big Bird was singing:

One of these kids is not like the others.

One of these kids just isn’t the same.

One of these kids is not like the others.

Now it’s time to play our game.<sup>9</sup>

The point of the game was, of course, to figure out that the child in the yellow raincoat was different than the three children in the red raincoats. The pedagogical purpose of the game was to teach children to categorize—to notice differences between objects and to group those objects, on the basis of those differences, into categories.

---

5. Donohue & Siegelman, *supra* note 3, at 1019 (citing American Bar Foundation Employment Discrimination Litigation Survey, *supra* note 3).

6. 411 U.S. 792 (1973).

7. See text accompanying notes 56-67 *infra* for a discussion of the *McDonnell Douglas* model.

8. Existing disparate treatment jurisprudence in many ways equates a finding of pretext in plaintiff’s favor with a finding that the employer has lied about the reasons for its decision. See *St. Mary’s Honor Ctr. v. Hicks*, 113 S. Ct. 2742, 2763, 2764 (1993) (Souter, J., dissenting).

9. *Sesame Street* (PBS television broadcast).

Children must learn to categorize. They must categorize to understand speech, to move safely through their environment ("Don't touch a strange dog; it might bite you. But yes, it's okay to touch a strange cat; cats don't bite." "Don't get into a stranger's car; but yes, you can get into Uncle Hurley's car; he's a relative.").

Children must learn to categorize before they can learn much of anything else. And when they get a little older, before they can learn to read, they have to learn to stereotype. You simply can't read if you can't stereotype. You have to minimize all those differences between the ways different people write an "F." Without even thinking about it, you have to exaggerate the subtle differences between a capital "D" and a capital "P." Your mind has to fill in when part of a line is missing, or ignore a stray mark that your eyes indeed see, but your mind knows does not really go with an "a."

It seemed ironic. There I sat, watching with maternal satisfaction as Big Bird taught my son to notice and categorize by color differences, while the plant manager's indignant claim of colorblindness echoed in my mind. It unsettled me as I sat there, and left me with a vague sense of disquiet.

In retrospect, I see that this experience was a turning point in my thinking about intergroup relations, discrimination, and equal employment opportunity. In the years that followed, I became increasingly uneasy about the enterprise in which I, as a Title VII lawyer for over a decade, had engaged. As I encountered more offended, defensive decisionmakers accused of discrimination, and as I counseled and consoled more embittered employees who knew they had been treated differently because of their race or gender or ethnicity but could not, as the law requires in such cases, prove that their employer harbored a discriminatory motive or intent, I became convinced that something about the way the law was defining and seeking to remedy disparate treatment discrimination was fundamentally flawed. This article represents my endeavor to understand what that "something" might be.

In this article, I argue that the way in which Title VII jurisprudence constructs discrimination, while sufficient to address the deliberate discrimination prevalent in an earlier age, is inadequate to address the subtle, often unconscious forms of bias that Title VII was also intended to remedy.<sup>10</sup> These subtle forms of bias, I suggest, represent today's most prevalent type of discrimination. While Title VII jurisprudence gives lip service to the notion that actionable intergroup bias can be subtle or unconscious, courts have so far failed to develop doctrinal models capable of addressing such phenomena—especially subtle or unconscious race and national origin discrimination.

This failure, I propose, stems from the assumption that disparate treatment discrimination, whether conscious or unconscious, is primarily motivational, rather than cognitive, in origin.<sup>11</sup> This one-sided understanding of bias leads

---

10. See, e.g., *McDonnell Douglas*, 411 U.S. at 801 (explaining that "Title VII tolerates no racial discrimination, subtle or otherwise").

11. It is here that my perspective differs from that so compellingly offered by Professor Charles R. Lawrence III in *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317 (1987). Drawing on psychoanalytic theory, Professor Lawrence argues that much of what

courts to approach every disparate treatment case as a search for discriminatory motive or intent. To the extent that intergroup bias stems from other sources, current models may either fail to identify discrimination or wrongfully attribute discriminatory motive to a well-intentioned, though biased, decisionmaker. We need a deeper, more nuanced understanding of what intergroup discrimination is, how and why it occurs, and what we can do to reduce it.

My endeavor is divided into four parts. Part I examines the analytical structure of Title VII's disparate treatment model of discrimination and the rhetoric courts employ in analyzing disparate treatment cases. I explore the unexamined central premise of Title VII's disparate treatment paradigm—that the origin of intergroup bias is motivational. In Part I, I also examine a number of assumptions regarding the nature of social judgment and decisionmaking reflected in Title VII's disparate treatment jurisprudence and trace those assumptions to personality and social psychological theories prevalent in the 1930s through the 1960s.

In Part II, I test these assumptions against insights derived from more contemporary empirical and theoretical research in cognitive and cognitive-social psychology. I conclude that, while the assumptions undergirding disparate treatment theory generally reflect the thinking about intergroup bias and human inference accepted into the 1970s, these assumptions have been so undermined, both empirically and theoretically, that they can no longer be considered valid.

Specifically, I suggest in Part II that a broad class of biased employment decisions now analyzed under Title VII's disparate treatment theory results not from discriminatory motivation, but from a variety of categorization-related judgment errors characterizing normal human cognitive functioning. Consequently, I propose, there now exists a substantial discontinuity between the jurisprudential construction of discrimination and the real life phenomenon it purports to represent.

In Part III, I examine the consequences of this discontinuity and the implications of a cognitive bias approach to discrimination for various aspects of equal employment opportunity law and policy. First, I suggest that the lack of fit between the disparate treatment model and the phenomenon it purports to represent has led to analytical incoherence in Title VII jurisprudence. This uncertainty and confusion, I argue, can be expected to decrease the validity of employment discrimination adjudications, increase adjudication costs, and discourage the voluntary settlement of employment discrimination cases. Furthermore, by attributing disparate treatment to discriminatory intent, current doctrine may exacerbate rather than reduce intergroup tensions.

Part IV searches for solutions. First, I argue that the pretext model of disparate treatment proof should be eliminated entirely and replaced by a "motivating factor" analysis similar to, but in some ways quite different from, that

---

is classified as disparate treatment discrimination results from subconscious instincts and motivations. While Professor Lawrence does mention cognitive bias as a potential source of discriminatory decision-making, he focuses primarily on discussing motivational rather than cognitive antecedents. *Id.* at 331-36.

adopted in *Price Waterhouse v. Hopkins*.<sup>12</sup> Second, I advocate grafting onto Title VII a 2-tier approach to liability and relief, similar to that used in cases under the Age Discrimination in Employment Act<sup>13</sup> under which “willful” and “nonwillful” discrimination are distinguished and accorded different remedies.

Finally, Part IV suggests that the nondiscrimination principle, currently interpreted as a proscriptive duty “not to discriminate,” must evolve to encompass a prescriptive duty of care to identify and control for category-based judgment errors and other forms of cognitive bias in intergroup settings. However, unlike other scholars who advocate a “negligence” approach to employment discrimination,<sup>14</sup> I suggest that additional empirical and theoretical work must be done before the contours of such a duty can be precisely defined, let alone crafted into practical and effective legal rules.

## I. THE CONSTRUCTION OF DISPARATE TREATMENT DISCRIMINATION IN TITLE VII JURISPRUDENCE

Cases tell stories. Indeed, judicial opinions and the legal theories they expound function somewhat like a society’s core stories, structuring the interpretation of experience and providing the authors and audiences of future stories with commonly recognized plots, symbols, themes, and characters.

In a sense, one’s task as a litigator is to choose a core story from existing jurisprudence and then to construct, from the available facts, a new story that resembles the core story as closely as possible. Thus, to bring one’s experience into the world of the law by filing a lawsuit is implicitly to consent to its being interpreted and judged through the lenses of the law’s core stories. To be “dragged into court” as a defendant is to find oneself and one’s experience cast in the story’s mold. Doctrinal models define for the litigants which facts belong in their stories and which do not. The assumptions upon which those models are constructed—the “meta-stories,” if you will—define how those facts will be interpreted in legal decisionmaking.

I knew that if I wanted to understand why so many of my clients and their opponents felt frustrated and misunderstood in their encounters with civil rights enforcement, I would have to examine the core stories told by Title VII disparate treatment caselaw. How do the stories define intergroup bias and how it operates? What are the stock plots? Who are the stock characters? How are characters’ actions interpreted, and what assumptions shape those interpretations? In this section, I attempt to answer these questions.

One sees in the stories that Title VII cases tell certain assumptions about human inference and judgment—both generally and in intergroup contexts. First, intergroup discrimination, even when subtle and unconscious, is assumed to result from discriminatory motive or intent. Thus, to ask whether an em-

---

12. 490 U.S. 228 (1989).

13. Age Discrimination in Employment Act, Pub. L. No. 90-202 § 1, 81 Stat. 602 (1967) (codified as amended at 29 U.S.C. § 621 (1994)).

14. See David Benjamin Oppenheimer, *Negligent Discrimination*, 141 U. PA. L. REV. 899, 915-17 (1993).

ployer discriminated against an individual because of group status is seen as equivalent to asking whether a discriminatory purpose motivated the employer's decision. Under this equation of causation and intentionality, even discrimination that results from applying racial, ethnic, or gender stereotypes is understood as a product of discriminatory motivation. Indeed, evidence that a decisionmaker holds stereotypes is relevant in disparate treatment analysis because it presumedly unmaskes discriminatory intent. In the stories told by disparate treatment caselaw, there is no discrimination without an invidiously motivated actor. Every successful disparate treatment story needs a villain.

Second, disparate treatment analysis assumes that, unless they harbor discriminatory intent or motive, decisionmakers will act objectively and judge rationally. Evidence that a particular decisionmaking strategy was suboptimal in light of all decision-relevant events is often construed as indicating that a discriminatory purpose really motivated the decision in question, and that the employer is now lying about his "true" intentions. In the process of fitting the facts into the doctrinal forms, evidence of judgment errors is interpreted as indicating that the nondiscriminatory reasons now given for an employment decision are "phony reasons"—"cover-ups" for the "real" discriminatory motivation. As a result of this "real reason/phony reason" distinction, the adjudication of most disparate treatment claims sinks inevitably into a thinly disguised brawl over whether the accused employer is lying about the reasons a particular decision was made.

Third, disparate treatment jurisprudence construes the process of employment decisionmaking as being functionally distinct from the processes of perception, encoding, and retention in memory of decision-relevant events. Title VII disparate treatment caselaw constructs intergroup discrimination as something occurring at the moment a decision is made. Current disparate treatment jurisprudence does not recognize that categorization based on race, sex, or national origin may distort perception, memory, and recall for decision-relevant events such that, at the moment of decision, an employer may be entirely unaware of the effect of an employee's group membership on the decisionmaking process.

Finally, disparate treatment jurisprudence—indeed the entire normative structure of anti-discrimination law—is based on an assumption that decisionmakers possess "transparency of mind," that they can accurately identify why they are about to make, or have already made, a particular decision. According to this view, if an employee's protected group status is playing a role in an employer's decisionmaking process, the employer will be aware of that role, even if he is not honest (or careless) enough to admit it. Equipped with conscious self-awareness, well-intentioned employers become capable of complying with the law's proscriptive injunction not to discriminate. They will monitor their decisionmaking processes and prevent prohibited factors from affecting their judgments.

Let us examine more closely how these assumptions undergird disparate treatment jurisprudence and shape the stories told by Title VII caselaw.

### A. *The Equation of Causation and Intentionality*

Section 703 of Title VII<sup>15</sup> prohibits employers from failing or refusing to hire, or from discharging or otherwise discriminating against any individual because of his or her race, color, religion, sex, or national origin. It would be reasonable to interpret this language as simply requiring proof of causation without proof of intent. In other words, a Title VII claimant would need only establish that his or her protected status "made a difference" or "played a role" in a challenged employment decision.

This is not, however, how section 703 has been construed. Under existing law, the disparate treatment plaintiff, whether proceeding under Title VII or under 42 U.S.C. sections 1981 or 1983, must prove not only that she was treated differently, but that such treatment was caused by purposeful or intentional discrimination.<sup>16</sup> Particularly in the context of race and national origin, discrimination is represented as resulting from the decisionmaker's discriminatory animus towards members of the plaintiff's racial or ethnic group.<sup>17</sup> This stands in marked contrast to the law's construction of age discrimination. Con-

15. Specifically, § 703(a) provides, in pertinent part:

(a) It shall be an unlawful employment practice for an employer—

(1) to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's race, color, religion, sex, or national origin.

42 U.S.C. § 2000e-2(a) (1994).

16. See, e.g., *St. Mary's Honor Ctr. v. Hicks*, 113 S. Ct. 2742, 2752 (1993) ("[The plaintiff has] the ultimate burden of persuading the court that she has been the victim of intentional discrimination.") (quoting *Texas Dep't of Community Affairs v. Burdine*, 450 U.S. 248, 256 (1981)). *Accord* *EEOC v. Flasher Co.*, 986 F.2d 1312, 1312 (10th Cir. 1992) (holding that plaintiff had to prove termination of employment was result of intentional discrimination based on plaintiff's national origin); *Warren v. Halstead Indus., Inc.*, 802 F.2d 746, 752-53 (4th Cir. 1986) (holding discriminatory intent means actual motive and cannot be presumed based upon a factual showing of less than actual motive); *Smith v. Honeywell, Inc.*, 735 F.2d 1067, 1068-69 (8th Cir.) (holding individual alleging disparate treatment has the burden of showing not only a difference in treatment, but that he is a victim of intentional discrimination), *cert. denied*, 469 U.S. 1077 (1984); *Smithers v. Bailar*, 629 F.2d 892, 898 (3rd Cir. 1980) (holding disparate treatment plaintiff is required to prove not only disparate treatment, but that such disparate treatment was caused by purposeful or intentional discrimination).

The plaintiff's burden in a Title VII disparate treatment case mirrors the showing required of an equal protection claimant proceeding under 42 U.S.C. §§ 1981 and 1983. See, e.g., *Patterson v. McLean Credit Union*, 491 U.S. 164, 186-87 (1989) (requiring both disparate treatment and intent); *Gutzwiller v. Fenik*, 860 F.2d 1317, 1325 (6th Cir. 1988) (noting that the required showings are similar under Title VII and § 1983). *Accord* *AFSCME v. State of Washington*, 770 F.2d 1401, 1405 (9th Cir. 1985) (applying the standard articulated in *Personnel Adm'r v. Feeney*, 442 U.S. 256, 279 (1979), that discriminatory purpose is more than awareness of consequences in a Title VII disparate treatment case).

Before the Supreme Court's decision in *Hicks*, the Second Circuit ruled that in disparate treatment cases, the plaintiff's group status need only have "made a difference" in the contested employment decision. See, e.g., *Ramseur v. Chase Manhattan Bank*, 865 F.2d 460, 465 (2d Cir. 1989); *Hagelthorn v. Kennecott Corp.*, 710 F.2d 76, 82 (2d Cir. 1983). No other circuit follows this approach, and its continued viability following *Hicks* is uncertain.

17. See *Gomez v. Medical College*, No. 92-5048, 1994 U.S. Dist. LEXIS 11274, \*9 (E.D. Pa. 1994) ("A plaintiff may not prevail on a mere showing that the defendant's proffered reasons are false, but must prove a discriminatory animus."); *EEOC v. Flasher Co.*, 986 F.2d 1312, 1321 (10th Cir. 1992) ("Merely finding that people have been treated differently stops short of the crucial question: *why* people have been treated differently."); *Minority Police Officers Ass'n v. City of South Bend*, 617 F. Supp. 1330, 1358 (N.D. Ind. 1985) ("Mere conclusory allegations of discrimination are clearly not sufficient to prove discriminatory intent.").

sider the following language, which originated in the Seventh Circuit and was subsequently adopted by many other courts:

Unlike race discrimination, age discrimination may simply arise from an *unconscious* application of stereotyped notions of ability rather than from a deliberate desire to remove older employees from the workforce: "Age discrimination is not the same as the insidious discrimination based on race or creed prejudices and bigotry. Those discriminations result in nonemployment because of feelings about a person entirely unrelated to his ability to do a job. This is hardly a problem for the older jobseeker. Discrimination arises for him because of assumptions that are made about the effects of age on performance."<sup>18</sup>

It is hard to understand why a court would assume that race discrimination could not, as easily as age discrimination, result from the unconscious application of stereotyped notions of ability or other characteristics. It is also difficult to understand why a court would assume that race discrimination results exclusively from a deliberate desire to exclude members of a particular racial group from the workforce.<sup>19</sup> Yet, in equating the causation requirement of section 703 with discriminatory intent, courts have constructed disparate treatment theory on these two dubious assumptions.

This is not to say that one cannot find language in disparate treatment cases acknowledging that race, sex, or national origin discrimination can be subtle or even unconscious. But one can literally count on one hand the number of published Title VII decisions in which, after acknowledging the existence of unconscious bias, the court rules for a race or national origin discrimination plaintiff or reverses a trial court ruling for the defendant.<sup>20</sup> Many more courts, after acknowledging the existence in society generally of subtle or unconscious

---

18. *Syvock v. Milwaukee Boiler Mfg. Co.*, 665 F.2d 149, 155 (7th Cir. 1981) (quoting 113 Cong. Rec. 34,742 (1967) (remarks of Rep. Burke)). *Accord* *Burlew v. Eaton Corp.*, 869 F.2d 1063, 1066 (7th Cir. 1989) (age discrimination may exist absent an intent to discriminate); *Brooks v. Woodline Motor Freight*, 852 F.2d 1061, 1064 (8th Cir. 1988) (citing *Syvock* for the proposition that unconscious stereotypes often underlie age discrimination); *Oxman v. WLS-TV*, 846 F.2d 448, 453 (7th Cir. 1988) (reaffirming *Syvock* regarding age discrimination); *La Montagne v. American Convenience Prods., Inc.*, 750 F.2d 1405, 1410 (7th Cir. 1984) ("Age discrimination may be subtle and even unconscious."); *McCorstin v. United States Steel Corp.*, 621 F.2d 749, 754 (5th Cir. 1980) (asserting that age discrimination "is more subtle but just as injurious to the worker").

19. *See* Oppenheimer, *supra* note 14, at 902; *see also* Lawrence, *supra* note 11, at 322.

20. *EEOC v. Inland Marine Indus.*, 729 F.2d 1229, 1236 (9th Cir.) (holding that racial discrimination occurs where subjective employment criteria embody racially discriminatory attitudes, even where intent is not established), *cert. denied sub nom.* *Inland Marine Indus. v. Houston*, 469 U.S. 855 (1984); *Johnson v. Stone*, 58 Fair Empl. Prac. Cas. (BNA) 656, 656 (D. Colo. 1992) (finding for plaintiff in race discrimination case where only evidence of discrimination was that plaintiff was not promoted and employer had nicknamed him "Bub").

One can find a somewhat larger number of sex discrimination cases resolved in plaintiff's favor on these grounds. *See, e.g.*, *Hopkins v. Price Waterhouse*, 825 F.2d 458, 469 (D.C. Cir. 1987) (asserting that unawareness of bias "neither alters the fact of its existence nor excuses it"), *aff'd sub nom.* *Price Waterhouse v. Hopkins*, 490 U.S. 228 (1989); *Lynn v. Regents of the Univ. of Cal.*, 656 F.2d 1337, 1343 (9th Cir. 1981) (asserting that "disdain for women's issues . . . is evidence of a discriminatory attitude towards women"); *Sweeney v. Board of Trustees of Keene State College*, 604 F.2d 106, 113 n.12 (1st Cir. 1979) (sex discrimination case affirming judgment for plaintiff because the district court reasonably concluded that the decision not to promote plaintiff was "determined by a subtle, if unexpressed, bias against women").

forms of bias, rule against the disparate treatment plaintiff on the grounds that she has failed to prove the existence of such bias in her case.<sup>21</sup>

It is easy to overlook this equation of causation and intentionality in disparate treatment jurisprudence. When we think about race "making a difference" in an employment decision, we usually think of a decisionmaker who, were he honest, would admit, "I just don't feel comfortable working around blacks" or "I just think men are, on the whole, more effective litigators than women." But the complete picture is more nuanced. To say that a decisionmaker made an employment decision *because* of someone's race or sex is not the same as saying that the decisionmaker meant to take that group status into account. An employee's group status may have affected the decisionmaker in completely nonconscious ways by affecting what he saw, how he interpreted it, the causes to which he attributed it, what he remembered, and what he forgot. Yet under current doctrine, if the factual record leads us to believe that race, gender, or national origin "made a difference," we must either find that the decisionmaker intended to discriminate or that no discrimination occurred. Disparate treatment doctrine is simply structured that way; these are the only two stories it knows how to tell.

Interestingly, courts long ago rejected this assumption in age discrimination jurisprudence and built a doctrinal and remedial scheme that differentiated non-conscious, unintentional bias and conscious, deliberate age discrimination. Courts have interpreted language in the Age Discrimination in Employment Act<sup>22</sup> identical to language in Title VII's section 703(a) as requiring only that age have "made a difference" or "played a part" in a decisionmaking process.<sup>23</sup> To establish "first tier" liability<sup>24</sup> entitling him to back pay, reinstatement, attorney's fees, and other equitable relief, an age discrimination plaintiff need not

---

21. This phenomenon is illustrated by the Sixth Circuit's decision in *Nichelson v. Quaker Oats Co.*, 752 F.2d 1153, 1156 (6th Cir. 1985). In reversing a jury verdict in favor of a black disparate treatment plaintiff, the court opined:

We are aware that employment discrimination based on race can occur both in subtle and obvious ways, both of which are contrary to the equal opportunity goals set out by Congress in Title VII . . . . We are also aware that subtle forms of discrimination may be difficult to prove, but the burden of proof is on the plaintiff . . . .

*Id.* at 1156; *see also* *O'Brien v. Sky Chefs, Inc.*, 670 F.2d 864, 866 (9th Cir. 1982) (holding that the plaintiff must show that more likely than not, the employer failed to promote women because of discriminatory intent).

22. The Age Discrimination in Employment Act, 29 U.S.C. § 623(a) (1988), provides in pertinent part:

(a) It shall be unlawful for an employer—

(1) to fail or refuse to hire or to discharge any individual or otherwise discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's age.

23. *See* *Lowe v. Commack Union Free Sch. Dist.*, 886 F.2d 1364, 1376 (2d Cir. 1989) (holding that court did not err in instructing jury that to prevail in an age discrimination claim, plaintiff must show age was a "significant contributing factor" in employer's decision); *Grant v. Hazelett Strip-Casting Corp.*, 880 F.2d 1564, 1568 (2d Cir. 1989) (holding that an inference of discrimination can be drawn absent direct evidence of discriminatory intent); *Burlew v. Eaton Corp.*, 869 F.2d 1063, 1067 (7th Cir. 1989) (holding that a "finding that age was a 'determining factor' . . . is tantamount to a finding of intentional discrimination").

24. The ADEA establishes two tiers of liability. Under the first tier, an employer can be held liable for damages for a simple violation of the ADEA. Under the second tier, the employer can be held

prove invidious intent—only that age “played a role” in motivating, in the sense of animating or inducing, the decisionmaker’s action. Unlike Title VII caselaw, in which motivation and intentionality are used interchangeably, age discrimination jurisprudence has distinguished these two constructs.<sup>25</sup> Consider the following language from the Seventh Circuit’s 1989 opinion in *Burlew v. Eaton Corp.*:

This standard—that age was a determining factor—does not in itself require a finding as to defendant’s state of mind, for in law there is a distinction between motive and intent. “Motive is what prompts a person to act, or fail to act. Intent refers only to the state of mind with which the act is done or omitted.” Indeed, in 1981, we stated:

Congress, in our opinion, intended that liability under the ADEA could be established without any showing as to the defendant’s state of mind. . . .<sup>26</sup>

It is difficult to understand why the insight the *Burlew* Court articulated—that motive and intent are not homologous constructs—could have escaped courts interpreting Title VII. But this is the state of equal employment opportunity jurisprudence.<sup>27</sup>

The requirement that a Title VII plaintiff prove purposeful discriminatory intent applies both in “pretext” cases, and in “mixed-motives” cases analyzed under the Supreme Court’s 1989 decision in *Price Waterhouse v. Hopkins*.<sup>28</sup> One might reasonably assume that *Price Waterhouse*, in which liability was based on the finding that gender-based stereotypes infected the employer’s decisionmaking process, would have shaken Title VII disparate treatment analysis free from the equation of causation and intentionality and moved it into greater alignment with the understanding of first-tier ADEA liability articulated by the Seventh Circuit in cases such as *Burlew* and *La Montagne*.<sup>29</sup> But a

---

liable for double damages if it committed a willful violation of the ADEA. See 29 U.S.C. § 626(b) (1994).

25. *Hazen Paper Co. v. Biggins*, 113 S. Ct. 1701, 1708 (1993) (holding that “willful violation” refers to an employer’s act evincing knowledge or reckless disregard of ADEA prohibitions).

26. *Burlew*, 869 F.2d at 1066 (quoting BLACK’S LAW DICTIONARY 727 (5th ed. 1979); Syvock v. Milwaukee Boiler Mfg. Co., 665 F.2d 149, 144-55 (7th Cir. 1981)) (emphasis in original); see also EEOC v. Century Broadcasting Corp., 957 F.2d 1446, 1458 (7th Cir. 1992) (stating that “the legislative history of the ADEA suggests that the Congressional framers thought that nonwillful discrimination directed towards an individual was quite possible”) (quoting Syvock v. Milwaukee Boiler Mfg. Co., 665 F.2d 149, 155 (7th Cir. 1981)); MacDonald v. Eastern Wyo. Mental Health Ctr., 941 F.2d 1115, 1119 (10th Cir. 1991) (stating that because age discrimination may be subtle, ADEA plaintiffs “may establish discrimination indirectly”); Graefenhain v. Pabst Brewing Co., 827 F.2d 13, 20 (7th Cir. 1987) (noting that “federal courts have long permitted [ADEA] plaintiffs to submit indirect evidence of discrimination”); Wood v. Southern Bell Tel. & Tel. Co., 725 F. Supp. 1244, 1252 (N.D. Ga. 1989) (holding that an ADEA plaintiff can establish a prima facie case with circumstantial evidence); Trent v. AT & T Technologies, Inc., 716 F. Supp. 1461, 1465 (N.D. Ga. 1989) (holding that “[a] jury could find that an employer intentionally discriminated even though . . . he did not do so willfully”).

27. See *Patterson v. McLean Credit Union*, 491 U.S. 164, 186 (1989) (requiring plaintiff to prove purposeful discrimination); *Personnel Adm’r v. Feeney*, 442 U.S. 256, 279 (1979) (“‘Discriminatory purpose’ . . . implies more than intent as volition or intent as awareness of consequences.”); *Warren v. Halstead Indus., Inc.*, 802 F.2d 746, 751-52 (4th Cir. 1986) (“Discriminatory intent means actual motive and is not a legal presumption to be drawn from a factual showing of something less than actual motive.”).

28. 490 U.S. 228 (1989).

29. *La Montagne v. American Convenience Prod., Inc.*, 750 F.2d 1405 (7th Cir. 1984).

careful reading of *Price Waterhouse* reveals that there, as in pretext cases, the concepts of motive, intent, and causation are confounded and liability is premised on the presence of conscious discriminatory animus.

For example, while Justice Brennan's lead opinion, joined by Justices Blackmun, Marshall, and Stevens, rejects the requirement that a plaintiff show "but for" causation, it does not reject the assumption that causation in employment discrimination is commensurate with conscious discriminatory intent. States Justice Brennan:

In saying that gender played a motivating part in an employment decision, we mean that, if we asked the employer at the moment of the decision what its reasons were and if we received a truthful response, one of those reasons would be that the applicant or employee was a woman.<sup>30</sup>

Justice Brennan states that a plaintiff proceeding under a mixed-motives theory must prove that an employer "relied upon sex-based considerations in coming to its decision."<sup>31</sup> And while at one point Justice Brennan writes that a mixed-motives plaintiff must prove only that her group status "played a part" in the employer's decision, it then defines "playing a part" as the conscious inclusion of the plaintiff's group status in the decisionmaker's judgmental calculus.<sup>32</sup>

Thus, the plurality opinion in *Price Waterhouse* frames causation not simply as an attempt to discern what actuated an employer's decision, but as an inquiry into the employer's conscious state of mind at the moment a decision was made. The distinction between motive and intent made by the Seventh Circuit in *Burlew* and subsequent cases appears to have escaped the *Price Waterhouse* Court. In both the plurality opinion and Justice O'Connor's influential concurrence, evidence that a decisionmaker holds stereotyped views of the plaintiff's group is deemed evidentially significant not in and of itself, but because it is assumed to betoken discriminatory animus. Nor is evidence of stereotyping itself sufficient to establish a Title VII violation—the plaintiff must prove the connection between stereotyping and discriminatory intent to prevail.<sup>33</sup>

30. *Price Waterhouse*, 490 U.S. at 250.

31. *Id.* at 242.

32. *Id.* at 250.

33. This is the only point on which the *Price Waterhouse* dissenters agreed with their colleagues. Justice Rehnquist wrote:

Although the District Court's version of Title VII liability is improper under any of today's opinions, I think it important to stress that Title VII creates no independent cause of action for sex stereotyping. Evidence of use by decisionmakers of sex stereotypes is, of course, quite relevant to the question of discriminatory intent. The ultimate question, however, is whether discrimination caused the plaintiff's harm.

*Id.* at 294 (Rehnquist, C.J., dissenting).

Numerous post-*Price Waterhouse* cases express this view of the connection between stereotyping and discriminatory motive. See, e.g., *Hong v. Children's Memorial Hosp.*, 993 F.2d 1257, 1265-66 (7th Cir. 1993) (remarks evidencing ethnic stereotypes do not necessarily prove intentional discrimination); *Bruno v. City of Crown Point*, 950 F.2d 355, 362 (7th Cir. 1991) (holding that the fact that decisionmaker asked only female applicant family-oriented questions was relevant, though insufficient, to prove discriminatory intent); *Sischo-Nownejad v. Merced Community College Dist.*, 934 F.2d 1104, 1112 (9th Cir. 1991) (stating that stereotyped remarks indicated presence of discriminatory intent); *Lin-*

One can reasonably conclude that evidence of stereotyping indicates intentional discrimination if one understands the relationship between stereotyping and biased decisionmaking in any of three particular ways. First, the assumed connection is warranted in situations where stereotypes bias decisionmaking through the *conscious* use of race, sex, national origin, or age as a proxy for some other characteristic stereotypically associated (or disassociated) with group membership. Without the assumption that the use of group membership as a proxy for some other characteristic is a conscious process, one cannot reasonably conclude that this type of stereotype-based discrimination is necessarily intentional.

Cases in which plaintiffs have prevailed under a stereotyping/proxy theory are relatively common in age and sex discrimination jurisprudence,<sup>34</sup> but they are increasingly rare in the race and national origin contexts. Seldom is an employer willing to admit, or a plaintiff able to prove, that the decisionmaker consciously used race or national origin as a proxy for some job-related trait. But, if we take the plurality opinion in *Price Waterhouse* at face value, this is precisely what a plaintiff must establish if he is to show that the employer "relied upon"<sup>35</sup> or "took into account"<sup>36</sup> race or ethnicity in coming to a decision.

Second, the assumption that evidence of stereotyping indicates discriminatory intent makes sense if stereotypes are understood as normative constructs. A normative stereotype is a composite of one's expectations as to how members of a particular group should behave. Price Waterhouse decisionmakers' comments that as a woman partner candidate, plaintiff Ann Hopkins should "walk more femininely, talk more femininely, dress more femininely, wear make-up, have her hair styled, and wear jewelry"<sup>37</sup> reflect normative stereotypes.

---

dahl v. Air France, 930 F.2d 1434, 1439 (9th Cir. 1991) (holding that decisionmaker comments revealing stereotypes regarding the relative abilities of male and female flight attendants indicated discriminatory motive); Grant v. Hazelett Strip-Casting Corp., 880 F.2d 1564, 1569 (2d Cir. 1989) (finding comments reflecting stereotypes are direct evidence of discriminatory intent); see also note 4 *supra*.

34. With respect to the use of age as a proxy, see, e.g., Hazen Paper Co. v. Biggins, 113 S. Ct. 1701, 1706 (1993) ("[A]n employer cannot rely on age as a proxy for an employee's remaining characteristics, such as productivity, but must instead focus on those factors directly."). Accord Baker v. Delta Air Lines, Inc., 6 F.3d 632, 645 (9th Cir. 1993) (age can not be used as a proxy for other disqualifying factors); Abbott v. Federal Forge, Inc., 912 F.2d 867, 876 (6th Cir. 1990) (ADEA is "directly aimed at the evil of taking age into account in making employment decisions, or using age as a proxy for some legitimate factor, with which it is somewhat, but not totally, correlated."). With respect to the use of gender as a proxy, see, e.g., Arizona Governing Comm. v. Norris, 463 U.S. 1073, 1086 (1983) (holding that employers cannot calculate retirement benefits based on sex); Dothard v. Rawlinson, 433 U.S. 321, 332 (1977) (holding employers may not use gender as a proxy for strength, although it might be a fairly accurate one). In the constitutional context, see, e.g., Craig v. Boren, 429 U.S. 190, 197-98 (1976) (holding that use of gender as a proxy for other, more germane bases of classification violates the Equal Protection Clause of the 14th Amendment).

35. *Price Waterhouse*, 490 U.S. at 241. Justice Brennan asserted that the critical inquiry in determining whether an employer violated Title VII is "whether gender was a factor in the employment decision at the moment it was made." *Id.* The focus is on the decisionmaker's thoughts at the moment of decision.

36. *Id.* at 239.

37. *Id.* at 235 (citing Hopkins v. Price Waterhouse, 618 F. Supp. 1109, 1117 (D.C. Cir. 1985)).

Expressions of gender-based normative stereotypes do indeed indicate that an employee's gender is entering into decisionmaking; it is logical to infer from such statements that males and females are being subjected to differing role expectations and behavioral standards because of gender. If one understands stereotypes in this way, it is reasonable to conclude that statements reflecting stereotypes, made in connection with a particular employment decision, indicate the presence of conscious discriminatory intent.

Third, the assumption that statements reflecting stereotyped views betoken discriminatory animus makes sense if one understands discrimination as resulting from prejudice and further understands prejudice as comprising a cognitive component (stereotypes), an affective component (aversion or dislike), and a behavioral component (discrimination aimed at creating or enforcing social distance). If we believe that these three components necessarily function as inseparable parts of an integrated whole, then the presence of one can be assumed to evidence the others. Significantly, this view of the relationship between intergroup animus, stereotyping, and discrimination characterized social psychological theories about intergroup relations prevalent in the 1950s, 1960s, and into the 1970s.<sup>38</sup>

As these observations indicate, the current jurisprudential understanding of the nature and causes of discrimination did not originate in federal discrimination caselaw. Rather, current jurisprudence reflects the perspective taken by personality and social psychologists who studied and theorized about intergroup discrimination from the 1920s into the 1980s, before the emergence of a cognitive approach to intergroup relations. While a thorough examination of this history is beyond the scope of this piece, a brief review will illustrate points significant to our inquiry.

Prejudice emerged as a subject of psychological inquiry during the 1920s.<sup>39</sup> Between that time and the present, one can distinguish various stages reflecting different theoretical and empirical approaches to the subject.<sup>40</sup> However these stages are divided, each is associated with particular social circumstances and

---

38. See notes 103-111 *infra* and accompanying text.

39. Before that time, few academics questioned the central premise of "race theory"—that non-white peoples were inferior to whites and that racial distinctions were a rational response to those "obvious" differences. See generally Franz Samelson, *From "Race Psychology" to "Studies in Prejudice": Some Observations on the Thematic Reversal in Social Psychology*, 14 J. HIST. BEHAVIORAL SCI. 265 (1978) (examining the shift in psychology from embracing the idea of racial inferiority to labeling it as irrational bias); see also JOHN DUCKITT, *THE SOCIAL PSYCHOLOGY OF PREJUDICE* (1992) (exploring the social and psychological variables contributing to prejudice and the means by which their expression can be mitigated); John F. Dovidio & Samuel L. Gaertner, *Prejudice, Discrimination, and Racism: Historical Trends and Contemporary Approaches*, in *PREJUDICE, DISCRIMINATION, & RACISM* 1, 1-34 (John F. Dovidio & Samuel L. Gaertner eds., 1986) (examining the extent to which contemporary egalitarian ideals have affected racial attitudes and tracking trends in attitudes and stereotypes about race); Graham M. Vaughan, *The Psychology of Intergroup Discrimination*, 17 N.Z. J. PSYCHOL. 1 (1988) (tracing the development of theories about prejudice from genetic theories to current sociopsychological models).

40. Scholars divide the history in different ways. Phyllis Katz, for example, discerns four distinct research periods. Phyllis A. Katz, *Racism and Social Science: Towards a New Commitment*, in *TOWARDS THE ELIMINATION OF RACISM* 3, 8-11 (Phyllis A. Katz ed., 1978). John Duckitt, in contrast, describes seven distinct periods. DUCKITT, *supra* note 39, at 47.

historical events, and each focuses attention on specific theoretical constructs and empirical agendas.

The 1920s and 1930s saw a radical change in how the academy approached the study of race relations. Samelson characterized the change as follows: "In 1920, most psychologists believed in the existence of mental differences between races; by 1940, they were searching for the sources of 'irrational prejudice.'"<sup>41</sup> Thus, from the 1920s to the early 1930s, social psychologists were busily repudiating the "accepted wisdom" of nonwhite racial inferiority posited by nineteenth century race theory and taking their first steps to understand the longevity of the belief in white superiority. In this endeavor, they invented the concept of prejudice,<sup>42</sup> designed instruments to measure it,<sup>43</sup> and began to theorize about its psychological origins.<sup>44</sup>

During the late 1930s through the early 1950s, research and theoretical work on prejudice centered primarily around the question of intrapsychic etiology. The problem was straightforward: If prejudice were irrational and unjustifiable, as the academy had collectively come to believe, how could its geographical and temporal ubiquity be explained? Psychodynamic theory provided a ready explanation: Prejudice is a defense mechanism. It subconsciously diverts inner conflicts, needs, and externally induced frustrations onto less threatening external targets.<sup>45</sup> The universality of these intrapsychic phenomena explained the temporal and geographical ubiquity of prejudice.<sup>46</sup>

After World War II, the theoretical paradigm shifted somewhat. While prejudice was still seen as psychodynamically based, its origins were no longer sought in universal psychodynamic processes. Rather, prejudice was seen as stemming from a particular pathological personality structure. Thus the problem confronting students of intergroup relations became identifying the prejudice-prone personality—the bigot. Primary symptoms of this pathological personality structure included "stereotypy," the tendency toward "either-or," categorical thinking, and hostility towards racial or ethnic "outgroups."

41. Samelson, *supra* note 39, at 265.

42. See generally GARDNER MURPHY, LOIS BARCLAY MURPHY & THEODORE M. NEWCOMB, *EXPERIMENTAL SOCIAL PSYCHOLOGY* (1931) (containing parts on measuring and changing attitudes concerning race and ethnic relations).

43. See Daniel Katz & Kenneth Braly, *Racial Stereotypes of One Hundred College Students*, 28 *J. ABNORMAL & SOC. PSYCHOL.* 280 (1933).

44. See T.W. ADORNO, ELSE FRENKEL-BRUNSWIK, DANIEL J. LEVINSON, & R. NEVITT SANFORD, *THE AUTHORITARIAN PERSONALITY* (1950); JOHN DOLLARD, LEONARD W. DOOB, NEAL E. MILLER, O.H. MOWRER, & ROBERT R. SEARS, *FRUSTRATION AND AGGRESSION* (1939); JOHN DOLLARD, *CASTE AND CLASS IN A SOUTHERN TOWN* (1937).

45. See, e.g., DOLLARD, DOOB, *supra* note 44; I.D. MACCRONE, *RACE ATTITUDES IN SOUTH AFRICA: HISTORICAL, EXPERIMENTAL AND PSYCHOLOGICAL STUDIES* (1937); Gordon W. Allport & Bernard M. Kramer, *Some Roots of Prejudice*, 22 *J. PSYCHOL.* 9 (1946); Carl Iver Hovland & Robert R. Sears, *Minor Studies in Aggression VI: Correlation of Lynchings with Economic Indices*, 9 *J. PSYCHOL.* 301 (1940); Helen V. McLean, *Psychodynamic Factors in Racial Relations*, 244 *ANNALS AM. ACAD. POL. & SOC. SCI.* 159 (1946). For a review of the psychodynamic approach to prejudice, see GORDON W. ALLPORT, *THE NATURE OF PREJUDICE* (1954).

46. The most influential work under this paradigm was ADORNO, *supra* note 44.

By 1954, when Gordon Allport wrote his now classic work, *The Nature of Prejudice*,<sup>47</sup> a mind-dazzling number of definitions of prejudice had been proposed. Among them all, there existed at least two points of agreement: first, that prejudice is a negative orientation, and second, that prejudice is an "attitude."<sup>48</sup>

During the entire period preceding the emergence of the cognitive perspective, discrimination was seen as resulting from prejudice. Prejudice, in turn, was understood as an attitude, consisting of various dimensions or components. During the 1920s and 1930s, and into the 1940s, attitudes were conceived as unidimensional, consisting of only an affective component.<sup>49</sup> During the late 1940s and into the 1960s, attitudes came to be understood as consisting of three components: beliefs about the attitude object (the cognitive component), feelings toward the attitude object (the affective component), and behavioral dispositions toward the attitude object (the behavioral component).<sup>50</sup> During the 1970s and into the 1980s, a 2-component model was advanced, in which the attitude of prejudice was defined as "[a] learned disposition consisting of . . . (1) negative beliefs or stereotypes ([the] cognitive component), and (2) negative feelings or emotions ([the] affective component)."<sup>51</sup> Under this 2-component approach, the behavioral element was split off and viewed as an independent construct referred to as the "behavioral intention."<sup>52</sup> The behavioral intention was seen as comprising a consciously formed intent to act towards the attitude object in a particular way.<sup>53</sup> Seen in this way, prejudice, an attitude, causes discrimination, a behavior. Mediating between the two is conscious behavioral intention. Discrimination, in other words, is the intentional, behavioral manifestation of prejudice.<sup>54</sup>

47. ALLPORT, *supra* note 45.

48. See, e.g., Richard Ashmore, *The Problem of Intergroup Prejudice*, in SOCIAL PSYCHOLOGY 1 (B.E. Collins ed., 1970); John Harding, Bernard Kutner, Harold Proshansky, & Isidor Chein, *Prejudice and Ethnic Relations*, in HANDBOOK OF SOCIAL PSYCHOLOGY 1021-22 (Gardner Lindzey ed., 1954) (explaining that a prejudice is an attitude toward members of some outgroup and in which the evaluative tendencies are predominantly negative). For a useful review of various definitions of prejudice encompassing these components, see DUCKITT, *supra* note 39, at 10.

49. See, e.g., L.L. THURSTONE & E.J. CHAVE, THE MEASUREMENT OF ATTITUDE 6-7 (1929); R. LIKERT, A TECHNIQUE FOR THE MEASUREMENT OF ATTITUDES (1931); Louis Guttman, *A Basis for Scaling Qualitative Data*, 9 AM. SOC. REV. 139 (1944).

50. See, e.g., DAVID KRECH & RICHARD S. CRUTCHFIELD, THEORY AND PROBLEMS OF SOCIAL PSYCHOLOGY (1948); THEODORE M. NEWCOMB, RALPH H. TURNER, & PHILIP E. CONVERSE, SOCIAL PSYCHOLOGY 40-79 (1965); Isidor Chein, *Notes on a Framework for the Measurement of Discrimination and Prejudice*, in RESEARCH METHODS IN SOCIAL RELATIONS 381, 386-90 (Marie Jahoda, Morton Deutsch, & Stuart W. Cook eds., 1951); John Harding, Harold Proshansky, Bernard Kutner, & Isidor Chein, *Prejudice and Ethnic Relations*, in 5 HANDBOOK OF SOCIAL PSYCHOLOGY 1, 4, 7-11 (Gardner Lindzey & Elliot Aronson eds., 2d ed. 1969); Daniel Katz & Ezra Stotland, *A Preliminary Statement to a Theory of Attitude Structure and Change*, in 3 PSYCHOLOGY: A STUDY OF A SCIENCE 423, 428-32 (Sigmund Koch ed., 1959).

51. JACK LEVIN & WILLIAM C. LEVIN, THE FUNCTIONS OF DISCRIMINATION AND PREJUDICE 66 (1982).

52. ICEK AJZEN & MARTIN FISHBEIN, UNDERSTANDING ATTITUDES AND PREDICTING SOCIAL BEHAVIOR 42-43 (1980).

53. *Id.* at 41-42.

54. See, e.g., Walter G. Stephan & David Rosenfield, *Racial and Ethnic Stereotypes*, in IN THE EYE OF THE BEHOLDER: CONTEMPORARY ISSUES IN STEREOTYPING 92, 93 (Arthur G. Miller ed., 1982).

The parallels between these psychological models of discrimination and current disparate treatment jurisprudence are apparent. Discrimination—at least in race and national origin contexts—is construed as resulting from hostile animus towards and accompanying negative beliefs about an individual because of his or her membership in a particular group. Mediating between prejudice and discrimination is the “discriminatory motive,” which is seen in both the psychology and race/national origin discrimination jurisprudence as a conscious behavioral intention to create social distance by denying outgroup members certain benefits and opportunities.

### B. *The Assumption of Rational Decisionmaking and the Presumption of Invidiousness*

Disparate treatment plaintiffs face a thorny problem. In short, courts have construed section 703 of Title VII, like 42 U.S.C. sections 1981 and 1983, to require proof of intent to discriminate in disparate treatment cases. But, as numerous courts have acknowledged, proving such intent is particularly difficult in employment-related disputes. As the Seventh Circuit observed: “Proof of such discrimination is always difficult. Defendants of even minimal sophistication will neither admit discriminatory animus nor leave a paper trail demonstrating it; and because most employment decisions involve some discretion, alternative hypotheses (including that of simple mistake) will always be possible and often plausible.”<sup>55</sup>

The pretext model of disparate treatment proof first articulated in *McDonnell Douglas* and later clarified in *Burdine* was intended to address this problem by providing a 3-step analytical framework for proving intentional discrimination in situations where only circumstantial evidence is available. Specifically, pretext analysis commences with the plaintiff’s presentation of a prima facie case of discrimination. In a termination case such as my client Miguel’s, to establish a prima facie case the plaintiff need only show that: (1) he is a member of a protected class (in his case Latino); (2) he was working in a job for which he was qualified; (3) his employment was terminated; and (4) his position remained open or was subsequently filled by someone of similar qualifications.<sup>56</sup> Upon making this showing, a burden of production—not a burden of

---

Allport also endorsed this view in *The Nature of Prejudice*, maintaining that the attitude of prejudice tends to manifest itself in negative actions towards the object of that prejudice. ALLPORT, *supra* note 45, at 14-15. As the prejudice becomes more intense, the negative actions increase in severity in the following five-step sequence: (1) antilocution; (2) avoidance; (3) discrimination, defined as “an attempt to exclude all members of the group in question from certain kinds of employment, from residential housing, political rights, educational or recreational opportunities, churches, hospitals or from some other social privileges”; (4) physical attack; and (5) extermination. *Id.*

55. *Riordan v. Kempiners*, 831 F.2d 690, 697 (7th Cir. 1987).

56. See *Texas Dep’t of Community Affairs v. Burdine*, 450 U.S. 248, 253 n.6 (1981) (quoting *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 802 (1973)); *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 802 (1973). The elements of a plaintiff’s prima facie case change slightly depending on the employment decision at issue. For example, in failure to hire or promote cases, some courts have held that subjective criteria forming part of the applicant’s qualifications—normally part of the plaintiff’s prima facie case—are more properly considered at the pretext stage, based on the rationale that the

proof<sup>57</sup>—shifts to the defendant to articulate one or more legitimate, nondiscriminatory reasons for its decision to terminate the plaintiff's employment.<sup>58</sup> This is what the box manufacturer defendant was doing when it identified tardiness and safety violations as the reasons underlying its decision. Because no one is perfect, few employers will be unable to articulate some plausible reason for firing or failing to hire or promote any particular employee or applicant. As a result, virtually all individual treatment cases turn on the third step in the *McDonnell Douglas* minuet: proof of pretext.

Under the *McDonnell Douglas/Burdine* model of disparate treatment proof, after a defendant articulates a legitimate, nondiscriminatory reason for the contested employment decision, the plaintiff can prevail only by proving that the proffered reason was not the "true reason"<sup>59</sup> for the decision, but a "pretext for discrimination."<sup>60</sup> In any cause adjudicated under the *McDonnell Douglas/Burdine* paradigm, the reason an employer offered to explain the negative action taken against a target employee must accordingly be classified as either the "true" reason for its action or a "phony reason"<sup>61</sup>—a "sham,"<sup>62</sup> "mask,"<sup>63</sup> "facade,"<sup>64</sup> or "cover-up"<sup>65</sup> for the "true" discriminatory motive.<sup>66</sup> Given the rhetoric of *McDonnell Douglas* and its progeny, finding against an employer at the third stage of proof is, in essence, finding that the employer has lied to the plaintiff and the court.<sup>67</sup>

---

"evaluation[ ] that a plaintiff lacks these qualities [is] more susceptible of abuse and more likely to mask pretext." *Fowle v. C & C Cola*, 868 F.2d 59, 64 (3d Cir. 1989).

57. See, e.g., *St. Mary's Honor Ctr. v. Hicks*, 113 S. Ct. 2742, 2753 (1993); *Burdine*, 450 U.S. at 254.

58. See, e.g., *Hicks*, 113 S. Ct. at 2747; *Burdine*, 450 U.S. at 254; *McDonnell Douglas*, 411 U.S. at 802.

59. *Burdine*, 450 U.S. at 253; see also *Foster v. MCI Telecom. Corp.*, 773 F.2d 1116, 1118 (10th Cir. 1985).

60. *Hicks*, 113 S. Ct. at 2752; *Burdine*, 450 U.S. at 254-56; see also *McDonnell Douglas*, 411 U.S. at 804.

61. *Visser v. Packer Eng'g Ass'n, Inc.*, 924 F.2d 655, 657 (7th Cir. 1991) (defining pretext as a "phony reason"); *Chipollini v. Spencer Gifts, Inc.*, 814 F.2d 893, 898 (3d Cir.) (defining pretext as a "fabricated justification"), cert. denied, 483 U.S. 1052 (1987).

62. *Acrey v. American Sheep Indus. Ass'n*, 981 F.2d 1569, 1581 (10th Cir. 1992); see also *Wilson v. Belmont Homes, Inc.*, 970 F.2d 53, 54 (5th Cir. 1992); *Ramsey v. City and County of Denver*, 907 F.2d 1004, 1007 (10th Cir. 1990), cert. denied, 113 S. Ct. 302 (1992); *Gunby v. Pennsylvania Elec.*, 840 F.2d 1108, 1116, 1119 (3d Cir. 1988) (holding that jury could have concluded that [defendant's] alleged discriminatory reasons for not promoting [plaintiff] were a sham), cert. denied, 492 U.S. 905 (1989).

63. *Ezold v. Wolf, Block, Schorr and Solis-Cohen*, 983 F.2d 509, 512 (3d Cir. 1992) ("pretext contrived to mask sex discrimination"), cert. denied, 114 S. Ct. 88 (1993); *Galbraith v. Northern Telecom, Inc.*, 944 F.2d 275, 282 (6th Cir. 1991) (noting that defendant's reasons may be "unworthy of belief"), cert. denied, 112 S. Ct. 1497 (1992); see also *Chipollini*, 814 F.2d at 898.

64. *Acrey*, 981 F.2d at 1574.

65. *McDonnell Douglas*, 411 U.S. at 804-08; *EEOC v. Flasher Co.*, 986 F.2d 1312, 1317 (10th Cir. 1992) (quoting *McDonnell Douglas*, 411 U.S. at 825); *Williams v. Valentec Kisco, Inc.*, 964 F.2d 723, 726 (8th Cir.), cert. denied, 113 S. Ct. 635 (1992); *Loeb v. Textron, Inc.*, 600 F.2d 1003, 1011-12 (1st Cir. 1979).

66. *Ezold*, 983 F.2d at 522 (quoting *Lockhart v. Westinghouse Credit Corp.*, 879 F.2d 43, 49 n.7 (3d Cir. 1989)).

67. As the Seventh Circuit stated in *Visser v. Packer Eng'g Ass'n, Inc.*, 924 F.2d 655, 657, "[a] pretext, in employment law, is a reason that the employer offers for the action claimed to be discriminatory and that the court disbelieves, allowing an inference that the employer is trying to conceal a dis-

Within the pretext paradigm, it is simply not possible for an employment decision to be both motivated by the employer's articulated reasons *and* tainted by intergroup bias; the trier of fact must decide between the two.<sup>68</sup> As this implicit dichotomy was described by the Seventh Circuit in *Riordan v. Kempiners*,<sup>69</sup> once an employer comes forward with explanatory evidence, the issue becomes: "(1) Was [defendant's] proffered reason—a noninvidious reason—the true reason? (2) If not, was [plaintiff's] sex the real reason? Only if the first question was answered 'no' and the second 'yes' was [defendant] guilty of sex discrimination."<sup>70</sup>

How is the trier of fact to choose between these two mutually exclusive characterizations of an employer's proffered justification and thereby resolve the central issue of discriminatory intent? As mentioned above, in the vast majority of cases, nothing even approaching "smoking gun" evidence is available. Furthermore, while statistical evidence of underutilization of members of plaintiff's group in relation to its representation in the relevant, qualified labor market is probative of discriminatory intent,<sup>71</sup> it is widely deemed insufficient to prove pretext in an individual case.<sup>72</sup>

The most common method of proving pretext is to show that the employer's proffered reason is not worthy of credence either because it appears implausible in light of data upon which such an employment decision should have been based, or because it appears inconsistent with decisions reached in similar cases involving employees outside of plaintiff's protected class. So, in Miguel's case, for example, because the employer maintained that my client was discharged because of tardiness, absence from work, and safety violations, pretext analysis required me to obtain his attendance and discipline records and compare them with the records of his Anglo coworkers.

---

criminary reason for his action." See also *St. Mary's Honor Ctr. v. Hicks*, 113 S. Ct. 2742, 2764 ("The majority's scheme therefore leads to the perverse result that employers who fail to discover non-discriminatory reasons for their own decisions to hire and fire employees not only will benefit from lying, but must lie, to defend successfully against a disparate-treatment action.") (citations omitted) (Souter, J., dissenting).

68. See, e.g., *Loeb v. Textron, Inc.*, 600 F.2d 1003, 1012 (1st Cir. 1974) ("If an employer were to prove that he was motivated by a legitimate reason, there would be no room left for showing that reason was a 'pretext,' as pretext is 'a purpose or motive alleged or an appearance assumed in order to cloak the real intention or state of affairs.'" (citation omitted); see also *Lockhart v. Westinghouse Credit Corp.*, 879 F.2d 43, 49 n.7 (3d Cir. 1989) (holding that the issue in a pretext case is "whether either illegal or legal motives, but not both, were the 'true' motives behind the decision") (quoting *NLRB v. Transportation Management Corp.*, 462 U.S. 393, 400 n.5 (1983)).

69. *Riordan v. Kempiners*, 831 F.2d 690, 696 (7th Cir. 1987). This dichotomy reflects the approach taken in *United States Postal Serv. Bd. of Governors v. Aikens*, 460 U.S. 711, 716 (1983), in which the Court stated that "the district court must decide which party's explanation of the employer's motivation it believes."

70. *Riordan*, 831 F.2d at 696.

71. See, e.g., *Lowe v. City of Monrovia*, 775 F.2d 998, 1008 (9th Cir. 1985), *modified*, 784 F.2d 1407 (9th Cir. 1986); *Box v. A & P Tea Co.*, 772 F.2d 1372, 1379 (7th Cir. 1985), *cert. denied*, 478 U.S. 1010 (1986).

72. See, e.g., *Furnco Constr. Corp. v. Waters*, 438 U.S. 567, 580 (1978); *Person v. J.S. Alberici Constr. Co.*, 640 F.2d 916, 919 (8th Cir. 1981); *Hudson v. IBM Corp.*, 620 F.2d 351, 355 (2d Cir.), *cert. denied*, 449 U.S. 1066 (1980); *Krodel v. Young*, 748 F.2d 701, 710 (D.C. Cir. 1979), *cert. denied*, 474 U.S. 817 (1985).

More specifically, the hundreds of disparate treatment cases I had read instructed me that the conscious, discriminatory purpose required to prevail in a disparate treatment case might be inferred from the following types of evidence:

- (1) Evidence that the objective data maintained by the defendant did not support the result reached by the decisionmaker;<sup>73</sup>
- (2) Evidence that the decisionmaker seemed to undervalue or ignore facts favorable to the employee;<sup>74</sup>
- (3) Evidence that the decisionmaker made a judgment about the plaintiff without being able to point to specific events which would reasonably support such a judgment;<sup>75</sup> or
- (4) Evidence showing that similarly situated Anglo employees were on occasion treated more favorably.<sup>76</sup>

In the end, if I succeeded in uncovering evidence of this sort, pretext theory not only permitted,<sup>77</sup> but indeed compelled<sup>78</sup> me to argue that the plant manager's stated reasons were a "sham," a post hoc fabrication to cover up intentional discrimination. It would not suffice to urge that the employer was a well-intentioned "'good' person" who, through lack of care, did a "'bad'

73. *See, e.g.,* *Duffy v. Wheeling Pittsburgh Steel Corp.*, 738 F.2d 1393, 1396-98 (3d Cir.) (evidencing that actual sales data showing plaintiff's performance was in fact superior to that of nonprotected employees retained in a reduction in force), *cert. denied*, 469 U.S. 1087 (1984); *Jauregui v. City of Glendale*, 852 F.2d 1128, 1135 (9th Cir. 1988) (employer's assertion that plaintiff was denied a promotion due to lack of interpersonal skill found pretextual because plaintiff's evaluations had failed to mention specific examples of this shortcoming).

74. *See, e.g.,* *Chipollini v. Spencer Gifts, Inc.*, 814 F.2d 893, 901 (3d Cir. 1987) (showing that defendant, which attempted to justify plaintiff's termination on grounds of uncooperative behavior, could not explain highly positive material in plaintiff's prior employment evaluations).

75. *Id.* at 901 (decisionmaker testified that plaintiff was "less cooperative . . . than [he] would have liked" and "not adaptive," but was unable to offer any specific examples of uncooperative or nonadaptive behavior to justify such a conclusion); *see also Jauregui*, 852 F.2d at 1135.

76. *See, e.g.,* *Wilson v. Stroh Cos.*, 952 F.2d 942, 945 (6th Cir. 1992) (holding that plaintiff may establish pretext by establishing that other employees who engaged in similar misconduct received less severe sanctions); *Bennun v. Rutgers State Univ.*, 941 F.2d 154, 179-80 (3d Cir. 1991), *cert. denied*, 502 U.S. 1066 (1992) (finding pretext in plaintiff's showing that nonminority professor was granted tenure under more lenient standards than those applied to plaintiff); *Jauregui*, 852 F.2d at 1135 (holding that fact that white police officer was promoted despite noted deficiencies in "interpersonal skills" discredited employer's assertion that lack of interpersonal skills was basis for Hispanic officer's nonpromotion).

Were I to succeed in uncovering evidence of Anglo coworkers treated more favorably, the employer would then attempt to demonstrate that those employees are not really "similarly situated." The employer would argue, for example, that these comparable employees had "better reasons" for having been late or that the safety violations disclosed in discovery were not really "their fault," thus justifying a lesser sanction. *See, e.g.,* *Lieberman v. Gant*, 630 F.2d 60, 68 (2d Cir. 1980); *Fong v. American Airlines, Inc.*, 626 F.2d 759, 762 (9th Cir. 1980). Similarly situated individuals are those who "have been subject to the same standards and have engaged in the same conduct without such differentiating or mitigating circumstances that would distinguish their conduct or the employer's treatment of them for it." *Mitchell v. Toledo Hosp.*, 964 F.2d 577, 583 (6th Cir. 1992).

77. *See, e.g.,* *Duffy*, 738 F.2d at 1395-96; *Loeb v. Textron, Inc.*, 600 F.2d 1003, 1012 n.6 (1st Cir. 1974) ("The reasonableness of the employer's reasons may of course be probative of whether they are pretexts.").

78. *See* notes 6-8 *supra* and accompanying text.

thing.”<sup>79</sup> The pretext story boards required me to paint him as an intentional wrongdoer who was lying to the court.

In fairness, it does not necessarily follow that because a judgment was irrational, suboptimal, or inconsistent with judgments made in similar situations, it must have resulted from a deliberate intent to discriminate. Some sort of presumption must serve as a bridge between the two. Pretext analysis permits this inferential leap from an apparently irrational or inconsistent judgmental process to an intentionally discriminatory one through the operation of a “presumption of invidiousness” first articulated by the Supreme Court in *Furnco Construction Corp. v. Waters*.<sup>80</sup> The Court in *Furnco* stated:

[W]e know from our experience that more often than not people do not act in a totally arbitrary manner, without any underlying reasons, especially in a business setting. Thus, when all legitimate reasons for rejecting an applicant have been eliminated as possible reasons for the employer’s actions, it is more likely than not the employer, whom we generally assume acts only with *some* reason, based his decision on an impermissible consideration such as race.<sup>81</sup>

Or, as the presumption was subsequently described by Judge Adams of the Third Circuit: “Our intuitive understanding of bigotry supports an inference of pretext from virtually any defect in an employer’s explanation of a decision to disfavor a member of a Title VII protected class.”<sup>82</sup>

Pretext analysis thus rests on the assumption that, absent discriminatory animus, employment decisionmakers are rational actors. They make even-handed decisions using optimal inferential strategies in which all relevant behavioral events are identified and weighted to account for transient situational factors beyond the employee’s control. If an employer’s proffered explanation for its decision is shown to be irrational or implausible in light of the relevant data set, the trier of fact may conclude, and to find for the plaintiff, must conclude, that the reasons given did not really motivate the decisionmaker, but were simply contrived to mask discriminatory intent. The presumption of invidiousness permits the trier of fact to infer discriminatory intent from flaws in a decisionmaker’s inferential process. Without this presumption, one could only infer that an irrational decision was made; such a decision, in the absence of a duty to discharge only for good cause, would not be actionable.

### C. *The Moment of Decision Doctrine and the Bifurcation of Perception and Judgment*

If, according to disparate treatment theory, the inferential strategy used by a nondiscriminatory decisionmaker conforms to idealized normative standards, how can we understand the judgmental strategy employed by an employer whose decisionmaking is tainted by discriminatory motivation? Current disparate treatment jurisprudence answers this question in three different ways.

---

79. Oppenheimer, *supra* note 14, at 970.

80. 438 U.S. 567 (1978).

81. *Id.* at 577.

82. *Duffy*, 738 F.2d at 1401 (Adams, J., dissenting) (contrasting age discrimination proof, where invidiousness cannot reasonably be presumed, from race discrimination, where it can).

First, it recognizes the possibility that discrimination occurs when a decisionmaker refuses to consider an individual for a particular position either because he holds antipathy towards members of that individual's social group, or because placing the individual in the position in question would violate role expectations for members of the individual's social group. This particular model of how and why discrimination occurs is reflected in the law and economics scholars' concept of a "taste for discrimination."<sup>83</sup> In this situation, the employer's desire to exclude members of the disfavored group operates independently of—and in a sense preempts—any effort to assess the applicant's objective qualifications.

According to this approach, intergroup bias affects decisionmaking not by distorting the decisionmaker's perceptual or inferential processes, but by corrupting decisionmaking at the moment of decision. Current disparate treatment jurisprudence assumes that the biased decisionmaker, like *Furnco's* "rational actor," is or could easily choose to become aware of all decision-relevant events and of the contextual factors by which the value of those events should be utility-adjusted in decisionmaking. He could, if he chose, reach an objective judgment in light of the relevant data. However, at the moment of decision, the discriminatory decisionmaker consciously "puts his thumb on the scale;" he adds the employee or applicant's group status as a factor in the equation.

Alternatively, current disparate treatment jurisprudence recognizes that stereotypes can cause discrimination when group status is consciously used as a "proxy" for some other job-relevant trait. In this way, stereotypes are seen as influencing decisionmaking through the operation of a syllogistic reasoning process. For example, an employer who fails to select a particular woman for promotion to a highly demanding, responsible upper management position may reason:

[P<sub>M</sub>] Women with young children are preoccupied with family responsibilities and do not put their jobs first;

[P<sub>m</sub>] Margaret is a woman with young children;

[C] Margaret cannot be expected to put the job first.

According to current jurisprudence, if this reasoning serves as the sole basis for the employer's decision, the pretext model would be applied to adjudicate any resulting employment dispute. If such reasoning serves as one factor among others leading to the decision, mixed-motives analysis is applied. In other words, under current doctrine, "proxy-based discrimination" occurs when an employer uses the target employee's sex as a "decision-relevant event." If we were to model the employer's decisionmaking process symbolically, the employee's group status would appear as a factor in the equation, its predictive value varying with the strength of the employer's stereotypic expectancy. In the mixed-motives context, it would appear as one factor among others; in the pretext context, it would be the only factor considered. But in either case, the

---

83. See, e.g., GARY S. BECKER, *THE ECONOMICS OF DISCRIMINATION* 39-54 (2d ed., 1971); RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 615-17 (3d ed., 1986); John J. Donohue III, *Is Title VII Efficient?*, 134 U. PA. L. REV. 1411, 1415-20 (1986).

stereotype of working mothers is seen to operate as an explicit expectancy, by which I mean that the employee's gender is consciously taken into account by the decisionmaker. Group status, as such, functions as an "additive constant" in the employer's decisionmaking calculus.

The Supreme Court vividly illustrates this understanding of discrimination as occurring at the moment of decision in *Price Waterhouse*:

The present, active tense of the operative verbs of § 703(a)(1) ("to fail or refuse") . . . turns our attention to the actual moment of the event in question, the adverse employment decision. The crucial inquiry, the one commanded by the words of § 703(a)(1), is whether gender was a factor in the employment decision *at the moment it was made*.<sup>84</sup>

*Price Waterhouse* in essence directs the trier of fact to take a snapshot<sup>85</sup> of the decisionmaker's mental state at the moment the allegedly discriminatory decision was made. If the "additive constant" appears in the picture, liability is established. If it is absent, the trier of fact should conclude that the decision was untainted by intergroup bias and find for the defendant.

The *Price Waterhouse* plurality's focus on the moment of decision is echoed in Justice O'Connor's concurrence, which attempts to delineate the relative spheres of application of the mixed-motives and pretext paradigms. Unlike the plurality, who remained silent on the issue, Justice O'Connor would limit the applicability of mixed-motives analysis to cases in which the plaintiff presents "direct evidence of discriminatory animus in the decisional process."<sup>86</sup> "[S]tray remarks in the workplace," which Justice O'Connor defines as, "statements by nondecisionmakers, or statements by decisionmakers unrelated to the decisional process itself,"<sup>87</sup> are not sufficient in her view to trigger mixed-motives analysis.<sup>88</sup>

---

84. *Price Waterhouse v. Hopkins*, 490 U.S. 228, 240-41 (1989).

85. *See Sabree v. United Bhd. of Carpenters & Joiners Local 33*, 921 F.2d 396, 403-04 (1st Cir. 1990) (stating that *Price Waterhouse* directs the trial court "to essentially take a snapshot at the moment of the allegedly discriminatory act").

86. *Price Waterhouse*, 490 U.S. at 278.

87. *Id.* at 277.

88. Following Justice O'Connor's lead, most courts now decide whether to apply mixed-motives analysis, rather than pretext analysis, by determining whether comments reflecting stereotypes or other forms of bias were made "during the decisionmaking process." *See, e.g., Hook v. Ernst & Young*, 28 F.3d 366, 375 (3d Cir. 1994) (finding that sexually demeaning remarks made to plaintiff by her supervisors were "unrelated" to decisionmaking and thus insufficient to trigger mixed-motives analysis); *Radabaugh v. Zip Feed Mills*, 997 F.2d 444, 449-50 (8th Cir. 1993); *Griffiths v. Cigna Corp.*, 988 F.2d 457, 472 (3rd Cir. 1993); *Ostrowski v. Atlantic Mutual Ins. Co.*, 968 F.2d 171, 184-85 (2d Cir. 1992); *McCarthy v. Kemper Life Ins. Co.*, 924 F.2d 683, 686-687 (7th Cir. 1991) (holding that racial remarks must be relatively contemporaneous to the termination of employment and "related to the employment decision in question" to suffice as evidence of racial animus); *Smith v. Firestone Tire and Rubber Co.*, 875 F.2d 1325, 1330 (7th Cir. 1989) (requiring a "nexus between the statements made by defendant and the demotion of the plaintiff to demonstrate that plaintiff's race was a 'substantial factor' in the defendant's decision"); *Williams v. Mead Coated Bd., Inc.*, 836 F. Supp. 1552, 1571 (M.D. Ala. 1993) (finding racially derogatory remarks not related to the employment decision at issue were insufficient to trigger mixed-motives analysis); *O'Connor v. Consolidated Coin Caterers Corp.*, 829 F. Supp. 155, 158 (W.D.N.C. 1993) (holding that decisionmaker's comments including, "[i]t's about time we started to get some young blood in this company," were unrelated to the decision to fire plaintiff and thus did not show age was a motivating factor in the employer's decision).

This notion that discrimination occurs at the moment of decision has led to astounding results in recent disparate treatment cases. Perhaps the most shocking example is *Heim v. State of Utah*.<sup>89</sup> In *Heim*, the plaintiff alleged that she was denied access to training and advancement opportunities because of her sex. At trial, she proved that her supervisor had made the statement, "Fucking women, I hate having fucking women in the office."<sup>90</sup> Shortly after making this statement, the supervisor denied plaintiff a desirable assignment. Despite this proof, the trial court refused to give a mixed-motives instruction. Unable to prove that her gender was the sole motive for her treatment, the plaintiff lost. The Tenth Circuit affirmed, stating that, although they might have been "inappropriate and boorish," the supervisor's statements were nonetheless "stray remarks in the workplace," insufficiently connected with plaintiff's claims to prove that her gender had been a motivating factor in her supervisor's decisionmaking.<sup>91</sup>

Similarly, in *Jackson v. Harvard University*,<sup>92</sup> a tenure review case, the court dismissed as not probative of sex discrimination numerous decisionmaker practices and comments evidencing sex stereotypes. The court, relying on Justice O'Connor's concurrence in *Price Waterhouse*, opined that because those comments were made prior to the commencement of the plaintiff's tenure review, they did not tend to prove that gender played a part in the tenure decision at the time it was made.<sup>93</sup>

Since *Price Waterhouse*, the federal courts' delineation of precisely where the realm of "stray remarks" ends and the "decisionmaking process" begins has reached a level of absurdity rivaling the grotesqueries of *Jarndyce vs. Jarndyce*.<sup>94</sup> Consider, for example, *Young v. Lukens Steel Co.*,<sup>95</sup> one of a number of recent age discrimination cases applying Justice O'Connor's "stray remarks" doctrine to disparate treatment cases brought under the ADEA. In *Young*, the plaintiff's coworkers and superiors consistently referred to him as "the old man" and "the oldest guy in maintenance."<sup>96</sup> It does not require a great leap to infer from these statements that plaintiff's age was salient in his superiors' minds. Nonetheless, the trial court refused to give a mixed-motives instruction to the jury, reasoning that because "there was no reference to age in the October 4th meeting at which the plaintiff's employment was terminated,"<sup>97</sup> the comments were not made in the decisionmaking process and thus constituted mere "stray remarks in the workplace."<sup>98</sup>

This bizarre distinction between bias-revealing remarks made before rather than during the decisionmaking process makes sense only if one views deci-

---

89. 8 F.3d 1541 (10th Cir. 1993).

90. *Id.* at 1546.

91. *Id.* at 1547.

92. 721 F. Supp. 1397 (D. Mass. 1989).

93. *Id.* at 1403, 1431-32.

94. CHARLES DICKENS, *BLEAK HOUSE* 7 (Everyman's Library 1991) (n.p. 1853).

95. 1994 U.S. Dist. LEXIS 5681, 65 Empl. Prac. Dec. (CCH) 43, \*193 (E.D. Pa. 1994).

96. *Id.* at \*15.

97. *Id.* at \*18.

98. *Id.* at \*16 (citing 490 U.S. at 277 (O'Connor, J., concurring)).

sionmaking as distinct from the processes of perception and interpretation of, and memory for, decision-relevant events. If, however, subtle forms of intergroup bias intrude into decisionmaking by somehow distorting the perceived data from which decisions are made, it makes little sense to distinguish between bias-revealing comments made prior, as opposed to during, the "process of decisionmaking." As I demonstrate in Part II, cognitive forms of intergroup bias affect decisionmaking at all points along a perceptual/inferential/judgmental continuum. When interpersonal judgment is understood as an integrated system involving perception, interpretation, attribution, memory, and decisionmaking, the distinction between stereotype-revealing comments made during decisionmaking and before decisionmaking utterly breaks down.

#### D. *The Assumption of Decisionmaker Self-Awareness*<sup>99</sup>

The final assumption underlying disparate treatment analysis is the most obvious and, perhaps for that reason, the easiest to miss. Not only disparate treatment analysis, but the entire normative structure of Title VII's injunction "not to discriminate," rests on the assumption that decisionmakers possess "transparency of mind"—that they are aware of the reasons why they are about to make, or have made, a particular employment decision. Possessed of such knowledge, well-intentioned decisionmakers are able to comply with Title VII's injunction "not to discriminate." Ill-intended decisionmakers know when they are taking an employee's group status into account; when challenged, they design "pretexts" to cover their tracks.

Nowhere is the assumption of decisionmaker self-awareness better illustrated than in the "moment of decision" language from the *Price Waterhouse* plurality, considered above.<sup>100</sup> As Justice Brennan opined:

In saying that gender played a motivating part in an employment decision, we mean that, if we asked the employer at the moment of the decision what its reasons were and if we received a truthful response, one of those reasons would be that the applicant or employee was a woman.<sup>101</sup>

As the *Price Waterhouse* plurality constructs reality, employment decisionmakers have ready access to the workings of their own inferential process. If they simply chose to be truthful, they could tell us whether an employee's race, ethnicity, or gender had influenced their decision. Gilbert Ryle describes this belief, which he refers to as "the official theory," in the following terms:

[A]ccording to the official theory, a person has direct knowledge of the best imaginable kind of the workings of his own mind. Mental states and processes are (or are normally) conscious states and processes, and the consciousness which irradiates them can engender no illusions and leaves the door open for no doubts. A person's present thinkings, feelings and willings, his perceivings, rememberings and imaginings are intrinsically "phosphorescent"; their existence and their nature are inevitably betrayed to their owner.<sup>102</sup>

---

99. GILBERT RYLE, *THE CONCEPT OF MIND* 154 (1949).

100. See notes 83-88 *supra* and accompanying text.

101. *Price Waterhouse*, 490 U.S. at 250; see text accompanying note 84 *supra*.

102. RYLE, *supra* note 99, at 13-14.

Because we so readily assume that people have privileged access to the content of their own thought processes, we may easily overlook the significance of this assumption of decisionmaker self-awareness. But if one thinks about it, one must immediately recognize that the normative utility of a rule prohibiting discrimination depends entirely on decisionmaker self-awareness. One can refrain from "discriminating" only to the extent that one can accurately identify the factors inducing one's actions or decisions. Absent decisionmaker self-awareness, the nondiscrimination principle—if framed solely as a prohibitory injunction "not to discriminate"—loses its normative mooring.

## II. QUESTIONING THE ASSUMPTIONS: A COGNITIVE BIAS APPROACH TO INTERGROUP JUDGMENT AND DECISIONMAKING

In 1963, social psychologists Henri Tajfel and A.L. Wilkes performed a series of simple experiments, asking subjects to estimate the relative lengths of lines that differed in length by a constant ratio of 5 percent.<sup>103</sup> In one condition, eight lines had not been presented as belonging to separate groups. In a second condition, the longer four of the eight lines had been previously presented as belonging to "Group A," the shorter four lines as belonging to "Group B."

Tajfel and Wilkes found that once they introduced the concept of "groupness" into the situation, subjects perceived objects in different groups as more different from each other, and objects in the same group as more similar to each other, than was in fact the case. In the second condition, where the lines were previously presented as belonging to one of two groups, subjects consistently overestimated the variation between lines belonging to different groups and underestimated the variation between lines belonging to the same group. Similar distortions were not observed in the "unclassified" condition. In other words, subjects tended to "stereotype" lines based on their group membership.

Tajfel and Wilkes' findings replicated those of another social psychologist, Donald Campbell, who in 1956 had performed a series of similar experiments using groups of nonsense syllables located along a spatial continuum. Campbell's results, like Tajfel and Wilkes', demonstrated that when subjects were led to think about objects as belonging to separate groups, they systematically exaggerated the variation between objects from different groups.<sup>104</sup> Nonsense syllables, like lines of varying lengths, can fall prey to human stereotyping.

Experiments such as these represented the earliest empirical investigation of a theoretical claim made by psychologists such as W.E. Vinacke, who argued that stereotypes should be understood as cognitive structures no different

103. Henri Tajfel & A.L. Wilkes, *Classification and Quantitative Judgement*, 54 *BRIT. J. PSYCHOL.* 101, 104 (1963). For a more comprehensible description of these experiments and their theoretical implications than can be found in the original article, see Henri Tajfel, *Cognitive Aspects of Prejudice*, 25 *J. Soc. ISSUES* 79, 83-86 (1969).

104. Donald T. Campbell, *Enhancement of Contrast As Composite Habit*, 53 *J. ABNORMAL & SOC. PSYCHOL.* 350, 355 (1956).

from other categorization-related constructs.<sup>105</sup> Vinacke also suggested that principles derived from the investigation of human cognition be applied to the study of intergroup perception and judgment.

Campbell, Tajfel, and Wilkes interpreted their findings along these lines. In his 1956 study, Campbell suggested that what he termed the "bias of enhancement of contrast," is, in either person or object perception, "a natural, automatic, and inevitable aspect of imperfect learning about the individual members of overlapping groups."<sup>106</sup> Similarly, Tajfel wrote in 1969 that intergroup bias was not necessarily motivational in origin. It may result, he suggested, from the same processes of categorization, assimilation, and search for coherence that underlie all human cognition, whether the objects judged be persons of different races or lines assigned to different groups.<sup>107</sup> Campbell, Tajfel, and Wilkes' findings lent the first empirical support to what became known as the cognitive approach to intergroup bias, or social cognition theory.

The emergence of social cognition theory represented a profound shift in psychologists' thinking about intergroup bias. As was discussed in Part I,<sup>108</sup> until well into the 1970s, intergroup prejudice was generally understood as stemming from motivational processes.<sup>109</sup> Stereotypes of members of "out-groups" were seen as developing out of prejudice, and as serving to rationalize it. While psychologists such as Gordon Allport recognized that stereotyping was functionally similar to categorization,<sup>110</sup> stereotypes were seen as something "special," discontinuous with "normal" cognitive process.<sup>111</sup> Before the 1970s, few psychologists seriously entertained the notion that normal cognitive processes related to categorization might *in and of themselves* produce and perpetuate intergroup bias.

This is a central premise of social cognition theory—that cognitive structures and processes involved in categorization and information processing can in and of themselves result in stereotyping<sup>112</sup> and other forms of biased intergroup judgment previously attributed to motivational processes. The social cognition approach to discrimination comprises three claims relevant to our present inquiry. The first is that stereotyping, as Vinacke suggested in 1957, is

---

105. W. Edgar Vinacke, *Stereotypes As Social Concepts*, 45 J. Soc. PSYCHOL. 229, 229 (1957); see also Joshua A. Fishman, *An Examination of the Process and Function of Social Stereotyping*, 43 J. Soc. PSYCHOL. 27, 45-52 (1956) (implicating normal psychological and sociological processes involved in language and communication in stereotyping and suggesting the study of these processes by researchers investigating intergroup relations).

106. Campbell, *supra* note 104, at 354.

107. Tajfel, *Cognitive Aspects of Prejudice*, *supra* note 103, at 81.

108. See text accompanying notes 47-54 *supra*.

109. See, e.g., DUCKITT, *supra* note 39; Richard D. Ashmore & Frances K. Del Boca, *Conceptual Approaches to Stereotypes and Stereotyping*, COGNITIVE PROCESSES IN STEREOTYPING AND INTERGROUP BEHAVIOR 1, 10 (David L. Hamilton ed., 1981).

110. See ALLPORT, *supra* note 45, at 17-27.

111. See Ashmore & Del Boca, *supra* note 109, at 10. But see Fishman, *supra* note 105, at 27-39; Vinacke, *supra* note 105, at 230.

112. For present purposes, I am adopting Hamilton and Trolier's definition of the term stereotype as a cognitive structure that contains the perceiver's knowledge, beliefs, and expectancies about some social category. David L. Hamilton & Tina K. Trolier, *Stereotypes and Stereotyping: An Overview of the Cognitive Approach*, in PREJUDICE, DISCRIMINATION, AND RACISM, *supra* note 39, at 127, 133-37.

nothing special.<sup>113</sup> It is simply a form of categorization, similar in structure and function to the categorization of natural objects. According to this view, stereotypes, like other categorical structures, are cognitive mechanisms that *all* people, not just "prejudiced" ones, use to simplify the task of perceiving, processing, and retaining information about people in memory. They are central, and indeed essential to normal cognitive functioning.

The second claim posited in social cognition theory is that, once in place, stereotypes bias intergroup judgment and decisionmaking. According to this view, stereotypes operate as "person prototypes" or "social schemas." As such, they function as implicit theories, biasing in predictable ways the perception, interpretation, encoding, retention, and recall of information about other people. These biases are *cognitive* rather than *motivational*. They operate absent intent to favor or disfavor members of a particular social group.<sup>114</sup> And, perhaps most significant for present purposes, they bias a decisionmaker's judgment long before the "moment of decision," as a decisionmaker attends to relevant data and interprets, encodes, stores, and retrieves it from memory. These biases "sneak up on" the decisionmaker, distorting bit by bit the data upon which his decision is eventually based.

The third claim follows from the second. Stereotypes, when they function as implicit prototypes or schemas, operate beyond the reach of decisionmaker self-awareness. Empirical evidence indicates that people's access to their own cognitive processes is in fact poor. Accordingly, cognitive bias may well be both unintentional and unconscious.

#### A. *Stereotypes As Categorical Structures*

Every person, and perhaps even every object that we encounter in the world, is unique, but to treat each as such would be disastrous. Were we to perceive each object *sui generis*, we would rapidly be inundated by an unmanageable complexity that would quickly overwhelm our cognitive processing and storage capabilities. Similarly, if our species were "programmed" to refrain from drawing inferences or taking action until we had complete, situation-specific data about each person or object we encountered, we would have died out long ago. To function at all, we must design strategies for simplifying the perceptual environment and acting on less-than-perfect information. A major way we accomplish both goals is by creating categories. As cognitive psychologist Eleanor Rosch wrote in 1977, "Since no organism can cope with infinite diversity, one of the most basic functions of all organisms is the cutting up of the environment into classifications by which nonidentical stimuli can be

---

113. See Vinacke, *supra* note 105, at 241.

114. Social cognition theorists do not maintain that affective processes play no role in intergroup perception and judgement. Indeed, the most recent and in many ways most interesting investigations in social cognition explore the interface between cognitive and affective determinants of intergroup stereotyping. For a thorough and thought-provoking collection of articles on this subject, see *AFFECT, COGNITION, AND STEREOTYPING: INTERACTIVE PROCESSES IN GROUP PERCEPTION* (Diane M. Mackie & David L. Hamilton eds., 1993).

treated as equivalent."<sup>115</sup> Categories and categorization permit us to identify objects, make predictions about the future, infer the existence of unobservable traits or properties, and attribute the causation of events.

What happens when we group objects into categories? First, we tend to perceive members of the same category as being more similar to each other, and members of different categories as more dissimilar to each other, than when all the objects are viewed in aggregate.<sup>116</sup> The same results adhere when the "objects" we categorize are other human beings.<sup>117</sup>

This should come as no surprise. Categories are guardians against complexity. Their purpose is to simplify the perceptual field by distorting it, so that we experience it as less complex and more predictable than it actually is. As Eleanor Rosch observes, to categorize a stimulus object means to consider it, for some significant purpose, as both equivalent to other stimuli in the same category and different from stimulus objects not in that category.<sup>118</sup> Categorical structures can simplify the perceptual environment only if "fuzzy" differences are transformed into clear-cut distinctions. Complexity continually threatens the balance of our categorical structures. Assimilation and enhancement of contrast "guard the guardians." Thus, with object and social categories, one can predict a tendency towards thinking that "all 'x's are alike."

Second, although some debate exists on this issue,<sup>119</sup> it appears that we create a mental prototype, often visual, of the "typical" category member.<sup>120</sup> To determine whether an item is a member of a particular category, we match the object perceived with the category prototype and determine the "distance" between the two. We experience an object first as a member of its "basic"

115. Eleanor Rosch, *Human Categorization*, in *STUDIES IN CROSS-CULTURAL PSYCHOLOGY* 1, 1-2 (Neil Warren ed., 1977) [hereinafter *Human Categorization*]; see also Eleanor Rosch, *Principles of Categorization*, in *COGNITION AND CATEGORIZATION*, 27, 27-48 (Eleanor Rosch & Barbara B. Lloyd eds., 1978) [hereinafter *Principles of Categorization*]; Eleanor Rosch, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, & Penny Boyes-Braem, *Basic Objects in Natural Categories*, 8 *COGNITIVE PSYCHOL.* 382, 385-428 (1976) (describing natural object categorization experiments).

116. Campbell, *supra* note 104, at 355; Tajfel, *Cognitive Aspects of Prejudice*, *supra* note 103, at 81; see also Virginia Hensley & Shelley Duval, *Some Perceptual Determinants of Perceived Similarity, Liking, and Correctness*, 34 *J. OF PERSONALITY AND SOC. PSYCHOL.* 159-168 (1976) (describing experiments finding exaggerated perceptions of similarity and dissimilarity with respect to ingroup and outgroup members).

117. See notes 170-175 *infra* and accompanying text. I am mindful that natural object categorization will differ from social categorization in significant respects. However, the commonalities to which I refer in this section appear widely accepted. See, e.g., John H. Lingle, Mark W. Altom, & Douglas L. Medin, *Of Cabbages and Kings: Assessing the Extensibility of Natural Object Concept Models to Social Things*, in 1 *HANDBOOK OF SOCIAL COGNITION* 71, 73-78 (Robert S. Wyer, Jr. & Thomas K. Srull eds., 1984).

118. *Principles of Categorization*, *supra* note 115, at 28.

119. Compare, e.g., Lingle, *supra* note 117, at 90-98 (questioning the "exemplar" or "visual prototype" approach to category representation and suggesting a mixed exemplar and probabilistic model of category structure) with Nancy Cantor & Walter Mischel, *Prototypes in Person Perception*, in 12 *ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY* 2, 28-31 (Leonard Berkowitz ed., 1979) and *Human Categorization*, *supra* note 115, at 2-15, and *Principles of Categorization*, *supra* note 115, at 27-48 (supporting an exemplar or visual prototype approach).

120. See, e.g., *Principles of Categorization*, *supra* note 115, at 35-38; Cantor & Mischel, *supra* note 119, at 28-31.

category—the category most accessible at the moment.<sup>121</sup> Only with additional mental processing do we identify it as a member of its superordinate or subordinate categories. According to this view, we carry in our heads images of the “typical letter a,” the “typical chair,” the “typical law school professor,” and the “typical urban gang member.”

Cognitive psychologists refer to these categorical structures as “schemas.” Schemas function somewhat like theories. As Rumelhart states:

[I]t is useful to think of a schema as a kind of informal, private, unarticulated theory about the nature of events, objects or situations which we face. The total set of schemata we have available for interpreting our world in a sense constitutes our private theory of the nature of reality. The total set of schemata instantiated at a particular moment in time constitutes our internal model of the situation we face at that moment in time.<sup>122</sup>

According to schema theory, when one must construe or make predictions about an event or situation, a salient aspect of that event or situation will activate a relevant schema. Once activated, the schema influences the interpretation, encoding, and organizing of incoming information and mediates the drawing of inferences or the making of predictions about the schematized object or event.<sup>123</sup> Schemas, like other categorical structures, “enable the perceiver to identify stimuli quickly, . . . fill in information missing from the stimulus configuration, and select a strategy for obtaining further information, solving a problem, or reaching a goal.”<sup>124</sup> But the price of this cognitive economy is that categorical structures—whether prototypes, stereotypes, or schemas—bias what we see, how we interpret it, how we encode and store it in memory, and what we remember about it later. In intergroup relations, these biases, mediated through perception, inference, and judgment, can result in discrimination, whether we intend it or not, whether we know it or not.

## B. *The Cognitive Origins of Stereotypes*

Social cognition theory posits that categorization and related cognitive biases can themselves result in and perpetuate stereotypes. “Prejudice,” understood as systematic biases in intergroup judgment, can flow directly from those

121. See generally Jerome S. Bruner, *On Perceptual Readiness*, 64 *PSYCHOL. REV.* 123, 132 (1957) (discussing how the accessibility of a category to the perceived object depends on both the qualities of the object and the needs of the perceiver); E. Tory Higgins, William S. Rholes, & Carl R. Jones, *Category Accessibility and Impression Formation*, 13 *J. EXPERIMENTAL SOC. PSYCHOL.* 141, 141-42 (1977) (proposing that “prior activation” of a category increases its accessibility to subsequent perceptions); Thomas K. Srull & Robert S. Wyer, Jr., *The Role of Category Accessibility in the Interpretation of Information About Persons: Some Determinants and Implications*, 37 *J. PERSONALITY & SOC. PSYCHOL.* 1660 (1979) (examining how individuals use the most accessible category of an object or event to make subsequent judgments).

122. David E. Rumelhart, *Schemata and the Cognitive System*, in 1 *HANDBOOK OF SOCIAL COGNITION*, *supra* note 117, at 161, 166.

123. *Id.* at 167; see also David L. Hamilton, *Cognitive Representations of Persons*, in 1 *SOCIAL COGNITION: THE ONTARIO SYMPOSIUM* 135, 149-56 (E. Tory Higgins, C. Peter Herman, & Mark P. Zanna eds., 1981); Reid Hastie, *Schematic Principles in Human Memory*, in *SOCIAL COGNITION*, *supra*, at 39, 60-66.

124. Shelley E. Taylor & Jennifer Crocker, *Schematic Bases of Social Information Processing*, in *SOCIAL COGNITION*, *supra* note 123, at 89, 93-94.

stereotypes, largely independent of motivational processes.<sup>125</sup> Research conducted under this paradigm can be divided into three classes, each supporting one or more propositions significant to our present inquiry. These include the minimal group experiments, studies of salience and distinctiveness, and studies concerning illusory correlation.

### 1. *The minimal group experiments.*

After his 1963 experiment showing what happens to length perception when lines are categorized into groups,<sup>126</sup> Henri Tajfel began wondering whether similar results would obtain when the categorized objects were people. To investigate, Tajfel and his colleagues designed what is now called the "minimal group paradigm." Under this research design, subjects are grouped according to what they are told<sup>127</sup> is some minimal similarity, and are then asked to evaluate members of their own and the other group or to allocate rewards between the two groups. In some studies, subjects were told that they had been grouped according to whether they tended to underestimate or overestimate the sizes of dots.<sup>128</sup> In other studies, subjects were told that their group assignment had been based on preferences for different paintings or photographs,<sup>129</sup> and in yet others, subjects were informed that group assignment was random.<sup>130</sup>

The experiments showed that, as soon as a people are divided into groups—even on a trivial or even random basis—strong biases in their perception of differences, evaluation, and reward allocation result. As soon as the concept of "groupness" is introduced, subjects perceive members of their group as more similar to them, and members of different as more different from them, than

125. Social cognition theorists do not maintain that all stereotypic conceptions of social groups arise exclusively from cognitive processes. They acknowledge the role of social learning processes occurring in the course of development. See, e.g., David L. Hamilton, *A Cognitive-Attributional Analysis of Stereotyping*, in 12 *ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY*, *supra* note 119, at 53, 64. As Henri Tajfel explained:

The value and cognitive functions of social accentuation provide a basis for the understanding of the structure and direction of biases in intergroup attitudes and stereotypes, but they cannot tell us very much about the contents of the groups' reciprocal conceptions. . . . Conceptions of outgroups are generated in their social and historical contexts and then transmitted to individual members of groups and widely shared through a variety of channels of social influence.

Henri Tajfel, *Social Psychology of Intergroup Relations*, 33 *ANN. REV. PSYCHOL.* 1, 22 (Mark R. Rosenzweig & Lyman W. Porter eds., 1982).

126. See text accompanying note 103 *supra*.

127. In all cases, group assignment was, in fact, arbitrary.

128. See, e.g., Henri Tajfel, M.G. Billig, R.P. Bundy, & Claude Flament, *Social Categorization and Intergroup Behaviour*, 1 *EUR. J. SOC. PSYCHOL.* 149, 154-55 (1971); John W. Howard & Myron Rothbart, *Social Categorization and Memory for In-Group and Out-Group Behavior*, 38 *J. PERSONALITY & SOC. PSYCHOL.* 301, 302-03 (1980).

129. See, e.g., Vernon L. Allen & David A. Wilder, *Categorization, Belief Similarity, and Intergroup Discrimination*, 32 *J. PERSONALITY & SOC. PSYCHOL.* 971, 971 (1975); W. Doise, G. Csepeli, H.D. Dann, C. Gouge, K. Larsen, & A. Ostell, *An Experimental Investigation into the Formation of Intergroup Representations*, 2 *EUR. J. SOC. PSYCHOL.* 202, 202 (1972); Tajfel, *supra* note 128, at 165-66.

130. See, e.g., Michael Billig & Henri Tajfel, *Social Categorization and Similarity in Intergroup Behaviour*, 3 *EUR. J. SOC. PSYCHOL.* 27, 34 (1973); Anne Locksley, Vilma Ortiz, & Christine Hepburn, *Social Categorization and Discriminatory Behavior: Extinguishing the Minimal Intergroup Discrimination Effect*, 39 *J. PERSONALITY & SOC. PSYCHOL.* 773, 776-83 (1980).

when those same persons are simply viewed as noncategorized individuals.<sup>131</sup> Indeed, when offered a choice minimal group subjects prefer to view information indicating their similarity with ingroup members and their distinctiveness from outgroup members.<sup>132</sup>

While ingroup members perceive similarities between themselves and others in their group, they perceive outgroup members as being even more homogeneous. In other words, subjects tend to perceive outgroup members as an undifferentiated mass, while ingroup members are more highly differentiated.<sup>133</sup> For a social cognition theorist, the plant manager's reference to my Salvadoran client as a "Mexican" would be highly relevant, as it tended to show that he perceived Latinos as an undifferentiated outgroup. It belies his contention that he "didn't even notice" Miguel's ethnicity. Or perhaps it is true that he "didn't notice." That, in a sense, was part of the problem.

Group assignment results in biased evaluation of ingroup and outgroup members even when subjects are divided into groups on an explicitly random basis. Again, subjects evaluate ingroup members more favorably than outgroup members on a variety of trait dimensions.<sup>134</sup> Subjects are better able to recall undesirable behavior of outgroup members than similar behavior of ingroup members,<sup>135</sup> and they disproportionately attribute ingroup members' failures to situational factors and outgroup members' failures to dispositional factors.<sup>136</sup> Subjects significantly overrate the product of their own group in comparison to their evaluation of the product of an outgroup.<sup>137</sup>

Dividing subjects into even minimal groups also affects their allocation of rewards between ingroup and outgroup members. In a 1973 experiment using

131. David A. Wilder, *Perceiving Persons as a Group: Categorization and Intergroup Relations*, in COGNITIVE PROCESSES IN STEREOTYPING AND INTERGROUP BEHAVIOR, *supra* note 109, at 213, 217. Similar results obtain when subjects are asked to evaluate the personalities of persons of different national origins, in this case, Canadians and Indians. See Henri Tajfel, A. A. Sheikh, & R.C. Gardener, *Content of Stereotypes and the Inference of Similarity Between Members of Stereotyped Groups*, 22 ACTA PSYCHOLOGICA 191 (1964).

132. David A. Wilder & Vernon L. Allen, *Group Membership and Preference for Information About Others*, 4 PERSONALITY & SOC. PSYCHOL. BULL. 106, 106 (1978).

133. Hamilton & Trolie, *supra* note 112, at 131; see also Edward E. Jones, George C. Wood, & George A. Quattrone, *Perceived Variability of Personal Characteristics in In-Groups and Out-Groups: The Role of Knowledge and Evaluation*, 7 PERSONALITY & SOC. PSYCHOL. BULL. 523, 523-24 (1981); Bernadette Park & Myron Rothbart, *Perception of Out-Group Homogeneity and Levels of Social Categorization: Memory for the Subordinate Attributes of In-Group and Out-Group Members*, 42 J. PERSONALITY & SOC. PSYCHOL. 1051 (1982) (finding outgroup homogeneity effect in gender-based groups). For a review of studies reflecting the outgroup homogeneity effect, in both minimal group and social group contexts, see Patricia W. Linville, Gregory W. Fischer, & Peter Salovey, *Stereotyping and Perceived Distributions of Social Characteristics: An Application to Ingroup-Outgroup Perception*, in PREJUDICE, DISCRIMINATION & RACISM, *supra* note 39, at 165 (examining group perception bias as a function of differentiation within categories).

134. See, e.g., Doise, *supra* note 129, at 203-04 (discussing intergroup bias in physical trait evaluation); Locksley, *supra* note 130, at 776-83 (exploring intergroup bias in reward allocation using explicitly random group assignment).

135. Howard & Rothbart, *supra* note 128, at 303-06.

136. See *id.*; Wilder, *supra* note 131, at 217.

137. Bernard M. Bass & George Dunteman, *Biases in the Evaluation of One's Own Group, Its Allies and Opponents*, 7 J. CONFLICT RESOL. 16, 18-20 (1963) (assigning subjects randomly to competing and collaborating groups); Robert R. Blake & Jane Srygley Mouton, *Overevaluation of Own Group's Product in Intergroup Competition*, 64 J. ABNORMAL & SOC. PSYCHOL. 237, 238 (1962).

explicitly random groups, Michael Billig and Henri Tajfel demonstrated that subjects permitted to allocate monetary rewards among ingroup and outgroup members chose a strategy that would maximize differences in rewards between ingroup and outgroup members. They chose that strategy over both fairness strategy and maximum mutual gain strategies.<sup>138</sup> In short, intergroup discrimination, in both evaluation and reward allocation, appears to result almost automatically from the categorization of people into groups.<sup>139</sup>

2. *The curse of salience: distinctiveness-based distortions in person perception and judgment.*

Under current disparate treatment doctrine, Miguel's status as the sole Latino in an otherwise Caucasian work group provided little evidentiary value in proving discrimination.<sup>140</sup> But to a social cognition theorist, it would have tripped a number of alarms.

One of those alarms was the subject of a now classic study by Harvard psychologist Shelley Taylor and his colleagues,<sup>141</sup> who explored the consequences of a minority group member's being a "token" or "solo" in a small group. In the first experiment in that study, subjects listened to a tape recording of a discussion among six males. As each person spoke, the experimenters projected a slide of that person's picture. By playing the same tape recording, while varying the race of the person supposedly speaking, Taylor and his colleagues compared subjects' judgments of a black person when he was the only black person in an otherwise all white group and when he was in a fully integrated group.

In the "solo" condition, participants judged the black participant in more extreme ways and perceived him more prominently in the discussion than in the "integrated" condition. In a second experiment, Taylor and his colleagues found similar, even stronger effects on perceptions of solo men and solo women in mixed groups. From this study and others that replicated its findings,<sup>142</sup> it appears that ratings of salient individuals on evaluative dimensions will be more extreme than ratings of nonsalient individuals. In the case of a solo, "When she was good she was very, very good, and when she was bad, she was horrid."

Similar results obtained when subjects evaluated the qualifications of members of other racial, gender, and age groups. In 1980, Patricia Linville and

---

138. Billig & Tajfel, *supra* note 130, at 37-48. These findings present an obvious challenge to the view that market forces will eventually eliminate discrimination.

139. For a review of these studies, see, e.g., Marilyn B. Brewer & Roderick M. Kramer, *The Psychology of Intergroup Relations*, 36 ANN. REV. PSYCHOL. 219, 226-227 (1985); Locksley, *supra* note 130, at 776-83.

140. The total number of employees was too small to make the disparity statistically significant.

141. For a thorough discussion of these two experiments, see Shelley E. Taylor, *A Categorization Approach to Stereotyping*, in COGNITIVE PROCESSES IN STEREOTYPING AND INTERGROUP BEHAVIOR, *supra* note 109, at 83, 90-94.

142. See generally Leslie Z. McArthur, *What Grabs You?: The Role of Attention in Impression Formation and Causal Attribution*, in SOCIAL COGNITION, *supra* note 123, at 201-06; Shelley E. Taylor & Susan T. Fiske, *Salience, Attention, and Attribution: Top of the Head Phenomena*, in 11 ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY 249, 264-65 (Leonard Berkowitz ed., 1978).

Edward Jones conducted an experiment in which white male and female subjects evaluated law school applications containing incidental indications of the applicants' race and sex.<sup>143</sup> In one experiment, evaluators judged black applicants with strong credentials more favorably than otherwise identical white applicants. However, in a second experiment, where black applicants' credentials were weak, they evaluated them less favorably than otherwise identical white candidates. Similar effects were found in subjects' ratings of members of the opposite sex on some, but not all, evaluative dimensions.<sup>144</sup> Such findings again obtained in studies varying subject and target ages.<sup>145</sup>

Polarized evaluation of outgroup members could, of course, be attributed to subjects' motivations and attempts to balance two competing needs: a desire to enhance their relative social standing by disparaging outgroup members on the one hand, and a need to preserve a belief in their own unbiased evaluative fairness on the other. However, some social cognition theorists suggest an alternative explanation based on cognitive rather than motivational factors.

According to one such explanation,<sup>146</sup> we pay more attention to salient or distinctive stimulus objects than to those that are not salient or distinctive. The more attention we pay to an object, the more information about it we perceive, encode, and store in memory. Indeed, the more salient an object is, the more likely we are to use visual rather than verbal processes to encode and retain information about it, and substantial evidence exists supporting the proposition that we recall visually encoded information more readily than verbally encoded material.<sup>147</sup> Accordingly, an impression formed under conditions of high attention can be expected to have a stronger valence, whether positive or negative, than one formed under conditions of low attention, because the data supporting that impression are more available to the perceiver. In a sense then, the polarized evaluation phenomenon can be understood as a variant of the availability heuristic.<sup>148</sup>

Looking through a somewhat different lens, Patricia Linville and Edward Jones posit that polarized appraisal is a product of inferior outgroup differentiation, referred to earlier as the "outgroup homogeneity effect."<sup>149</sup> Drawing on the minimal group research reviewed above and on their own experimental re-

143. Patricia W. Linville & Edward E. Jones, *Polarized Appraisals of Out-Group Members*, 38 J. PERSONALITY & SOC. PSYCHOL. 689, 693 (1980).

144. *Id.* at 698.

145. See, e.g., Patricia W. Linville, *The Complexity-Extremity Effect and Age-Based Stereotyping*, 42 J. PERSONALITY & SOC. PSYCHOL. 193, 206-09 (1982); Linville, *supra* note 133, at 173-74.

146. See Hamilton, *supra* note 125, at 60-61; Hamilton & Troler, *supra* note 112, at 134-36; Russell A. Jones, *Perceiving Other People: Stereotyping as a Process of Social Cognition*, in IN THE EYE OF THE BEHOLDER: CONTEMPORARY ISSUES IN STEREOTYPING 41, 46-49 (Arthur G. Miller ed., 1982); Taylor & Fiske, *supra* note 142, at 264-68. But see McArthur, *supra* note 142, at 201-09 (explaining how studies of salience and physical attributes on impressions have produced mixed results).

147. See Taylor & Fiske, *supra* note 142, at 271-73.

148. "A person is said to employ the availability heuristic whenever he estimates frequency or probability by the ease with which instances or associations could be brought to mind." Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, 5 COGNITIVE PSYCHOL. 207, 208 (1973). For further discussion of the availability heuristic, see text accompanying note 196 *infra*.

149. See, e.g., Park & Rothbart, *supra* note 133, at 1054-58.

sults, Linville and Jones posit that people have a more complex schema of their own group than of other groups.<sup>150</sup> This greater schematic complexity results in more moderate judgments of ingroup than outgroup members.

But why would schematic complexity result in evaluative moderation? Linville and Jones suggest the following explanation. First, the more familiar one is with a group of people, the more experiences one has had with members of that group. As familiarity increases, one's awareness of the number and complexity of evaluative dimensions along which individual group members may be characterized increases. Consequently, one feels the need for more data before confidently evaluating group members. This increased appreciation of the complexity of the evaluative task, and an awareness of the relative inadequacy of available information, makes us more conservative. It pulls our evaluations, whether positive or negative, towards the "safer" mean. In contrast, we evaluate members of an undifferentiated outgroup in global terms, as either "good" or "bad."<sup>151</sup>

### 3. *Illusory correlation in the formation and maintenance of stereotypes.*

Stereotypes are correlational constructs. "Women are emotional." "Blacks are athletic." "Hawaiians are friendly." Each of these statements expresses a belief that membership in a particular group (female, black, Hawaiian) correlates highly with a particular trait (emotionalism, athleticism, friendliness). In situations where such correlational judgments are objectively unfounded, they represent a phenomenon cognitive psychologists call "illusory correlation"—the subjective perception of a strong correlation that objectively does not exist. From a cognitive perspective, stereotypes can be understood as one variant of this very common judgmental error. Thus, knowing something about the cognitive processes underlying illusory correlation might inform our understanding of how stereotypes form and what makes them so enduring in the face of contradictory evidence.

Psychologist Loren J. Chapman coined the term "illusory correlation" in 1967 and defined it as "the report by observers of a correlation between two classes of events which, in reality, (a) are not correlated, or (b) are correlated to a lesser extent than reported, or (c) are correlated in the opposite direction from that which is reported."<sup>152</sup>

In his original demonstration of this phenomenon,<sup>153</sup> Chapman presented subjects with a series of twelve word pairs, constructed by pairing each word from one list with each word from a second list. Each word pair was shown to the experimental subjects the same number of times. When subjects were

150. Linville & Jones, *supra* note 143, at 689.

151. *Id.* at 691-92. If Linville and Jones are correct, increased familiarity with an outgroup should correspond with decreased evaluative extremity. In a 1982 follow-up study, Linville demonstrated precisely such an effect—the more schematic complexity young, male subjects demonstrated with reference to the category of older males, the less polarized were their evaluations of older male targets. Linville, *supra* note 145, at 193.

152. Loren J. Chapman, *Illusory Correlation in Observational Report*, 6 J. VERBAL LEARNING & VERBAL BEHAV. 151, 151 (1967).

153. *Id.*

asked to estimate the frequency with which various words co-occurred, certain systematic biases appeared. First, subjects consistently overestimated the co-occurrence of words with a strong associative connection, such as "lion-tiger" and "bacon-eggs." Second, each of the two lists contained one word which was longer than the others; subjects overestimated the co-occurrence of these distinctively long words.<sup>154</sup>

Some years later, psychologist David Hamilton sought to extend the illusory correlation concept to the problem of social stereotypes. Specifically, he sought to explore the possibility that stereotypic judgments could form out of the same cognitive biases in the assessment of covariation that Chapman observed. In two separate experiments,<sup>155</sup> Hamilton and his colleague Robert Gifford read a series of sentences describing behaviors performed by members of two groups, labelled simply as "Group A" and "Group B." In both experiments, there appeared twice as many statements about members of Group A as Group B. Group B membership became novel or salient by virtue of its relative infrequency. The sentences described were either moderately desirable or moderately undesirable behaviors. In the first experiment, sentences described desirable behaviors twice as often as undesirable behaviors. The second experiment reversed that proportion. In both experiments, the ratio of desirable to undesirable behaviors was the same for both groups.

After the series of sentences had been presented, subjects were given a list of the behavioral descriptions and asked to indicate whether each had been performed by a member of Group A or of Group B. In both experiments, subjects significantly overestimated the extent to which members of the minority group performed the less frequent type of behavior.<sup>156</sup> Hamilton and Gifford interpreted these findings as supporting, in the context of social judgment, Chapman's hypothesis that people tend to overestimate the co-occurrence of distinctive stimulus events.

In later theoretical treatments of the connection between stereotyping and illusory correlation, Hamilton drew out the implications of these findings for

---

154. Over the following two years, Loren and Jean Chapman replicated these findings in a variety of experimental contexts. The results of those studies provide insight into how illusory correlation can explain the persistence of flawed beliefs in the validity of certain commonly accepted but objectively invalid psychodiagnostic signs. See Loren J. Chapman & Jean P. Chapman, *Genesis of Popular But Erroneous Psychodiagnostic Observations*, 72 J. ABNORMAL PSYCHOL. 193 (1967) (studying illusory correlations between psychological symptoms and drawing characteristics on the "Draw A Person" test); Loren J. Chapman & Jean P. Chapman, *Illusory Correlation as an Obstacle to the Use of Valid Psychodiagnostic Signs*, 74 J. ABNORMAL PSYCHOL. 271 (1969) (explaining how practicing psychodiagnosticians failed to report valid correlations and reported invalid correlations between male homosexuality and Wheeler-Rorschach signs; naive observers replicated the practitioners' illusory correlations).

155. David L. Hamilton & Robert K. Gifford, *Illusory Correlation in Interpersonal Perception: A Cognitive Basis of Stereotypic Judgments*, 12 J. EXPERIMENTAL SOC. PSYCHOL. 392, 392 (1976).

156. Hamilton and Gifford's findings have been replicated using favorable versus unfavorable personality traits rather than specific behaviors. See, e.g., Russell A. Jones, Joanne Scott, Jose Solernou, Audrey Noble, Joe Fiala, & Kathy Miller, *Availability and Formation of Stereotypes*, 44 PERCEPTUAL & MOTOR SKILLS 631, 631 (1977). See generally Jones, *supra* note 146, at 41-77.

stereotype formation.<sup>157</sup> He observed that majority group members tend not to have much contact with minority group members. This alone could make minority group members "salient" to members of the majority group. Negative behaviors, which occur less frequently than positive behaviors, and which trigger a kind of "automatic vigilance,"<sup>158</sup> are salient as well.<sup>159</sup> Accordingly, one can conclude that stereotyped conceptions of minority groups could result from illusory correlations between two salient variables: minority group membership and negative behavioral events.<sup>160</sup>

In addition to showing that people tend to overestimate the co-occurrence of distinctive events, the Chapmans' original research indicated that people exaggerate covariations between associatively related stimuli.<sup>161</sup> This observation is significant for understanding how cognitive processes can perpetuate stereotypes once they are in place.

In their 1967 study,<sup>162</sup> Chapman and Chapman sought to understand the persistence of psychological clinicians' belief in the diagnostic value of certain signs on the Draw-A-Person Test in the face of mounting objective evidence indicating their patent invalidity. What they found was that both clinicians and untrained college undergraduate students consistently "discovered" in the data objectively nonexistent correlations between psychological symptoms and drawing characteristics which had been shown in pre-tests to be associatively related.<sup>163</sup> So, for example, both professional and undergraduate subjects overestimated the frequency with which patients manifesting the symptom "suspiciousness" had produced drawings with atypical eye features.<sup>164</sup> These distortions in subjects' views of the presented data persisted even when subjects were provided monetary incentives for making accurate covariation estimates.<sup>165</sup> In the Chapmans' view, semantic associations between a particular psychological symptom and a particular body part served as a prior expectancy,

---

157. Hamilton, *supra* note 125, at 59-64; David L. Hamilton, *Illusory Correlation as a Basis for Stereotyping*, in *COGNITIVE PROCESSES IN STEREOTYPING AND INTERGROUP BEHAVIOR*, *supra* note 109, at 115, 124-31.

158. See Felicia Pratto & Oliver P. John, *Automatic Vigilance: The Attention-Grabbing Power of Negative Social Information*, 61 *J. PERSONALITY & SOC. PSYCHOL.* 380, 380 (1991) (observing and defining automatic vigilance as "a mechanism that serves to direct attentional capacity to undesirable stimuli").

159. Although Hamilton and Gifford found that both desirable and undesirable behaviors could be rendered salient by virtue of infrequency, other researchers subsequently demonstrated that negative events are generally more salient than positive events. See, e.g., *id.* at 381.

160. This conception of how biases in assessing covariation can lead to stereotype formation makes sense only if we assume that people encode incoming information about a target person's behavior according to that person's race, sex, ethnicity, or other social group membership. In a 1978 study, Shelley Taylor and his colleagues at Harvard provided strong empirical evidence supporting this proposition by demonstrating the existence of distinctive patterns in the distribution of within-group versus between-group errors in attributing certain statements to members of different gender or racial groups. Shelley E. Taylor, Susan T. Fiske, Nancy L. Etoff, & Audrey J. Ruderman, *Categorical and Contextual Bases of Person Memory and Stereotyping*, 36 *J. PERSONALITY & SOC. PSYCHOL.* 778, 784-85 (1978).

161. Chapman & Chapman, *Genesis of Popular But Erroneous Psychodiagnostic Observations*, *supra* note 154, at 193.

162. *Id.*

163. *Id.* at 197-99.

164. *Id.* at 196.

165. *Id.* at 202-03.

which distorted subjects' assessments of the covariations between paired symptoms and signs.<sup>166</sup>

These early observations suggest a mechanism by which social stereotypes, operating as correlational expectancies, might distort incoming information about members of different social groups and thereby insulate those stereotypes from the corrective effects of disconfirming evidence. As was discussed above, a stereotype can be understood as a "schema," a network of elements representing a person's accumulated knowledge, beliefs, and expectations about a particular category. The elements of any schema are associatively related to one another. Thus, an associative basis for an illusory correlation exists whenever a schema is activated.

Our schematic expectations about members of social groups come, no doubt, from a variety of sources. As H.J. Ehrlich stated:

Stereotypes about ethnic groups appear as a part of the social heritage of society. They are transmitted across generations as a component of the accumulated knowledge of society. They are as true as tradition, and as pervasive as folklore. No person can grow up in a society without having learned the stereotypes assigned to the major ethnic groups.<sup>167</sup>

Stereotypic expectancies may arise from one's own experiences with members of other social groups. Alternatively, as Ehrlich suggests, they may be absorbed in much the same way as folklore is learned, through the vicarious experiences of stories, television shows, movies, newspaper reports, and so forth. But wherever they come from, stereotypes can be expected to bias our subjective perception of the covariation between group membership and traits stereotypically associated with that group. We do not ignore evidence and choose to act instead on the stereotype. Rather, the stereotype, acting as an associative construct, biases the way we see the evidence. We recall, through the same cognitive processes that result in other forms of illusory correlation, stereotype-confirming instances as having occurred more frequently than they actually did.<sup>168</sup>

In summary, research conducted under the cognitive approach indicates that normal cognitive processes can lead to the creation and maintenance of social stereotypes. Seen in this way, stereotypes represent simply one manifestation of generalized cognitive biases resulting from categorization, differential attention to salient events, and the search for meaning and coherence. In a sense, we can say that human cognitive organization predisposes us to stereotyping.

---

166. *Id.* at 201.

167. H.J. EHRLICH, *THE SOCIAL PSYCHOLOGY OF PREJUDICE* 35 (1973).

168. See David L. Hamilton & Terrence L. Rose, *Illusory Correlation and the Maintenance of Stereotypic Beliefs*, 39 J. PERSONALITY & SOC. PSYCHOL. 832 (1980). See generally Hamilton, *supra* note 157.

C. *How Stereotypes Cause Discrimination: The Information Processing Approach*

How exactly do stereotypes cause discrimination? As was discussed above in Part I,<sup>169</sup> current disparate treatment caselaw answers this question in two ways. First, it recognizes that stereotypes, operating as role expectations, may cause discrimination when members of certain groups are excluded from certain roles or occupations deemed “inappropriate” for members of their group. In this circumstance, an employer doesn’t even consider an prospective employee’s qualifications; group membership trumps all other factors.

Alternatively, disparate treatment jurisprudence recognizes that stereotypes cause discrimination when group status is consciously used as a “proxy” for another trait. In either situation, however, current doctrine assumes that stereotypes operate explicitly and independently from the processes of perception, memory, or recall. They influence decisionmaking at the moment a decision is made.

Social cognition theory provides a fundamentally different explanation of how stereotypes cause discrimination. Stereotypes are viewed as social schemas or person prototypes. They operate as implicit expectancies that influence how incoming information is interpreted, the causes to which events are attributed, and how events are encoded into, retained in, and retrieved from memory. In other words, stereotypes cause discrimination by biasing how we process information about other people. Let us look more closely at the theoretical premises of this approach and at the empirical findings it has generated.

1. *Schematic bases of person perception and judgment.*

Schema theory<sup>170</sup> posits that, as people learn, they construct schemas, some of which relate to social cognition. These are referred to as “social schemas.”<sup>171</sup> They include trait schemas (“leader”), role schemas (“manufacturing plant foreman”), identity group schemas (“Hispanic men”), and person impressions (“Miguel”).<sup>172</sup>

Each schema contains a number of interrelated elements that represent a person’s accumulated knowledge, beliefs, experiences (both direct and vicarious), and expectancies regarding the schematized construct. A schema may also contain visual representations of “typical” schematic exemplars. Confronted with the task of construing a person or his behavior, one matches the incoming perception against the relevant subset of schemas.<sup>173</sup> An incoming

169. See text accompanying notes 28-38 *supra*.

170. For an overview of schema theory, particularly as it relates to social cognition, see, e.g., William F. Brewer & Glenn V. Nakamura, *The Nature and Function of Schemas*, in 1 HANDBOOK OF SOCIAL COGNITION, *supra* note 117, at 119; Hamilton, *supra* note 123; Rumelhart, *supra* note 122; Taylor & Crocker, *supra* note 124.

171. Taylor & Crocker, *supra* note 124, at 91.

172. See *id.*

173. Schemas appear to be organized in a hierarchical structure, with certain “sub-schemas” embedded into others. This structure is similar to that posited by Eleanor Rosch in her work on subordinate and superordinate natural object categories. See *Principles of Categorization*, *supra* note 115, at 30-35.

bit of information that "fits" an existing schema is said to "instantiate"<sup>174</sup> that particular schema. When this instantiation process activates a schema, other elements of the schema are then imposed on the incoming experience. In other words, once a particular schema is activated, incoming information tends to be ordered in a manner that reflects the structure of the schema.<sup>175</sup> In this way, a schema acts as an implicit expectancy: We implicitly expect incoming information instantiating a particular schema to be consistent with elements of that schema already present in its cognitive representation.

Schemas serve essential heuristic functions in human cognition. They lend structure to experience, enable us to identify the relationships between elements of a stimulus field, impute meaning to behavior or other events, and fill in data missing from a particular stimulus configuration. They provide a structure for determining what information will be encoded into memory, how it will be organized in relation to other stored information, and how it will be retrieved from memory when required.

Unfortunately, all this cognitive efficiency also results in a certain biasing of incoming information. These systematic biases may in turn lead to predictable types of error in social judgment. These include:

- errors in schema activation;
- errors in the interpretation and encoding of ambiguous events;
- errors in attributing the causes of events; and
- errors caused by biases in memorial storage and retrieval processes.

It is to these biases and the errors they may cause that our attention now turns.

## 2. *Representativeness, salience, and schema activation.*

A primary cognitive function of schemas is to help answer the questions, "What is it?" and "How is it likely to behave?" The initial matching of a stimulus object against a perceiver's existing schematic structures and the resulting activation of a particular schema represent a significant source of error in social perception and judgment.

Consider the following scenario. A twenty-five-year-old, 5-foot, 6-inch woman with a 2-year degree in paralegal studies enters the police academy in a major urban area. One of her classmates is a twenty-five-year-old 6-foot male with a 2-year college degree in physical education. Their instructors, who must evaluate them and eventually decide whether they will become part of the police force, ask as they meet these two new cadets, "Is it a police officer?" and "How will it perform as a police officer?" To answer these initial questions using normative rules of inference would be an elaborate process indeed. But whatever those normative rules might prescribe, the training officers are not likely to follow them. Rather, they are likely to form an initial impression through the use of a cognitive shortcut which psychologists Daniel Kahneman

---

174. David E. Rumelhart & Andrew Ortony, *The Representation of Knowledge in Memory*, in *SCHOOLING AND THE ACQUISITION OF KNOWLEDGE* 99, 105 (Richard C. Anderson, Rand J. Spiro, & William E. Montague eds., 1977).

175. Taylor & Crocker, *supra* note 124, at 97.

and Amos Tversky call the "representativeness heuristic."<sup>176</sup> According to Kahneman and Tversky, a person attempting to estimate the likelihood that a person or an object falls into a particular category or class matches salient attributes of the stimulus object (in our case, the cadet) with equally salient attributes characterizing the viewer's schema or prototype of the category in question ("police officer"). If these two sets of salient attributes match, the person judges the object likely to be a member of the category or class. If these sets do not match, the person judges the object unlikely to be a member of the category or class.

Thus, the specific content of the training officer's schema "police officer" can powerfully influence whether the officer initially perceives each of our two cadets as a good candidate for the job. If the most salient features in the training officer's schema for "police officer" include attributes such as "dominating," "physically imposing," or "authoritarian," or if he pictures a male as the prototypical police officer, he may less readily categorize the female cadet as "police officer material." If, instead, the training officer's schema for a police officer includes attributes such as "calms people down in tense situations," "is skilled at interviewing witnesses and writing reports," or "can understand legal rules and apply them in novel situations," he may judge the female cadet, with her degree in paralegal studies and her less intimidating physical aspect, more favorably in this initial implicit assessment. In short, the initial categorization of a person or event is influenced by schematic expectancies.<sup>177</sup> In forming our initial impressions, we judge other people by the content of our categories.

What determines which schema will be primed by an incoming perceptual object? Why does the training officer see a "woman," a "man," a "Caucasian," or an "Asian" at all when he meets a new recruit? And if he initially categorizes cadets along those lines, does that simply confirm a sexist or racist predisposition or does it evidence a more complex phenomenon?

As Eleanor Rosch observed thirty years ago, people categorize objects in their environment in a particular way because it proves useful in understanding their environment and predicting future events.<sup>178</sup> It would be difficult to argue credibly that racial, ethnic, or gender distinctions have no utility in understanding American society or negotiating experience within it. The antidiscrimination laws themselves make gender, racial, and ethnic distinctions salient. The American history of slavery, race-based immigration restrictions, Jim Crow laws, and persistent segregation in housing, schools, and economic status all further serve to make race salient. Indeed, a fear that one might be racist or sexist, or a desire not to be perceived as such, will make race and gender salient, even highly charged. As a theoretical matter, the notion that racial, ethnic,

---

176. Daniel Kahneman & Amos Tversky, *On the Psychology of Prediction*, 80 *PSYCHOL. REV.* 237, 237 (1973).

177. For further development of this claim, including a discussion of empirical studies demonstrating its application in several contexts, see Thomas K. Srull, Meryl Lichtenstein, & Myron Rothbart, *Associative Storage and Retrieval Processes in Person Memory*, 11 *J. EXPERIMENTAL PSYCHOL.: LEARNING, MEMORY, & COGNITION* 316, 339 (1985).

178. See note 115 *supra* and accompanying text.

or gender distinctions could be ignored in the priming of schematic expectancies is, at best, implausible. As an empirical matter, it is simply insupportable.<sup>179</sup> Because gender, ethnic, and racial distinctions are often perceptually apparent, and because these categories are made salient by our social and cultural context, we can expect race, ethnic, and gender-based schemas to be implicated in the processing of information about other people. Once activated, the content of a schema will profoundly affect how we interpret a person's subsequent behavior, what about that behavior we remember, and how we use the behavior in judging the person later.

### 3. *Schematic expectancies and the interpretation of ambiguous events.*

According to schema theory, incoming information is processed within the framework of the particular schema activated at that time.<sup>180</sup> The schema acts as a lens through which subsequent events are viewed. Consequently, the activation of a social schema can be expected to influence how information is interpreted and encoded into memory. A fair amount of empirical evidence supports this proposition.

For example, in 1980, H. Andrew Sagar and Janet Schofield conducted a study demonstrating the effect of social schemas on the interpretation of ambiguous information.<sup>181</sup> In the study, school-age children were presented cartoon-like drawings of two children and a verbal description of the scene, read by the experimenter. One drawing showed two students sitting in a classroom, one behind the other. The verbal description of the scene was: "Mark was sitting at his desk, working on his social studies assignment, when David started poking him in the back with the eraser end of his pencil. Mark just kept on working. David kept poking him for a while, and then he finally stopped."<sup>182</sup> In one condition, David was depicted as black, in the other condition, white. Subjects

179. I find no empirical support for the proposition that Americans no longer categorize by race, ethnicity, or sex. Indeed, available empirical evidence strongly suggests the opposite. See, e.g., John F. Dovidio, Nancy Evans, & Richard B. Tyler, *Racial Stereotypes: The Content of Their Cognitive Representations*, 22 J. EXPERIMENTAL PSYCHOL. 22 (1976) (explaining a category priming/reaction-time study demonstrating the presence of racial categories similar in form and function to the "natural object categories" described by Eleanor Rosch); Shelley E. Taylor, Susan T. Fiske, Nancy L. Eitcott, & Audrey J. Ruderman, *Categorical and Contextual Bases of Person Memory and Stereotyping*, 36 J. PERSONALITY & SOC. PSYCHOL. 778, 781-82 (1978) (concluding from error patterns in the attribution of statements to actor/accomplices that subjects organized incoming social information into racial and gender categories). For a review of more recent empirical findings indicating the continuing existence of racially differentiated categorical representations and accompanying race-based stereotypes, see Faye Crosby, Stephanie Bromley, & Leonard Saxe, *Recent Unobtrusive Studies of Black and White Discrimination and Prejudice: A Literature Review*, 87 PSYCHOL. BULL. 546 (1980); Samuel L. Gaertner & John F. Dovidio, *The Aversive Form of Racism*, in PREJUDICE, DISCRIMINATION, & RACISM, *supra* note 39, at 61.

180. See Rumelhart, *supra* note 122, at 163-66; Taylor & Crocker, *supra* note 124, at 91. For a more thorough consideration of schema activation or "priming" and its effect on the interpretation of ambiguous information, see Bruner, *supra* note 121, at 135-37; Higgins, *supra* note 121, at 140-44; Srull & Wyer, *supra* note 121, at 1661.

181. H. Andrew Sagar & Janet Ward Schofield, *Racial and Behavioral Cues in Black and White Children's Perceptions of Ambiguously Aggressive Acts*, 39 J. PERSONALITY & SOC. PSYCHOL. 590 (1980).

182. *Id.* at 593.

were asked to rate David's behavior on four scales, evaluating the extent to which they thought it was playful, friendly, mean, or threatening.<sup>183</sup>

The results demonstrated that the race of the actor had a significant impact on the manner in which subjects categorized his actions. If the actor was black, subjects judged his behavior to be more mean and threatening and less playful and friendly. The opposite result obtained when the actor was white. Thus, subjects interpreted the same behavior differently depending on the race of the actor performing it.<sup>184</sup>

Sagar and Schofield's results replicated the finding of a similar experiment conducted by Birt Duncan in 1976.<sup>185</sup> In Duncan's study, white college students watched one of four videotapes in which two males discussed alternative solutions to a particular problem. Subjects were told they were watching a live interaction happening in another room, and that its purpose was to develop a new system for rating interpersonal behavior. While the videotape played, a buzzer rang at specified intervals, signalling subjects to categorize the behavior they were then viewing in one of ten categories<sup>186</sup> and to indicate its intensity on an 8-point scale.

As the videotaped discussion progressed, the dialogue became increasingly heated. Finally, one of the participants (the protagonist) shoved the other (the victim). At that point, the buzzer rang—not for the first time—and subjects were asked to characterize and rate the intensity of the protagonist's behavior.<sup>187</sup>

As in the Sagar and Schofield study, the protagonist's race significantly affected how subjects characterized the shove. If the protagonist was white, his behavior was characterized as "playing around" or "dramatizes." If he was black, it was characterized as "aggressive" or "violent."<sup>188</sup>

Once behavior has been interpreted and encoded into memory, its meaning is in a sense "fixed," for once a person has constructed an explanation of an event, it is this construction, not the raw information that stimulated it, that is used in making subsequent judgments or predictions.<sup>189</sup> Indeed, once a behavior has been encoded as a trait, its effect on subsequent judgments increases over time.<sup>190</sup> It then supports and validates the preexisting stereotypic expectancy. So, for example, when a female or Asian cadet performs a particular behavior we may characterize it as "passive," yet characterize the same behavior as "prudent" or "restrained" when performed by a male Caucasian cadet.

183. *Id.* at 594.

184. *Id.* at 594-95.

185. Birt L. Duncan, *Differential Social Perception and Attribution of Intergroup Violence: Testing the Lower Limits of Stereotyping of Blacks*, 34 J. PERSONALITY & SOC. PSYCHOL. 590 (1976).

186. The ten categories included: "dramatizes, gives information, gives opinion, gives suggestion, asks for information, asks for opinion, asks for suggestion, playing around, aggressive behavior, and violent behavior." *Id.* at 594.

187. *Id.* at 593.

188. *Id.* at 595.

189. See, e.g., Lee Ross, Mark R. Lepper, Fritz Strack, & Julia Steinmetz, *Social Explanation and Social Expectation: Effects of Real and Hypothetical Explanations on Subjective Likelihood*, 35 J. PERSONALITY & SOC. PSYCHOL. 817, 827-28 (1977).

190. Srull & Wyer, *supra* note 121, at 1662.

One might invoke "racism" or "sexism" to describe this discrimination. Less dramatically, but perhaps more informatively, one could describe it as one way that social schemas subtly distort information processing about other people and result in intergroup bias.

#### 4. *Biases in causal attribution.*

The significance we attach to an event depends largely on the causes to which we attribute it. Without even being aware of it, we constantly ask and answer the question "why?" "Why was Miguel late to work today?" "Why were Mary's sales figures so high last year?" "Why do girls score lower than boys on the math section of the SAT?" Causal attributions shape and are shaped by our cognitive representations of persons, objects, and events. Most importantly, causal attribution substantially affects how past events inform our predictions about the future. If, for example, we attribute a behavioral transgression to a stable personality trait (e.g., "Miguel was late because he's lazy"), we are more likely to predict that the transgression will recur than if we attribute the transgression to a transient environmental causes (e.g., "Miguel was late because of the unusually bad traffic this morning"). Similarly, we are more likely to see success as predictive of future performance if we attribute it to stable dispositional factors (e.g., ability) than to transient situational ones (e.g., good luck, or even extraordinary effort).

Since the 1950s, cognitive and social psychologists have been studying the ways in which ordinary people try to understand the causes and implications of the events they witness. From this study has emerged a recognition that causal attribution is subject to systematic biases that distort person perception (how we see others) and interpersonal judgment (how we interpret their actions). While these biases generalize across attributional tasks, they are particularly pronounced in tasks involving intergroup perception and judgment. Thus, to understand the subtle ways that cognitive processes can cause intergroup discrimination, we need to know something about attribution theory.

Attribution theory stems in large part from Fritz Heider's 1958 work, *The Psychology of Interpersonal Relations*.<sup>191</sup> Heider noted that we look for the causes of an action in either the actor (the internal dimension) or the situation (the external dimension),<sup>192</sup> and that because people are so salient in any perceptual array, their behavior has a tendency to "engulf the field."<sup>193</sup> Subsequent empirical research has established quite convincingly that in attributing

---

191. FRITZ HEIDER, *THE PSYCHOLOGY OF INTERPERSONAL RELATIONS* (1958).

192. *E.g., id.* at 146-59. Later research elaborated upon this bivalent approach. Most importantly, Bernard Weiner and his colleagues noted that causal attributions can be characterized not only along an internal/external dimension, but also along a stable/transitory dimension. According to this view, attributions based on either internal or stable factors increase the predictive value assigned to an event, while attributions based on either external or transient factors tend to depress an event's perceived predictive utility. See Bernard Weiner, Irene Frieze, Andy Kukla, Linda Reed, Stanley Rest, & Robert M. Rosenbaum, *Perceiving the Causes of Success and Failure*, in *ATTRIBUTION: PERCEIVING THE CAUSES OF BEHAVIOR* 95, 113-17 (Edward E. Jones, David E. Kanouse, Harold H. Kelley, Richard E. Nisbett, Stuart Valins, & Bernard Weiner eds., 1971).

193. HEIDER, *supra* note 191, at 54.

the causes of behavior, people tend to underestimate the impact of situational factors and overestimate the impact of stable, dispositional factors.<sup>194</sup> This bias is so basic and widespread that it is referred to as the "fundamental attribution error."<sup>195</sup>

Susceptibility to the fundamental attribution error varies in a number of ways relevant to our present inquiry. First, susceptibility depends on who is doing the attributing: Actors tend to attribute their own behaviors to situational forces or constraints, whereas observers tend to attribute those same behaviors to the actor's personality, attitudes, or abilities.<sup>196</sup> Second, we tend to attribute events that confirm our prior expectancies to stable, internal factors, while attributing events that contradict prior expectancies to transient or environmental causes.<sup>197</sup> Much empirical evidence now suggests that people tend to attribute stereotype-consistent behaviors to dispositional factors and stereotype-inconsistent behaviors to transient or environmental factors.<sup>198</sup>

The most dramatic demonstration of this bias, and of its effects on behavioral prediction and the selection of sanctions, appears in two experiments conducted by Galen Bodenhausen and Robert Wyer in 1985.<sup>199</sup> In both experiments, subjects read a case file describing a behavioral transgression<sup>200</sup> by a target individual. In the first experiment, the case file contained only general biographic, demographic, and job-related information about the target. In the second experiment, the case file also contained information about the target's background and life circumstances at the time the transgression occurred.<sup>201</sup> After reading the case file, subjects were asked to predict whether the transgression would recur and to recommend a negative sanction. In one

194. For a representative sampling of this literature, see Edward E. Jones, *How Do People Perceive the Causes of Behavior?*, 64 *AM. SCIENTIST* 300 (1976); Jones, *supra* note 146, at 57-64; Lee Ross, *The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process*, in *COGNITIVE THEORIES IN SOCIAL PSYCHOLOGY*, 337, 347 (Leonard Berkowitz ed., 1978).

195. Ross, *supra* note 194, at 347.

196. The leading work on these actor/observer differences in susceptibility to the fundamental attribution error is Edward E. Jones & Richard E. Nisbett, *The Actor and the Observer: Divergent Perceptions of the Causes of Behavior*, in *ATTRIBUTION: PERCEIVING THE CAUSES OF BEHAVIOR*, *supra* note 192, at 79.

Actor/observer differences are widely thought to result from the differential salience of the actor and the environment to actors and observers. For the actor, the environment is more salient than himself; for the observer, the actor is perceptually more prominent. In a sense, then, the fundamental attribution error can be understood as a variant of the availability heuristic. For a thorough review of this perspective, see Taylor & Fiske, *supra* note 142, at 253.

197. Jennifer Crocker, Darlene B. Hannah & Renée Weber, *Person Memory and Causal Attributions*, 44 *J. PERSONALITY & SOC. PSYCHOL.* 55, 56 (1983).

For a discussion of how this differential attribution might contribute to the maintenance of social stereotypes in the face of stereotype-disconfirming evidence, see text accompanying notes 205-210 *infra*.

198. See, e.g., Kay Deaux, *Sex: A Perspective on the Attribution Process*, in 1 *NEW DIRECTIONS IN ATTRIBUTION RESEARCH* 335 (John H. Harvey, William John Ickes, & Robert F. Kidd eds., 1976); Galen V. Bodenhausen & Robert S. Wyer, Jr., *Effects of Stereotypes on Decision Making and Information-Processing Strategies*, 48 *J. PERSONALITY & SOC. PSYCHOL.* 267, 268, 279 (1985); Hamilton, *supra* note 125, at 65-68.

199. Bodenhausen & Wyer, *supra* note 198.

200. In the first experiment, the transgression involved job-related behavior, in the second, a criminal act. *Id.* at 270, 272.

201. *Id.* at 270, 273.

condition, the transgression was stereotypic of the target's ethnic group, in a second, the transgression was stereotype-inconsistent, and in a third, no stereotype was activated.

When the transgression was stereotypic of the target's ethnicity, subjects perceived it as more likely to recur. They therefore assigned more severe punishments in that situation than when the transgression was stereotype-inconsistent or when no stereotype had been activated.<sup>202</sup> Furthermore, when the transgression was stereotype-inconsistent with the target's ethnicity, or when no stereotype was activated, subjects were better able to recall information about relevant life circumstances presented in the file.<sup>203</sup> From these results, the experimenters reasonably inferred that subjects spent more time seeking and considering potential situational causes for the transgression in the no-stereotype and stereotype-inconsistent conditions than they spent when the behavior suggested a stereotype-consistent explanation.<sup>204</sup>

Bodenhausen and Wyer's findings bear substantial implications for how Title VII jurisprudence might interpret differences in the treatment of "similarly situated" workers. Assume that two employees, one male and one female, have performed poorly on a stereotypically male task, and that the female employee was disciplined more severely than the male. The employer maintains that her poor performance, not her sex, motivated the disciplinary action taken.

Under current disparate treatment doctrine, evidence that the female employee was treated less favorably than her male counterpart would be interpreted as indicating that the employer's stated reason for imposing the discipline was a "sham," a post hoc fabrication designed to cover up deliberate, intentional discrimination.<sup>205</sup> Attribution theory provides an alternative explanation for cases such as this. As we have seen, a decisionmaker tends to impose a more severe sanction for the same transgression when that transgression is attributed to stable/dispositional as opposed to transient/situational causes. And as we have also seen, stereotypes operate as judgment heuristics in causal attribution. Once a stereotype is activated, observers tend to attribute stereotype-consistent behaviors to stable, stereotypic traits without bothering to search for other potentially relevant information. The stereotype operates as a kind of cognitive shortcut, bringing the search for additional causal antecedents to a screeching halt.<sup>206</sup>

Thus, in our hypothetical, the decisionmaker attributes failure by the female employee to stable dispositional causes associated with her gender. This results in a more severe sanction than similar failure by the male employee, whose failure the decisionmaker treats as a product of transient, environmental causes.

---

202. *Id.* at 271, 275-76.

203. *Id.* at 277-78.

204. *Id.* at 280-81.

205. See text accompanying notes 73-82 *supra*.

206. As David Kanouse has recognized, people tend to think of unitary actions or events as having a single, as opposed to multiple causes. Consequently, they exert more cognitive effort in seeking a causal explanation when none has yet come to mind than they will exert in seeking additional explanations. David E. Kanouse, *Language, Labeling, and Attribution, in* ATTRIBUTION: PERCEIVING THE CAUSES OF BEHAVIOR, *supra* note 192, at 121, 131.

The decisionmaker may be wholly unaware that the employee's gender in any way influenced his decisionmaking.<sup>207</sup> As he sees it, the two employees simply are not similarly situated.

Attribution biases stem from other sources as well. First, people have a tendency to engage in "defensive self-attribution." That is, we tend to attribute our successes to internal causes and our failures to external causes.<sup>208</sup> We make similar allowances for our friends: an increasing level of intimacy correlates with an increasing attribution of success to internal and failure to external causes.<sup>209</sup> Perhaps the strongest source of attribution bias is that referred to by Thomas Pettigrew as the "ultimate attribution error."<sup>210</sup> As various researchers have demonstrated, we tend to attribute desirable behaviors by ingroup members to internal, dispositional causes, whereas we readily attribute similar behaviors by outgroup members to environmental causes.

The causal attribution of success and failure plays a critical role in performance evaluation, promotion, compensation, and discharge decisions. Absent corrective measures, systematic biases in causal attribution can be expected to disadvantage members of stereotyped groups or individuals who are socially "distant" from the decisionmaker or who, for whatever reason, the decisionmaker has grouped in a different cognitive category. And as I explicate below,<sup>211</sup> these biases operate largely outside of the realm of decisionmaker self-awareness, and can be expected to contaminate interpersonal judgment even in the absence of any intent to discriminate.

##### 5. *The role of memory.*

In 1979, Myron Rothbart and his colleagues investigated the extent to which implicit expectancies about other people bias our memory for information that either confirms or disconfirms our expectations.<sup>212</sup> To do this, they presented subjects with identical sets of fifty behavioral descriptions of a target male individual. The items presented described actions that could be characterized as either "friendly," "unfriendly," "intelligent," "unintelligent," or unrelated to any of these four constructs. Half of the subjects were led to believe that the target was intellectual and half that the target was friendly. In each

---

207. See notes 237-245 *infra* and accompanying text.

208. See, e.g., Jones, *supra* note 146, at 64-67; Ross, *supra* note 194, at 337, 346-47. It should be noted that defensive self-attribution theory is not without its detractors. See generally D.T. Miller & M. Ross, *Self-Serving Biases in the Attribution of Causality: Fact or Fiction?*, 82 *PSYCHOL. BULL.* 213 (1975).

209. See Shelley E. Taylor & Judith Hall Koivumaki, *The Perception of Self and Others: Acquaintanceship, Affect, and Actor-Observer Differences*, 33 *J. OF PERSONALITY & SOC. PSYCHOL.* 403, 404, 407 (1976).

210. Thomas F. Pettigrew, *The Ultimate Attribution Error: Extending Allport's Cognitive Analysis of Prejudice*, 5 *PERSONALITY & SOC. PSYCHOL. BULL.* 461, 464-66 (1979).

211. See notes 233-247 *infra* and accompanying text.

212. Myron Rothbart, Mark Evans, & Solomon Fulero, *Recall for Confirming Events: Memory Processes and the Maintenance of Social Stereotypes*, 15 *J. EXPERIMENTAL SOC. PSYCHOL.* 343, 344 (1979).

group, half of the subjects were given this expectancy before being presented with the behavioral descriptions and half afterward.<sup>213</sup>

The investigators found that subjects recalled behaviors which confirmed their prior expectancies more readily than those which disconfirmed or were unrelated to them. However, these results obtained only when the expectancy was induced prior to the presentation of the behaviors; subjects who viewed the behaviors before any expectancy was created were far more evenhanded in their memory for confirming and disconfirming information. From these findings, Rothbart and his colleagues inferred that selective retrieval alone could not account for the observed memorial bias in favor of expectancy-confirming events. Whatever was causing the bias operated much earlier in the attention/encoding/storage/recall continuum.

Over the following years, these results were replicated,<sup>214</sup> and their implications refined,<sup>215</sup> by various researchers across a variety of experimental contexts. Taken as a whole, these findings indicate that people do not evenhandedly encode, store, or recall information about other people. Rather, once a target individual has been perceived as a member of a particular category, people are more likely to remember the target as exhibiting attributes and behaviors commonly associated with that category. Indeed, once we have developed stereotypic expectancies of a person, we even tend to "remember" stereotype-consistent behaviors that did not actually occur.<sup>216</sup>

213. *Id.* at 346-48.

214. See, e.g., Nancy Cantor & Walter Mischel, *Prototypicality and Personality: Effects on Free Recall and Personality Impressions*, 13 J. RES. IN PERSONALITY 187, 188-92 (1979); Claudia E. Cohen, *Person Categories and Social Perception: Testing Some Boundaries of the Processing Effects of Prior Knowledge*, 40 J. PERSONALITY & SOC. PSYCHOL. 441, 447 (1981) (observing that subjects more accurately remembered features of a target woman that were consistent with their prototype of a person of her occupation); Mark Snyder & Seymour W. Uranowitz, *Reconstructing the Past: Some Cognitive Consequences of Person Perception*, 36 J. PERSONALITY & SOC. PSYCHOL. 941, 941 (1978); see also Myron Rothbart, *Memory Processes and Social Beliefs*, in COGNITIVE PROCESSES IN STEREOTYPING AND INTERGROUP BEHAVIOR, *supra* note 39, at 145, 157-64.

215. Some researchers, notably Reid Hastie and Thomas K. Srull, find that in certain circumstances, expectancy-inconsistent information is retained at higher rates than either expectancy-confirming or expectancy-irrelevant information. See, e.g., Reid Hastie, *Memory for Behavioral Information that Confirms or Contradicts a Personality Impression*, in PERSON MEMORY: THE COGNITIVE BASIS OF SOCIAL PERCEPTION 155, 174-75 (Reid Hastie, Thomas M. Ostrom, Ebbe B. Ebbesen, Robert S. Wyer, Jr., David L. Hamilton, & Donal E. Carlston eds., 1980); Reid Hastie & Purohit Anand Kumar, *Person Memory: Personality Traits as Organizing Principles in Memory for Behaviors*, 37 J. PERSONALITY & SOC. PSYCHOL. 25, 26 (1979); Thomas K. Srull, *Person Memory: Some Tests of Associative Storage and Retrieval Models*, 7 J. EXPERIMENTAL PSYCHOL.: HUM. LEARNING AND MEMORY 440, 441 (1981). These results have, however, been explained in a manner consistent with findings such as those reviewed in note 198 *supra*. Specifically, the observed recall bias in favor of expectancy disconfirming events has been explained as resulting either from an application of the availability heuristic built into Hastie & Kumar's experimental design, or from the increased processing time accorded these events in the process of causal attribution. See Crocker, *supra* note 197, at 63 (attributing heightened recall for incongruent information to the extra processing time necessary to encode it); Gordon D. Hemsley & Harvey H.C. Marmurek, *Person Memory: The Processing of Consistent and Inconsistent Person Information*, 8 PERSONALITY & SOC. PSYCHOL. BULL. 433 (1982) (describing the effects of Hastie & Kumar's experimental design). Accord Reid Hastie, *Causes and Effects of Causal Attribution*, 46 J. PERSONALITY & SOC. PSYCHOL. 44, 53-55 (1984).

216. See, e.g., Nancy Cantor & Walter Mischel, *Traits as Prototypes: Effects on Recognition Memory*, 35 J. PERSONALITY & SOC. PSYCHOL. 38, 41-45 (1978) (subjects led to think of a target person

The fact that a decisionmaker might harbor a "phantom memory" about another person does not, as Justice Souter concluded in *St. Mary's Honor Center v. Hicks*,<sup>217</sup> turn him into a liar. In cognitive terms, the phenomenon is no different than phantom memories for prototype-consistent but nonpresented material found in experiments involving visual dot patterns, digit-letter series, and stories about everyday experiences.<sup>218</sup> Once tagged with a category label, a target person, like any other perceptual object, "comes more and more to resemble our expectations for persons of that 'type.'"<sup>219</sup> Apparently, one mechanism by which this occurs is that events that confirm schematic expectancies are favored in memory over those that do not. This preference cannot be attributed to simple response bias; it stems as well from biases in attention, encoding, and memory storage processes.

Under current disparate treatment jurisprudence, courts construe facts showing that stereotype-confirming information was overutilized in a decision-making process as evidencing conscious discriminatory intent. Schema theory, however, provides an alternative explanation. According to this view, schemas, prototypes, or stereotypes operate as implicit knowledge structures which organize input into, storage in, and recall from memory. People learn and remember information by actively categorizing or coding that information according to well-learned conceptual schemas. Once a schema is activated, subsequent confirming events are, in a sense, "tagged" with a category label when stored in memory. When a person later attempts to recall specific events, the category label functions as a retrieval cue. Items not encoded with the schematic "tag" (i.e. expectancy-inconsistent information) are stored in a more diffuse manner, making retrieval more difficult.<sup>220</sup> Furthermore, at least in certain circumstances, perceivers attend more closely to and thus preferentially encode expected as opposed to unexpected events.<sup>221</sup>

Such a bias for expected events distorts the "subjective data set" from which an employment decision is ultimately made such that, at the moment of decision, a biased decisionmaker can in good faith believe that he was fairly basing his decision on "the facts." Seen in this way, it becomes clear that stereotypes, person prototypes, and other implicit knowledge structures bias decisionmaking long before the "moment of decision" upon which *Price Waterhouse* and its progeny focus Title VII's adjudicative attention.<sup>222</sup>

---

as either "extrovert" or "introvert" exhibited enhanced recall of conceptually related but nonpresented items).

217. 113 S. Ct. 2742, 2763-64 (1993) (Souter, J., dissenting).

218. See, e.g., John D. Bransford & Jeffrey J. Franks, *The Abstraction of Linguistic Ideas*, 2 COGNITIVE PSYCHOL. 331, 349-50 (1971). But see Judith S. Reitman & George H. Bower, *Storage and Later Recognition of Exemplars of Concepts*, 4 COGNITIVE PSYCHOL. 194, 205 (1973).

219. Cantor & Mischel, *supra* note 214, at 5.

220. Regarding the theorized effects of schematic structures on memory encoding, storage and retrieval, and empirical data supporting these views, see, for example, Cantor & Mischel, *supra* note 216, at 39; Taylor & Crocker, *supra* note 124, at 98-101; Endel Tulving & Donald M. Thomson, *Encoding Specificity and Retrieval Processes in Episodic Memory*, 80 PSYCHOL. REV. 352 (1973).

221. McArthur, *supra* note 142, at 233-34.

222. See text accompanying notes 32-39.

6. *The effects of prior theories on search and retrieval strategies.*

All employment decisionmaking entails a form of hypothesis testing. To decide, for example, whether to promote a particular employee, a decisionmaker must first formulate an implicit judgment or belief about that employee on one or more dimensions ostensibly relevant to the decision in question. Such a hypothesis might be that the applicant lacks initiative and is thus unlikely to perform adequately under minimal supervision.

A tentative hypothesis of this sort may stem from any of a number of sources. It may result from one particularly vivid event involving the target employee, which, by virtue of its memorability, is subjectively viewed as highly predictive of future behavior.<sup>223</sup> It may stem from even superficial similarities between this employee and another whose performance was unsatisfactory.<sup>224</sup> It may stem from schematic expectancies, stereotypes, or other applications of the representativeness heuristic.<sup>225</sup> Another person's opinion about the target individual may suggest the hypothesis. But from whatever source, an implicit hypothesis, once formulated, can be expected to exert a profoundly biasing effect on judgment and decisionmaking.

Now compare Title VII's model of social inference. The structure of proof under Title VII's pretext theory assumes that, in formulating judgments about other people, an employer will first implicitly comprehend a proposition ("Miguel lacks initiative") and then freely decide whether to believe it by testing it evenhandedly against all reasonably available, relevant data. This view boasts a long tradition, beginning with Descartes, in Western philosophy and psychology.<sup>226</sup> However, empirical investigation supports an alternative view,<sup>227</sup> originating with Spinoza, that people initially accept as true every proposition they comprehend and then go about deciding whether to "unbelieve" it. If people are mentally preoccupied or otherwise unmotivated or unable to engage in further processing, they may never move beyond their initial proposition.<sup>228</sup> Thus, people may experience particular difficulty in rejecting false propositions. The "intuitive scientist" approaches the task of evaluating a tentative hypothesis by building a case to support it rather than by seeking to disconfirm it.

Accordingly, when asked to assess the validity of a hypothesis, subjects systematically search available information for theory-confirming, at the ex-

223. See Tversky & Kahneman, *supra* note 148, at 207.

224. See Stephen J. Read, *Once Is Enough: Causal Reasoning From a Single Instance*, 45 J. OF PERSONALITY & SOC. PSYCHOL. 323, 323 (1983) (finding that prediction and explanation are often based on comparison to a single "representative" event). On representativeness, see Kahneman & Tversky, *supra* note 176.

225. See text accompanying note 177 *supra*.

226. Daniel T. Gilbert, *The Assent of Man: Mental Representation and the Control of Belief*, in HANDBOOK OF MENTAL CONTROL 57, 60 (Daniel M. Wegner & James W. Pennebaker eds., 1993) [hereinafter *The Assent of Man*]; Daniel T. Gilbert, *How Mental Systems Believe*, 46 AM. PSYCHOL. 107, 108 (1991) [hereinafter *Mental Systems*].

227. *The Assent of Man*, *supra* note 226, at 108; *Mental Systems*, *supra* note 226, at 61-63.

228. For a review of the literature supporting this view, see Mark Snyder, *Seek, And Ye Shall Find: Testing Hypotheses About Other People*, in SOCIAL COGNITION, *supra* note 123, at 277; for a review of more recent material, see *Mental Systems*, *supra* note 226.

pense of theory-disconfirming, data.<sup>229</sup> They not only evaluate theory-confirming evidence as being more relevant than theory-disconfirming evidence but also have better recall for the former than the latter. Indeed, people have a good deal of difficulty in searching for and recognizing theory-disconfirming information.<sup>230</sup> They accept theory-confirming evidence at face value, whereas they subject theory-disconfirming evidence to critical evaluation.<sup>231</sup> Unsurprisingly, this results in a rather powerful bias towards confirming tentative judgments or beliefs. This enormous quantity of empirical evidence suggests that Title VII's assumption of a blank slate from which employers make decisions is wholly unupportable.

#### H. *Rethinking the Assumptions Underlying Disparate Treatment Jurisprudence*

The assumptions underlying Title VII's disparate treatment theory have been so substantially undermined by social cognition theory that they can no longer be considered valid. As we have seen, disparate treatment does not necessarily manifest discriminatory motive or intent. Even among the well-intentioned, social schemas such as stereotypes, acting in concert with a variety of other judgment heuristics, can be expected to bias intergroup perception and judgment. Stereotyping, while it may in some contexts be socially undesirable or otherwise maladaptive, is neither aberrant nor indicative of discriminatory motivation. Stereotypes are simply a subset of the vast array of categorical structures, expectancies and heuristics that characterize, and indeed make possible, human cognitive functioning.

The notion that decisionmaking is somehow separate from the perceptive, interpretive, and memorial processes that precede it is utterly fallacious. These various processes comprise a functional continuum which is vulnerable to distortion at every point. Thus, discrimination is not necessarily something that occurs "at the moment of decision." Rather, it can intrude much earlier, as cognitive process-based errors in perception and judgment subtly distort the ostensibly objective data set upon which a decision is ultimately based. We must, therefore, reconsider the very foundations upon which current Title VII doctrine has developed.

---

229. Mark Snyder & William B. Swann, Jr., *Hypothesis Testing Processes in Social Interaction*, 36 J. OF PERSONALITY & SOC. PSYCHOL. 1202, 1202 (1978). Interestingly, neither the origins of the hypotheses being tested nor the structure of incentives for accuracy exerted an observable influence on whether subjects adopted a theory-confirming or theory-disconfirming strategy.

230. Hillel J. Einhorn & Robin M. Hogarth, *Confidence in Judgment: Persistence of the Illusion of Validity*, 85 PSYCHOL. REV. 395, 397 (1978).

231. Charles G. Lord, Lee Ross, & Mark R. Lepper, *Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence*, 37 J. PERSONALITY & SOC. PSYCHOL. 2098, 2106 (1979).

1. *Cognitive bias and the assumption of rational decisionmaking.*

At least since *Furnco*,<sup>232</sup> Title VII's disparate treatment theory has been premised on the assumption that, absent discriminatory motive or intent, employment decisionmakers can be expected to follow normative rules of inference and judgment. This is, in actuality, no small order. *Furnco*'s "rational businessman," when called upon to make a decision about a particular applicant or employee, must perform a complex series of inferential tasks.

First, he formulates a tentative hypothesis or belief about the applicant or employee on one or more dimensions relevant to the decision in question. Then he proceeds to test it against all reasonably available relevant data.

To do this, he retrieves from memory, or from the record before him, all decision-relevant traits or events which might logically confirm or disconfirm his tentative hypothesis. Next, he assesses the validity of each trait or event as a predictor of future job performance. In a sense, he subjectively calculates the expected utility of each piece of data to his ultimate decision. *Furnco*'s rational decisionmaker then weighs the quantum of confirming evidence against the quantum of disconfirming evidence, and makes a final decision to accept or reject the tentative hypothesis.

To perform these various tasks in accordance with normative rules of inference, our rational employer must be able to attend to, interpret, remember, recall, and utilize all relevant information evenhandedly. He must accurately assess the relationships between a wide array of events and future outcomes, and evenhandedly attribute the causes of an employee's success or failure. In other words, *Furnco*'s rational decisionmaker must be a " 'systematic' information processor."<sup>233</sup>

But real decisionmakers, even those of good intention, fall far short of *Furnco*'s presupposed ideal. As we have already seen, implicit knowledge structures and judgmental heuristics systematically bias perception and judgment at all points along the perceptual/judgmental continuum. Negative behaviors by minority group members are more salient and thus more readily remembered than negative behaviors by members of the majority. Stereotype-consistent behaviors are better remembered and considered more predictive of future behavior than stereotype-inconsistent behaviors. Indeed, stereotypic expectancies can go so far as to determine as a threshold matter whether a particular behavior is interpreted and subsequently encoded as positive, negative, or neutral. Finally, stereotypes, like any other implicit expectancy, can be expected to charge a decisionmaker's ostensibly objective review of available evidence with a powerful theory-confirming valence. The past thirty years of research in cognitive social psychology tell us that people are anything but "systematic information processors." Instead it is much more plausible to as-

---

232. *Furnco Constr. Corp. v. Waters*, 438 U.S. 567 (1978). See text accompanying notes 80-81 *supra*.

233. Shelley Chaiken, Akiva Liberman, & Alice H. Eagly, *Heuristic and Systematic Information Processing Within and Beyond the Persuasion Context*, in UNINTENDED THOUGHT 212, 212 (James S. Uleman & John A. Bargh eds., 1989).

sume that, in a large proportion of employment discrimination disputes, the following two propositions are true: first, that the nondiscriminatory reasons proffered for the disputed decision were the decisionmaker's "real reasons" and not a pretext for discrimination; and second, that the decision resulted from intergroup bias. The "real reason"/"phony reason" dichotomy built into pretext analysis does not accurately reflect the process by which cognitive sources of bias distort intergroup perception and judgment.

### 2. *Cognitive bias and the bifurcation of perception and judgment.*

Just as social cognition theory undermines the assumption that well-intentioned employers will be systematic information processors, it also subverts the notion that discrimination necessarily occurs at the moment of decision. Cognitive sources of intergroup bias corrupt decisionmaking not at the moment of decision, but long before it, by distorting the interpretive framework through which decisions are made. Social cognition theory teaches us that interpersonal decisionmaking is an integrated system comprising perception, interpretation, attribution, memory, and judgment. Decisionmaking is not, as the "moment of decision" fallacy assumes, structurally disjoined from those perceptual and inferential processes which comprise it.

Current disparate treatment jurisprudence, which focuses on decisionmaker intent at the moment of decision, is ill-equipped to deal with bias of this sort. An excellent illustration can be found in the Sixth Circuit's opinion in *Wilson v. Stroh Co.*<sup>234</sup> At trial, plaintiff Wilson had successfully argued that his immediate supervisor had been motivated by racial animus in reporting Wilson's behavioral transgression to upper management.<sup>235</sup> The trial court and the Sixth Circuit on appeal, assuming the truth of plaintiff's initial argument, based its decision against the plaintiff on the observation that members of upper management had conducted their own investigation of the transgression, and had made an independent determination that the discharge was warranted. Because the plaintiff could not show that the ultimate decisionmaker was racially motivated "at the moment of decision," the court treated the supervisor's established bias as irrelevant. Nowhere in the opinion does the court consider the possibility that the supervisor's discriminatory animus, combined with management's bias against disconfirming his initial hypothesis, constituted a form of discrimination actionable under Title VII.

### 3. *Cognitive bias and the assumption of decisionmaker self-awareness.*

As the plurality opinion in *Price Waterhouse* makes clear,<sup>236</sup> current disparate treatment caselaw reflects a belief that employment decisionmakers know when they are discriminating on a prohibited basis and when they are not. More fundamentally, the normative efficacy of Title VII's rule against disparate treatment discrimination rests entirely on the assumption that deci-

---

234. 952 F. 2d 942 (6th Cir. 1992).

235. *Id.* at 944 n.3.

236. See text accompanying notes 84-102 *supra*.

sionmakers know how particular stimuli influence their judgments and decisions. Far from supporting the assumption of decisionmaker self-awareness which disparate treatment theory takes for granted, a persuasive body of empirical and theoretical research suggests that, in large measure, people lack access to the mental processes involved in evaluation and judgment, and are quite poor at accurately attributing the causes of their actions and decisions.

For example, as Shelley Taylor and his colleagues demonstrated,<sup>237</sup> subjects remained unaware that they were using race and gender as structures for organizing social information in memory, even though the experimental results clearly demonstrated that they were doing so. This result should come as no great surprise, however. There is virtually no evidence supporting the proposition that people possess conscious awareness of higher level perceptual and memorial processes. This lack of awareness includes, of course, a lack of awareness about the perceptual and memory-biasing effects of schemas, stereotypes, and person prototypes examined earlier.

Along these lines, psychologists Laraine Winter and James Uleman have demonstrated that when we observe an actor behaving in a particular way, we do not simply store a representation of the event itself. Rather, along with information about the event, we automatically and unconsciously encode dispositional judgments about the actor.<sup>238</sup> These dispositional attributions, often supplied by stereotypes and other forms of schematic expectancies, may then serve to bias future judgments about the individual in question through the operation of a "halo effect,"<sup>239</sup> which also functions outside of decisionmaker self-awareness.

In a groundbreaking pair of articles published in the late 1970s,<sup>240</sup> Richard Nisbett and Timothy Wilson demonstrated that people are actually quite poor at identifying the effects of various stimuli on their evaluations, judgments, choices, and predictions.<sup>241</sup> In a series of experiments, Nisbett and Wilson

237. Taylor, *supra* note 141.

238. Laraine Winter & James S. Uleman, *When Are Social Judgments Made? Evidence for the Spontaneousness of Trait Inferences*, 47 J. PERSONALITY & SOC. PSYCHOL. 237 (1984). *Accord* Laraine Winter, James S. Uleman, & Cathryn Cunniff, *How Automatic Are Social Judgments?*, 49 J. PERSONALITY & SOC. PSYCHOL. 904 (1985) (evidencing how experimental data indicated that dispositional inferences are made at encoding, without intention and with little awareness); Leonard S. Newman & James S. Uleman, *Spontaneous Trait Inference*, in UNINTENDED THOUGHT, *supra* note 233, at 155.

239. See Richard E. Nisbett & Timothy DeCamp Wilson, *The Halo Effect: Evidence for Unconscious Alteration of Judgments*, 35 J. PERSONALITY & SOC. PSYCHOL. 250 (1977) (describing how global evaluations of a person induced altered evaluations of the person's specific attributes, even when there existed sufficient information to permit independent assessment of those attributes and where subjects were unaware of the influence of the global evaluations on their ratings of specific attributes).

240. Richard E. Nisbett & Timothy DeCamp Wilson, *Telling More Than We Can Know: Verbal Reports on Mental Processes*, 84 PSYCHOL. REV. 231, 233-48 (1977); Timothy DeCamp Wilson & Richard E. Nisbett, *The Accuracy of Verbal Reports About the Effects of Stimuli on Evaluations and Behavior*, 41 SOC. PSYCHOL. 118, 119 (1978).

241. The experiments were designed so as to minimize possibilities of self-report bias and maximize the applicability of their findings to real life situations. Thus, the processes studied were representative of those routinely occurring in daily life. Care was taken to assure that subjects were aware of the existence of both the critical stimuli and their own responses and choices. Most of the stimulus situations were designed to minimize self-report biases generated by social desirability or ego-maintenance effects. See Nisbett & Wilson, *supra* note 240, at 242.

systematically manipulated some component of a complex stimulus situation such that the impact of a particular stimulus component on subjects' choices and actions could be assessed. They found that subjects were virtually never accurate in identifying the causal efficacy of the manipulated stimulus. Where the stimulus had a significant effect on their responses, subjects typically reported that it was noninfluential; in cases where a particular stimulus component had no significant causal effect, subjects reported it as having determined their response. When it comes to identifying the reasons why we make particular decisions, concluded Nisbett and Wilson, people tend towards "telling more than we can know."<sup>242</sup>

Also significant was Nisbett and Wilson's finding that subjects' erroneous self-attributions were neither random nor haphazard; they closely matched a second group of subjects' predictions about why the actors they were observing might have responded as they did. Thus, when subjects reported on their cognitive process leading to a particular response, they did not appear to do so on the basis of actual introspection. Rather, their self-attributions appear to have been based on the same plausible, prior causal theories on which observers relied in predicting why people in certain situations might act in a particular way or make a particular choice.<sup>243</sup>

Nisbett and Wilson interpreted these findings as indicating that in making causal self-attributions, people rely largely on the representativeness heuristic. Specifically, they argue that when people are asked to attribute the causes of actions or choices, they do not actually examine their cognitive processes, but select from memory a causal theory which seems to them to be a plausible explanation for what they have done.<sup>244</sup>

If Nisbett and Wilson are even partially correct,<sup>245</sup> the implications for disparate treatment compliance and dispute adjudication policy are profound. It will be the rare employer indeed who can *accurately* identify the reasons why he hired or promoted one employee over another, fired another, or set salary increases as he did. Moreover, in the vast majority of cases, even after extensive and costly discovery, it will be impossible as an empirical matter for either party to prove whether or not a particular proffered causal theory in fact motivated an employer. Unlike Wilson and Nisbett's laboratory, in the real world one simply cannot control the multiplicity of possible causal antecedents so as to determine the causal efficacy of race, gender, national origin, or age. Thus, in the final analysis, each disparate treatment adjudication turns on the subjective attribution of causation by a trier of fact, a process prone to all the distortions in attributional judgment explored earlier in this section.

---

242. See note 251 *infra*.

243. Nisbett & Wilson, *supra* note 240, at 247.

244. *Id.* at 233, 248.

245. Nisbett and Wilson's 1977 article generated a fair amount of controversy. See generally Timothy D. Wilson & Julie I. Stone, *Limitations on Self-Knowledge: More on Telling More Than We Can Know*, in *SELF, SITUATIONS, & SOCIAL BEHAVIOR: REVIEW OF PERSONALITY AND SOCIAL PSYCHOLOGY* 167 (Philip Shaver ed., 1985) (reviewing and responding to a variety of methodological and theoretical challenges to the Nisbett & Wilson article).

In fact, disparate treatment theory's central premise, that discrimination results from the conscious application of invidious intent, can be viewed as an expression of one particular attributional bias. John Stuart Mill originally termed this bias the "resemblance criterion."<sup>246</sup> Myron Rothbart and Solomon Fulero refer to it as the "profound motive fallacy."<sup>247</sup> This bias leads us to assume that the conditions or causes of an event will resemble the effect of the event or the event itself. Thus, as Rothbart and Fulero observe, people have a strong tendency to attribute emotionally relevant events to emotionally relevant causes; a profoundly negative event tends to be attributed to profoundly negative causes. Accordingly, to the extent that we see employment discrimination as a profoundly negative phenomenon, we will tend to attribute it to profoundly negative causes, such as the morally reprehensible action of an invidiously motivated decisionmaker.

#### 4. *Rethinking the role of motivation in discrimination and equal employment opportunity.*

Viewed through the lens of social cognition theory, it appears that current disparate treatment jurisprudence construes the role of motivation in intergroup discrimination precisely backwards. A substantial body of empirical and theoretical work supports the proposition that cognitive biases in social judgment operate automatically and must be controlled, if at all, through subsequent "mental correction."<sup>248</sup> Intergroup discrimination, or at least that variant which results from cognitive sources of bias, is automatic.<sup>249</sup> It does not result from a motive or intent to discriminate; it is an unwelcome byproduct of otherwise adaptive cognitive processes. But, like many unwanted byproducts, it can be controlled, sometimes even eliminated, through careful process re-engineering.<sup>250</sup> Cognitive biases in intergroup perception and judgment, though unin-

---

246. See RICHARD NISBETT & LEE ROSS, *HUMAN INFERENCE: STRATEGIES AND SHORTCOMINGS OF SOCIAL JUDGMENT* 115-16 (1980) (discussing JOHN STUART MILL, *A SYSTEM OF LOGIC RATIOCINATIVE AND INDUCTIVE* (1848)).

247. Myron Rothbart & Solomon Fulero, *Attribution of Causality for Important Events: The Profound Motive Fallacy* 1 (unpublished manuscript, on file with the *Stanford Law Review*). Rothbart and Fulero had subjects read stories in which a protagonist engaged in behavior that had either mild, moderate or severely negative outcomes, and then had subjects attribute the causes of those outcomes. Even after controlling for other factors, the attribution of profound motives to the protagonist was positively correlated with the severity of the outcome: the more severe the predictable outcome, the more profound the motive attributed to the protagonist.

248. According to this view, social judgment consists of two sequential phases: an unconscious, largely automatic process of dispositional inference referred to as "characterization," followed by situational adjustment or "correction," a deliberate form of conscious reason that occurs only with effort. For a thorough exposition of this approach, see Timothy D. Wilson & Nancy Brekke, *Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations*, 116 *PSYCHOL. BULL.* 117 (1994); Daniel T. Gilbert, *Thinking Lightly About Others: Automatic Components of the Social Influence Process*, in *UNINTENDED THOUGHT*, *supra* note 233, at 189.

249. For additional support for the view that cognitive biases in social judgment operate automatically without intention or awareness, see John A. Barth, *Conditional Automaticity: Varieties of Automatic Influence in Social Perception and Cognition*, in *UNINTENDED THOUGHT*, *supra* note 233, at 3; Newman & Uleman, *supra* note 238, at 155.

250. See Wilson & Brekke, *supra* note 248, at 130-35 (discussing difficulties inherent in and potential techniques for correcting "mental contamination").

tentional and largely unconscious, can be recognized and prevented by a decisionmaker who is motivated not to discriminate and who is provided with the tools required to translate that motivation into action. Seen in this way, disparate treatment does not necessarily manifest discriminatory motive or intent, but a motive or intent not to discriminate must be present to prevent it.

Given this psychologically more accurate view of discrimination, what would a reformulated definition of nondiscrimination look like? To say that a decisionmaker lacks discriminatory motivation is not to say that his perceptions and judgments are unaffected by cognitive sources of intergroup bias. Indeed, we should expect that a self-professed "colorblind" decisionmaker will fall prey to the various sources of cognitive bias we have examined. For even if this decisionmaker's conscious inferential process is colorblind, the categorical structures through which he collects, sorts, and recalls information are not. In a culture in which race, gender, and ethnicity are salient, even the well-intentioned will inexorably categorize along racial, gender, and ethnic lines. And once these categorical structures are in place, they can be expected to distort social perception and judgment. Our decisionmaker is not colorblind; he is simply "color-clueless," likely unaware that his perceptions, judgements, and decisions are being distorted by cognitive sources of intergroup bias. He, like the plant manager in my box manufacturing case, would be genuinely shocked and profoundly offended if accused of discrimination, especially under a legal and popular construction which equates discrimination with invidious intention. And his plaintiff counterpart, experiencing disparities in treatment but unable to prove discriminatory intent, would, like my client Miguel, be left embittered and disillusioned by his encounter with the civil rights enforcement process.

### III. IMPLICATIONS OF THE COGNITIVE APPROACH FOR EQUAL EMPLOYMENT OPPORTUNITY LAW AND POLICY

What implications derive from rejecting the assumptions currently undergirding disparate treatment theory and accepting the proposition that a broad class of discriminatory employment decisions result not from discriminatory motivation, but from normal cognitive processes and strategies that tend to bias intergroup perception and judgment? The overwhelming conclusion is that there now exists a fundamental "lack of fit" between the jurisprudential construction of discrimination and the actual phenomenon it purports to represent. This lack of fit has created a number of serious theoretical and practical problems. First, it is responsible for the increasing proliferation and deepening theoretical incoherence of Title VII's various models of liability. Second, this incoherence can be expected to decrease the validity of disparate treatment adjudications, increase adjudication-related information costs, and discourage the voluntary settlement of employment discrimination cases. Third, and perhaps most perniciously, it may be exacerbating rather than reducing intergroup tensions. In each of the next three sections, I tackle these three main problems: the incoherence of current liability theories, the effect of this incoherence on dispute resolution, and its effect on intergroup relations.

### A. *The Increasing Incoherence of Title VII Liability Theories*

In *The Structure of Scientific Revolutions*,<sup>251</sup> Thomas Kuhn observes that the breakdown of a theoretical paradigm often follows the proliferation of ad hoc adjustments designed to explain, within the existing theoretical structure, phenomena for which the paradigm could not otherwise account. As this disintegration progresses, Kuhn observes, the paradigm becomes so increasingly complex, so incapable of consistent application, that it eventually loses its utility as a guiding framework.<sup>252</sup>

Kuhn's model aptly describes the development of Title VII's disparate treatment paradigm. Over the past decade or so, courts have made ad hoc adjustments to Title VII liability theory to address three important employment discrimination problems for which the original pretext model of discrimination proof elaborated in *McDonnell Douglas* and *Burdine* had proven inadequate. In the first class of cases, a plaintiff proved that his group status had played a role in a challenged employment decision, but failed to prove, as pretext theory required, that the nondiscriminatory reasons proffered by the defendant had played no role at all. In such cases, pretext theory required judgment in favor of the defendant, despite the fact that plaintiff's group status had been found to have affected the employer's decisionmaking.

In the second class of cases, an employer's proffered nondiscriminatory reasons were proven not to have motivated its decision, but the trier of fact remained unconvinced that the employer had been motivated by discriminatory intent. Pretext theory, with its rigid dichotomization of "real reasons" and "cover-ups" for discrimination, seemed to dictate judgment for the plaintiff in such cases, despite the fact that the plaintiff had failed to prove discriminatory intent to the trier of fact's satisfaction.

Third, courts increasingly encountered cases in which an employer's overall decisionmaking process consistently resulted in the statistical underutilization of qualified members of a particular ethnic, gender, or racial group, though no systematic intent to discriminate could be proven. Under systemic disparate treatment theory, statistical disparities, without more, were deemed insufficient to prove the required discriminatory intent.

The manifest injustice of results such as these functioned much like the experimental anomalies Kuhn described,<sup>253</sup> calling into question the disparate treatment paradigm's utility as a guiding framework. While at least one quasi-judicial decisionmaker reacted by rejecting pretext theory outright,<sup>254</sup> the far more prevalent response was as Kuhn's model predicts. Courts did not reject the intent-based, pretext model of disparate treatment discrimination. Instead, they devised ad hoc adjustments to disparate treatment and disparate impact theory to "resolve" the observed anomalies.

---

251. THOMAS S. KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* (2d ed., 1970).

252. *Id.* at 52-65.

253. *Id.* at 52-53.

254. *See* California Dep't of Fair Employment and Hous. v. Church's Fried Chicken, FEHC No. 90-11 (1990).

For example, in response to the first anomaly described above, various circuit courts and later a sharply fragmented Supreme Court in *Price Waterhouse*, added a mixed-motives variant to disparate treatment theory. However, neither the *Price Waterhouse* Court, nor the district and circuit courts which applied it, have been able to devise workable standards for delineating the respective spheres of the pretext and mixed-motives variants of disparate treatment proof. Indeed, under the current majority approach to this problem, neither the parties, the court, nor the jurors are apt to know which of the two competing variants will apply to determine liability until, at the earliest, the conclusion of the plaintiff's case.

The second anomaly gave rise to what became known as the "Pretext-Plus" controversy,<sup>255</sup> which culminated in the Court's bitterly divided decision in *St. Mary's Honor Center v. Hicks*.<sup>256</sup> Over a scathing dissent authored by Justice Souter, the *Hicks* majority held that the trier of fact's rejection of an employer's asserted reasons for its actions may be sufficient to justify a plaintiff's verdict, but it does not entitle a plaintiff to judgment as a matter of law.<sup>257</sup> This result, while it may have corrected one anomaly, undermined the logic supporting the pretext model of proof.

In response to the third problem, circuit courts divided for many years on the question of whether disparate impact theory could be used to challenge subjective decisionmaking systems.<sup>258</sup> And although in *Watson v. Fort Worth Bank and Trust*<sup>259</sup> and in *Wards Cove Packing Co. v. Atonio*,<sup>260</sup> the Supreme Court held that disparate impact theory could be used in this way, the Court was so troubled by how this might be accomplished without undermining competing values and interests that they fundamentally altered the burdens of proof and defense in disparate impact cases. When they were done, the disparate treatment paradigm had all but swallowed the disparate impact model. And while Congress in a sense overturned *Wards Cove* with the Civil Rights Act of 1991,<sup>261</sup> it simply ignored the grave practical and analytical problems that had so troubled the *Watson* and *Wards Cove* Courts. The Civil Rights Act of 1991 does not resolve the many practical and theoretical problems associated with applying disparate impact theory in subjective practices cases; it simply ignores

---

255. See, e.g., Catherine J. Lancot, *The Defendant Lies and the Plaintiff Loses: The Fallacy of the "Pretext-Plus" Rule in Employment Discrimination Cases*, 43 *HASTINGS L.J.* 57, 71-90 (1991) (describing split among several circuits in applying the "pretext-plus" rule).

256. 113 S. Ct. 2742 (1993).

257. *Id.* at 2749.

258. See, e.g., *Muller v. United States Steel Corp.*, 509 F.2d 923, 925 (10th Cir.), *cert. denied*, 423 U.S. 825 (1975) (applying disparate impact theory to foreman's uncontrolled discretion to allocate on-the-job training); *Pettway v. American Cast Iron Pipe Co.*, 494 F.2d 211, 241 (5th Cir. 1974), *cert. denied*, 439 U.S. 1115 (1979) (applying disparate impact theory to subjective evaluations by all-white superintendents); *Brown v. Gaston County Dyeing Mach. Co.*, 457 F.2d 1377, 1383 (4th Cir. 1972) (applying disparate impact theory to employer's practice of announcing job openings only to current employees); *Rowe v. General Motors Corp.*, 457 F.2d 348, 353 (5th Cir. 1972) (applying disparate impact theory to foreman's recommendations based on employee "ability, merit and capacity").

259. 487 U.S. 977 (1988).

260. 490 U.S. 642 (1989).

261. 42 U.S.C. § 2000e (1994).

them. As a result, a substantial segment of Title VII jurisprudence now stands in a state of practical uncertainty and analytical disarray.

In each of the next three subsections, I will explore the courts' treatment of these three anomalies and describe the serious theoretical and practical problems their responses now present.

1. *The mixed-motives and pretext variants of disparate treatment proof: which model for which case?*

In *Price Waterhouse v. Hopkins*,<sup>262</sup> the Supreme Court held for the first time that a disparate treatment plaintiff could prevail even if she were unable to disprove the motivating significance of every legitimate, nondiscriminatory reason proffered by the defendant. *Price Waterhouse* thus seemed to stand for the proposition that a plaintiff could shift the burden of proof to the defendant by showing simply that her group status "played a role" in the decision or action taken against her.<sup>263</sup> She did not have to prove that it was the sole reason, nor did she have to establish that the reasons proffered by the defendant were "cover-ups" for a real discriminatory reason.<sup>264</sup> Thus, it appeared, plaintiffs no longer bore the heavy burden associated with pretext proof; they could, if they chose, plead mixed-motives discrimination and avoid them.

However, as one commentator described,<sup>265</sup> courts have floundered in their efforts to define the respective spheres of application of the pretext and mixed-motives theories of liability, creating a theoretical and practical morass.<sup>266</sup> In what may well be a jurisprudentially unprecedented situation, most federal courts now determine whether pretext or mixed-motives theory will apply to a given case based on the type of evidence a plaintiff proffers. Presently, five or six of the circuits hold that a plaintiff can use the mixed-motives theory of liability only if she proves by "direct evidence" that her protected status was a motivating factor.<sup>267</sup> Three other circuits<sup>268</sup> permit a plaintiff to use a mixed-motives theory only if she proffers evidence "directly tied to" or "directly re-

262. 490 U.S. 228 (1989).

263. *Id.* at 244. The *Price Waterhouse* rule was modified by the Civil Rights Act of 1991. Under the Act, liability is now established upon a plaintiff's showing that his or her protected group status was a motivating factor in an employer's decision. 42 U.S.C. § 2000e-2(m) (1994). However, a prevailing "mixed motives" plaintiff may still be denied certain individualized relief if the employer is able to prove that he or she would have reached the same decision in the absence of the impermissible motivating factor. *Id.* at § 2000e-5(g)(2)(B).

264. *Price Waterhouse*, 490 U.S. at 246.

265. Michael A. Zubrensky, *Despite the Smoke, There Is No Gun: Direct Evidence Requirements in Mixed-Motives Employment Law After Price Waterhouse v. Hopkins*, 46 STAN. L. REV. 959, 960-69 (1994) (surveying various federal court interpretations of plaintiff's burden of proof).

266. *Id.* at 970-80 (presenting the array of evidentiary standards that have been applied).

267. *See* Heim v. Utah, 8 F.3d 1541, 1547 (10th Cir. 1993) (plaintiff failed to show direct evidence of discriminating intent); Brown v. East Mississippi Elec. Power Ass'n, 989 F.2d 858, 861 (5th Cir. 1993) (plaintiff successfully showed by direct evidence that race was a factor in a contested employment action); EEOC v. Alton Packaging Corp., 901 F.2d 920, 923 (11th Cir. 1990) (plaintiff showed direct evidence of social animus); Jackson v. Harvard Univ., 900 F.2d 464, 467 (1st Cir. 1990) (plaintiff failed to show direct evidence of gender discrimination in tenure decision); Gagne v. Northwestern Nat'l Ins. Co., 881 F.2d 309, 315 (6th Cir. 1989) (plaintiff failed to proffer direct evidence of age discrimination in her dismissal); Randle v. LaSalle Telecom., Inc., 876 F.2d 563, 569 (7th Cir. 1989) (plaintiff failed to proffer direct evidence of racial animus). The Seventh Circuit presently suffers from an in-

flecting" the alleged discrimination. Absent such evidence, a plaintiff's only recourse is to the pretext model of proof, under which she must disprove the motivating significance of each of the reasons proffered by the defendant and show that her group status was the "sole motive" underlying the defendant's decision.<sup>269</sup>

That the essential elements of proof and defense in a disparate treatment case should vary depending on the type of evidence proffered by the plaintiff would be perplexing enough. To make matters worse, no one agrees on the meanings of "direct evidence," or "evidence directly tied to discrimination."

Consider, for example, the Fifth Circuit's decision in *Brown v. East Mississippi Electric Power Ass'n*,<sup>270</sup> a race discrimination case in which the plaintiff proffered evidence showing that the decisionmaker had routinely used racial slurs in referring to African Americans, including the plaintiff. Brown attempted to proceed under a mixed-motives theory, but was not permitted to do so by the district court because he had not provided "direct evidence" of discrimination. Unable to prove that each of the reasons for his discharge proffered by his employer was pretextual, Brown lost at trial. The Fifth Circuit reversed, defining "direct evidence" in a relatively straightforward manner: "Direct evidence is evidence which, if believed, proves the fact without inference or presumption."<sup>271</sup> The Court concluded that the supervisor's use of racial slurs constitute "direct evidence that racial animus was a motivating factor in the contested disciplinary decisions."<sup>272</sup>

This is nonsensical. Without inference or presumption, one cannot logically conclude from an employer's use of racial slurs that he was motivated by race-based animus in making a particular decision. It may be a reasonable inference to make, but it is an inference nonetheless.

Every circuit applying a direct evidence test in determining whether a plaintiff may proceed under a mixed-motives theory defines direct evidence as the

---

tracircuit conflict on this issue. Compare *id.* (applying direct evidence rule) with *Robinson v. PPG Indus., Inc.*, 23 F.3d 1159, 1165 (7th Cir. 1994) (arguably loosening the direct evidence requirement).

The "direct evidence" requirement derives from Justice O'Connor's *Price Waterhouse* concurrence, in which she stated that "in order to justify shifting the burden on the issue of causation to the defendant, a . . . plaintiff must show by *direct evidence* that an illegitimate criterion was a substantial factor in the decision." *Price Waterhouse*, 490 U.S. at 276 (O'Connor, J., concurring) (emphasis added). No other justice expressed such a view.

268. See *Brown v. Polk County*, 37 F.3d 404, 410-11 (8th Cir. 1994) (plaintiff failed to produce evidence that directly reflects religious or racial basis for termination); *Griffiths v. Cigna Corp.*, 988 F.2d 457 (3rd Cir. 1993) (finding plaintiff presented no evidence that directly reflects a retaliatory motive for plaintiff's termination); *Ostrowski v. Atlantic Mut. Ins. Cos.*, 968 F.2d 171 (2d Cir. 1992) (holding plaintiff entitled to burden-shifting jury instruction where evidence reflecting a discriminatory attitude present). Only the Fourth Circuit imposes no evidence-based restrictions on plaintiff's attempting to proceed under a mixed motives theory. *White v. Federal Express Corp.*, 939 F.2d 157, 160 (4th Cir. 1991) ("plaintiff may carry its burden under ordinary principles of proof by any sufficiently probative direct or indirect evidence"). As of the writing of this article, the Ninth Circuit had not ruled on the issue.

269. See notes 55-82 and accompanying text *supra*.

270. 989 F.2d 858 (5th Cir. 1993).

271. *Id.* at 861.

272. *Id.*

Court did in *Brown*.<sup>273</sup> But none of these cases, whether decided in favor of the plaintiff or the defendant, actually involved “direct evidence” as the court had defined it. From such a logically incoherent analysis, it is impossible to derive a comprehensible standard or to predict what a court might do in any future case.

The situation is perhaps worse in circuits applying the “directly-tied-to” test. First, there are almost as many interpretations of the standard as there are cases applying it. One court’s “stray remarks in the workplace” are another court’s “evidence directly reflecting” discriminatory animus.<sup>274</sup> Reasonable people can differ far too easily on whether a particular piece of circumstantial evidence reflects discriminatory attitudes with a directness sufficient to trigger mixed-motives analysis. Given the substantial strategic advantage at stake, we can expect Title VII litigants to delay settlement and to invest significant time and financial resources into discovery, motion practice, and post-judgment appeals calculated to improve their position on this important issue.<sup>275</sup>

To make matters worse, the applicability of mixed-motives analysis is not necessarily amenable to resolution by the court as a matter of law. Determining whether a particular piece of evidence “directly reflects” discriminatory animus involves questions of credibility, contextualized interpretation, and inference. These are matters for a jury, not a court. Thus, as the Seventh Circuit observed, one cannot necessarily determine whether mixed-motives analysis is appropriate until after the trier of fact has decided what inferences to draw from the plaintiff’s evidence.<sup>276</sup> The result, as Justice Kennedy predicted in his dissent in *Price Waterhouse*, is apt to be chaos in the formulation of comprehensible jury instructions.<sup>277</sup>

273. See note 267 *supra*.

274. Compare cases discussed in text accompanying note 267 *supra* with cases cited in note 268 *supra*.

275. See note 384 *infra*.

276. In deciding to remand the case to the district court for jury trial, the court in *Robinson v. PPG Indus., Inc.*, 23 F.3d 1159 (7th Cir. 1994) stated:

A factfinder might infer from these kinds of statements that PPG was motivated by age in its termination decisions. That is not the only inference, of course—[the employer’s regional manager] might simply have meant that the company would not hesitate to discharge any employee who did not meet its expectations, even if he or she were approaching retirement. But this is for the finder of fact to sort out.

*Id.* at 1165 (citation omitted).

Similarly, in *Griffiths v. Cigna Corp.*, 988 F.2d 457 (3d Cir. 1993), the Third Circuit stated: “A court will categorize a case as mixed motives where the plaintiff can ‘satisfy the fact finder that it is more likely than not that a forbidden characteristic played a part in the employment decision.’” *Id.* at 472 (quoting *Price Waterhouse*, 490 U.S. at 247 n.12).

277. Justice Kennedy warned:

Perhaps [disparate treatment] cases in the future will require a bifurcated trial, with the jury retiring first to make the credibility findings necessary to determine whether the plaintiff has proved that an impermissible factor played a substantial part in the decision, and later hearing evidence on the “same decision” or “pretext” issues. Alternatively, perhaps the trial judge will have the unenviable task of formulating a single instruction for the jury on all of the various burdens potentially involved in the case.

*Price Waterhouse*, 490 U.S. at 292 (Kennedy, J., dissenting); see also *Tyler v. Bethlehem Steel Corp.*, 958 F.2d 1176, 1185-87 (2d Cir. 1992) (warning of the potential for jury confusion accompanying the pretext-mixed motives distinction), *cert. denied*, 113 S. Ct. 82 (1993).

Why have the courts floundered so in their efforts to distinguish between pretext and mixed-motives cases? The answer lies in the problem which gave rise to mixed-motives analysis in the first place—the inadequacy of pretext theory as a model for explaining how human cognition works. Kuhn's analysis reminds us that communities of inquiry are reluctant to displace their guiding paradigms.<sup>278</sup> Courts are accordingly engaged in an attempt to carve out separate spheres for the two competing approaches. But the ad hoc solution they have devised is theoretically incoherent and practically unworkable.

There is no sound reason for basing a decision as to which jurisprudential model should apply in a given case on the type of evidence presented by the plaintiff. Rather, courts should formulate rules defining the respective spheres of the two models, if indeed both models are retained,<sup>279</sup> with reference to the type of discrimination each is best equipped to identify and address.

As *Price Waterhouse* describes, pretext theory applies best in cases where the defendant's discriminatory motive was the "sole cause"<sup>280</sup> underlying an employer's decision and where the employer's proffered reason in fact had no actuating significance in his decision, but was simply a post hoc fabrication designed to cover up his "real" discriminatory purpose.<sup>281</sup> Such cases of blatant, deliberate discrimination no doubt exist, but are likely to represent only a small fraction of all employment decisions in which intergroup bias has played a role. Ironically, it is in these cases, rather than in those involving mixed or complex motivations, that one is more likely to find direct evidence of discrimination, such as racial epithets or outright exclusion.

Mixed-motives theory reflects much more accurately than pretext theory the processes by which cognitive sources of bias result in intergroup discrimination. First, very few actions or decisions derive from a single cause. Indeed, a belief in monocausality, or what psychologists Richard Nisbett and Lee Ross call the "hydraulic" model of causation, represents a common source of error in attributional judgment.<sup>282</sup> Pretext theory, with its rigid dichotomization of "real reasons" and "pretexts for discrimination," both reflects and perpetuates this single cause fallacy.

Second, a particular employment decision may be based on *both* the reason proffered by the defendant *and* tainted by intergroup biases mediating the perception, interpretation, encoding, and recall of decision-relevant information and events. Thus, in a very real sense, every case of discrimination resulting from cognitive bias is a "mixed-motives" case. In very few of these cases is one likely to find either "direct" evidence of discriminatory animus or a prof-

---

278. Kuhn, *supra* note 251, at 52-53.

279. I argue below that pretext analysis should be eliminated entirely and replaced with a unified "causal factor" model of disparate treatment proof. See text accompanying notes 365-380 *infra*.

280. See, e.g., *Griffiths*, 988 F.2d at 472.

281. Both the plurality opinion and Justice White's concurrence in *Price Waterhouse* distinguish pretext from mixed-motives cases by asking whether a single motive or a mixture of legitimate and illegitimate motives informed the challenged employment decision. *Price Waterhouse*, 490 U.S. at 258 (White, J., concurring).

For a more extensive discussion of the pretext model, see text accompanying notes 59-67 *supra*.

282. NISBETT & ROSS, *supra* note 246, at 127-28.

ferred nondiscriminatory reason devoid of actuating significance, the subject to which we now turn.

## 2. St. Mary's Honor Center v. Hicks and the eroding presumption of invidiousness.

In June of 1990, Melvin Hicks' race discrimination case against St. Mary's Honor Center went to trial before United States District Judge Stephen N. Limbaugh<sup>283</sup> of the Eastern District of Missouri. After hearing the evidence and analyzing it under the pretext model of proof, Judge Limbaugh found that the nondiscriminatory reasons proffered by St. Mary's for Hicks' discharge were, in fact, not the real reasons his termination.<sup>284</sup>

Despite this finding, Judge Limbaugh ruled for the defendant on the grounds that, although plaintiff had proven the existence of a crusade to terminate him, he had not proven that the crusade was racially rather than personally motivated.<sup>285</sup> In so deciding, Judge Limbaugh refused to assume, as had the Supreme Court in *Furnco Construction Co. v. Waters*, that "when all legitimate reasons for rejecting an applicant have been eliminated as possible reasons for the employer's actions, it is more likely than not that the employer . . . based his decision on an impermissible consideration such as race."<sup>286</sup> In short, Judge Limbaugh rejected the presumption of invidiousness.

As was described in Part I, the logic of the *McDonnell Douglas/Burdine* model of proof had long been based on a willingness to presume discriminatory intent from a finding of pretext.<sup>287</sup> The Supreme Court had never explicitly labelled this a mandatory presumption, but it had long been treated as such.<sup>288</sup> In other words, until the late 1980s, proving that the reasons offered by a disparate treatment defendant were pretexts was implicitly equated with proving that these reasons were pretexts for intentional discrimination.<sup>289</sup> But by the time Hicks' case went to trial in June of 1990, a number of circuits had explicitly

283. For those readers who couldn't help but wonder, the answer is "yes—they are related." Judge Limbaugh is Rush Limbaugh's uncle. John McCollister, *The Rush Is On: Profile of Radio Personality Rush Limbaugh*, SATURDAY EVENING POST, May 1993, at 54.

284. Hicks had produced powerful evidence of pretext, showing that he was the only supervisor ever disciplined by St. Mary's for transgressions committed by his subordinates and that similar and even more serious violations committed by his white coworkers were disregarded or treated more leniently. Most importantly, he convinced the court that his supervisor deliberately provoked the final verbal confrontation used to justify his termination. *St. Mary's Honor Ctr. v. Hicks*, 113 S. Ct. 2742, 2750-51 (1993).

285. *Id.* at 2752.

286. *Furnco Constr. Corp. v. Waters*, 438 U.S. 567, 577 (1978).

287. See generally Lanctot, *supra* note 255.

288. See *United States Postal Serv. Bd. of Governors v. Aikens*, 460 U.S. 711, 718 (1983) (Blackmun, J., concurring) ("The *McDonnell Douglas* framework requires that a plaintiff prevail when . . . he demonstrates that the legitimate, nondiscriminatory reason given by the employer is in fact not the true reason for the employment decision."); *Texas Dep't of Community Affairs v. Burdine*, 450 U.S. 248, 254 (1981) (citing *Furnco* and noting that the successful prima facie case raises the inference that unexplained acts are more likely than not based on consideration of impermissible factors); *Furnco*, 438 U.S. at 577 (calling *McDonnell Douglas* burden of proof a "sensible, orderly way to evaluate the evidence in light of common experience").

289. See text accompanying notes 77-82 *supra*.

rejected this equation.<sup>290</sup> And by the time Hicks' case was over, the Supreme Court had rejected it as well.<sup>291</sup>

While many of my colleagues in the plaintiffs' employment bar saw the pretext-plus controversy and the Supreme Court's decision in *Hicks* as but another right wing assault on Title VII protections,<sup>292</sup> I hold a somewhat different view. The pretext-plus controversy, in my opinion, developed out of a gradual erosion of confidence in the premises underlying pretext theory itself—in particular, the assumption of rational decisionmaking and the presumption of invidiousness. As my earlier analysis and arguments indicate,<sup>293</sup> I do not consider this erosion unwarranted. Neither of these assumptions is logically supportable and both have been so thoroughly undermined empirically that they can no longer be considered valid.

But without the presumption of invidiousness, the entire pretext approach to disparate treatment proof loses its practical utility. If disproving the motivating significance of each of the purportedly nondiscriminatory reasons identified by the defendant in the adjudicative process is not sufficient to establish liability, little remains of *Burdine*'s promise that a Title VII plaintiff can prove discrimination "indirectly by showing that the employer's proffered explanation is un-

---

290. Of these, the 1st, 4th, 6th, 7th, and 10th Circuits had held that disproving the motivating significance of the defendant's proffered explanation was never, in and of itself, sufficient to prove intentional discrimination. See *EEOC v. Flasher Co.*, 986 F.2d 1312, 1321 (10th Cir. 1992) ("the plaintiff must show not merely that the proffered reasons are pretextual but that they are 'a pretext for discrimination' ") (citing *Burdine*, 450 U.S. at 253); *Galbraith v. Northern Telecom, Inc.*, 944 F.2d 275, 282-83 (6th Cir. 1991) (plaintiff bears ultimate burden of proving she was victim of discrimination), *cert. denied*, 112 S. Ct. 1497 (1993); *Samuels v. Raytheon Corp.*, 934 F.2d 388, 392 (1st Cir. 1991) (original case plus evidence of pretext, even if enough to raise inference of discrimination, not enough to satisfy plaintiff's burden); *Holder v. City of Raleigh*, 867 F.2d 823, 827-28 (4th Cir. 1989) (pretext in and of itself does not prove discrimination); *Benzie v. Illinois Dep't of Mental Health and Developmental Disabilities*, 810 F.2d 146, 148 (7th Cir. 1987) (spurious explanation by employer does not compel inference of discrimination as a matter of law), *cert. denied*, 483 U.S. 1006 (1988). On the other hand, the 2nd, 3rd, 5th, 8th, 11th, and D.C. Circuits had held that disproving the motivating significance of the proffered reasons did not compel—but could permit—the trier of fact to find in the plaintiff's favor. See *Hicks v. St. Mary's Honor Ctr.*, 970 F.2d 487, 492-93 (8th Cir. 1992) (holding that once defendant's proffered reasons for adverse treatment shown to be pretextual, plaintiff has met burden), *rev'd*, 113 S. Ct. 2742 (1993); *Caban-Wheeler v. Elsea*, 904 F.2d 1549, 1554-57 (11th Cir. 1990) (showing discriminatory reasons are more likely explanation than proffered reason for discharge satisfies plaintiff's burden); *King v. Palmer*, 778 F.2d 878, 881 (D.C. Cir. 1985) (discrediting defendant's rebuttal to plaintiff's prima facie case satisfies burden); *Thornbrough v. Columbus & Greenville R.R.*, 760 F.2d 633, 639-40, 646-47 (5th Cir. 1985) (disproving defendant's proffered reasons for treatment satisfies burden); *Duffy v. Wheeling Pittsburgh Steel Corp.*, 738 F.2d 1393, 1395-96 (3rd Cir. 1984) (showing that proffered justification is pretextual satisfies burden).

291. The majority opinion in *Hicks* is internally contradictory and could be interpreted in either of two ways. On the one hand, the opinion could be seen as simply supporting the proposition that the trier of fact may, but is not compelled, to conclude from proof of pretext that the plaintiff was the victim of intentional discrimination. On the other hand, it could stand for the proposition that proof of pretext alone is not in and of itself sufficient to support a plaintiff's verdict. In my view, the first interpretation is the better of the two. The majority explicitly disavows the dissent's interpretation: "rejection of the defendant's proffered reasons, will permit the trier of fact to infer the ultimate fact of intentional discrimination," *Hicks*, 113 S. Ct. at 2749.

292. See, e.g., Arnold Shep Cohen, *Discrimination Standard Abandoned*, N.J. LAWYER, July 19, 1993, at 6; *Management, Civil Rights Attorneys Differ on Effect of Hicks Decision*, 1993 DAILY LAB. REP., (BNA) No. 126, at 12 (July 12, 1993).

293. See text accompanying notes 55-82 *supra*.

worthy of credence."<sup>294</sup> Without presuming the presence of discriminatory intent from a finding of pretext, how will pretext theory serve its intended function, which was originally "to sharpen the inquiry into the elusive factual question of intentional discrimination?"<sup>295</sup>

If juries are permitted, but not required, to infer intentional discrimination from a showing that the defendant's proffered reasons are false, what standards should guide them in drawing one inference instead of another? Presumably, they will have heard no direct evidence of discrimination, nor, in many circuits, evidence "directly tied to" the alleged discrimination. In such cases, jury deliberations would proceed under a mixed-motives rather than a pretext approach. What, other than jurors' preexisting biases regarding the prevalence of intentional discrimination, or its plausibility as a causal theory, will inform disparate treatment factfinding after *Hicks*?

The presumption of invidiousness, flawed though it may have been, at least served to control the otherwise unrestrained biases about what might have caused the employer's action. Faced with a choice between intentional discrimination or some other explanation as the sole causal antecedent of the employer's decision, jurors can be expected to rely either on the relative salience of the competing explanations or on whatever a priori causal theory they find most representative of the situation at hand. Given that pretext theory applies only when relatively more salient direct evidence of animus is lacking, discriminatory intent is not likely to be viewed as a convincing explanation for what occurred. This leaves us with one overwhelming influence on jury decision-making: jurors' preexisting biases regarding the prevalence of conscious, intentional discrimination.

Ironically, it is those very employment decisions that are now analyzed under the pretext model of proof which are most apt to have been tainted by unintended cognitive biases in intergroup perception and judgment. These decisions are the least likely, however, to derive from a sole motive, in particular, a conscious, deliberate intent to discriminate. Paradoxically then, after *Hicks* and *Price Waterhouse*, courts instruct jurors to look for monocausal discriminatory intent in precisely those cases where it is least likely to exist.

One could, of course, devise further ad hoc modifications to disparate treatment theory in an attempt to correct these problems. I believe this approach would ultimately fail. The present state of theoretical incoherence and practical unworkability characterizing Title VII disparate treatment theory stems from its equation of causation with intentionality. If Title VII doctrine is to regain its theoretical coherence and practical utility, intentional and unintentional disparate treatment must be distinguished and doctrinal tools must be developed to address each. Before considering these modifications, let us examine one further anomaly generated by current disparate treatment doctrine and the ill-fated attempts of both the courts and Congress to rationalize it within the existing theoretical framework.

---

294. *Burdine*, 450 U.S. at 256.

295. *Id.* at 255, n.8.

3. *Subjective decisionmaking, systemic disparate treatment, and the disparate impact "patch."*

While Part I of this article focuses on individual disparate treatment theory, many of my observations apply equally to class-based or "systemic" disparate treatment doctrine. As in the individual context, Title VII jurisprudence constructs class-based disparate treatment as manifesting discriminatory intent. Thus, in order to prevail in a systemic disparate treatment case, the plaintiff must prove that the defendant maintains a pattern and practice of intentional discrimination as its "standard operating procedure."<sup>296</sup> Statistical disparities between a group's representation in the contested jobs and their representation in the relevant selection pool may support a finding of discrimination, but they are not sufficient to establish pattern and practice liability. Proof of discriminatory motivation is required.<sup>297</sup>

During the late 1970s and early 1980s, courts and employers became more sophisticated in their understanding of the various statistical techniques used to prove class-based discrimination.<sup>298</sup> At the same time, federal judges became increasingly hesitant to infer discriminatory intent from statistical disparities alone.<sup>299</sup> In response to these trends, plaintiffs' advocates sought alternate avenues of redress. In particular, various members of the plaintiffs' bar endeavored to apply Title VII's disparate impact paradigm, which did not require proof of intent, to cases which formerly would have been analyzed under the systemic disparate treatment model of proof.<sup>300</sup> In these cases, plaintiffs ar-

296. *International Bhd. of Teamsters v. United States*, 431 U.S. 324, 336 (1977).

297. As in the individual disparate treatment context, the Court in *Teamsters* equated causation and intentionality: "'Disparate treatment' such as is alleged in the present case is the most easily understood type of discrimination. The employer simply treats some people less favorably than others because of their race, color, religion, sex, or national origin. Proof of discriminatory motive is critical . . ." *Id.* at 335 n.15.

298. See, e.g., *Hazelwood Sch. Dist. v. United States*, 433 U.S. 299, 309-13 (1977) (finding error in disregarding relevant statistical data about school's practices in hiring of black teachers); *Gay v. Waiter's & Dairy Lunchmen's Union Local 30*, 694 F.2d 531, 555 (9th Cir. 1982) (concluding that plaintiff's statistical and circumstantial evidence fails to establish prima facie case of disparate treatment of black male writers); *Wilkins v. University of Houston*, 654 F.2d 388, 396-410 (5th Cir. 1981) (holding that female plaintiffs failed to make prima facie case with statistical evidence of hiring, recruitment, and compensation discrimination).

299. See, e.g., *Segar v. Smith*, 738 F.2d 1249, 1283 (D.C. Cir. 1984), *cert. denied*, 471 U.S. 1115 (1985) (citing with approval district court statement that statistics alone would not permit an inference of discrimination in promotions to positions above the GS-12 level.); *Sweeney v. Research Found. of State Univ. of N.Y.*, 711 F.2d 1179, 1187 (2d Cir. 1983); *Pegues v. Mississippi State Employment Serv.*, 699 F.2d 760, 773 (5th Cir.), *cert. denied*, 464 U.S. 991 (1983); *Wilmore v. City of Wilmington*, 699 F.2d 667, 670 (3d Cir. 1983); *Gay*, 694 F. 2d at 531.

300. See, e.g., *Wilmore*, 699 F.2d at 670 (applying disparate impact theory and using statistical evidence); *Wang v. Hoffman*, 694 F.2d 1146, 1148 (9th Cir. 1982) (holding that lack of objective criteria for promotion can be analyzed for disparate impact); *Rowe v. Cleveland Pneumatic Co.*, 690 F.2d 88, 94 (6th Cir. 1982) (finding that leaving rehiring decisions to the foreman's subjective judgment could be challenged on disparate impact grounds). For a discussion of the frustrations leading to the use of this approach, see generally Elizabeth Bartholet, *Application of Title VII to Jobs in High Places*, 95 HARV. L. REV. 947 (1982) (arguing that disparate treatment theory fails to identify discrimination in upper level jobs).

This strategy had its roots in pre-*Teamsters* cases such as *Pettway v. American Cast Iron Pipe Co.*, 494 F.2d 211, 241 (5th Cir. 1974), *cert. denied*, 439 U.S. 1115 (1979), and *Rowe v. General Motors Corp.*, 457 F.2d 348 (5th Cir. 1972), in which the courts applied disparate impact theory to cases where

gued that seemingly neutral, subjective decisionmaking systems which resulted in gender, ethnic, or racial selection disparities be found discriminatory under Title VII's disparate impact theory. Under this approach, a plaintiff would be required to prove only the existence of the subjective system and a statistically significant disparity in selection rates between the plaintiff class and the favored group. Unless the system was deemed justified as a business necessity, the defendant would be liable. Proof of intent would not be required.<sup>301</sup>

By the late 1980s, appellate courts had split on this use of disparate impact theory to challenge subjective employment decisionmaking. The Fourth, Fifth, Seventh, Eighth Circuits,<sup>302</sup> and certain panels within the Second and Sixth Circuits,<sup>303</sup> refused to apply the disparate impact model to subjective decisionmaking. These courts and some commentators maintained that the business necessity defense, and in particular the cost of validation studies, would pose too onerous a burden on employers in cases involving subjective selection criteria.<sup>304</sup> Additionally, they maintained that, as a conceptual matter, the disparate treatment paradigm provided the better analytical framework for subjective criteria cases because "any discrimination which results from subjective decisionmaking is inherently a result of disparate treatment."<sup>305</sup> Subjective decisionmaking systems were seen as problematic because they failed to provide safeguards against intentional discrimination by individual decisionmakers. Accordingly, class-based disparate treatment theory, not disparate impact theory, was deemed the more appropriate analytical framework.

In contrast, the Third, Ninth, Tenth, Eleventh, and D.C. Circuits permitted individual disparate impact challenges to subjective decisionmaking sys-

---

promotion decisions were left to the unfettered discretion of Caucasian supervisors applying subjective standards. By the early 1980s, however, the Fifth Circuit had rejected this use of disparate impact theory and required that courts analyze "subjective decisionmaking" cases under the systemic disparate treatment model of proof. *See, e.g., Pouncy v. Prudential Ins. Co.*, 668 F.2d 795, 800 (5th Cir. 1982).

301. *See Burdine*, 450 U.S. at 254; *Griggs v. Duke Power Co.*, 401 U.S. 424, 430-32 (1971) (employment practices that disproportionately exclude blacks and are unrelated to job performance struck down regardless of lack of discriminatory intent); *Chrisner v. Complete Auto Transit, Inc.*, 645 F.2d 1251, 1257 (6th Cir. 1981) (proof of discriminatory intent not required in a disparate impact case).

302. The 4th, 5th, 7th, and 8th Circuits rejected application of disparate impact theory. *See Griffin*, 795 F.2d at 1288-89; *Vuyanich v. Republic Nat'l Bank of Dallas*, 723 F.2d 1195, 1201-02 (5th Cir. 1983), *cert. denied*, 469 U.S. 1073 (1984); *Talley v. United States Postal Serv.*, 720 F.2d 505, 507 (8th Cir. 1983); *Carroll v. Sears Roebuck & Co.*, 708 F.2d 183, 190 (5th Cir. 1983); *Pegues*, 699 F.2d at 765; *EEOC v. Federal Reserve Bank of Richmond*, 698 F.2d 633, 639 (4th Cir. 1983), *rev'd on other grounds sub nom. Cooper v. Federal Reserve Bank of Richmond*, 467 U.S. 867 (1984); *Pope v. City of Hickory*, 679 F.2d 20, 22 (4th Cir. 1982); *Pouncy*, 668 F.2d at 800; *Harris v. Ford Motor Co.*, 651 F.2d 609, 611 (8th Cir. 1981).

303. The 2d and 6th Circuits demonstrated an internal division on the disparate impact/subjective practices issue. *Compare Rossini v. Ogilvy & Mather, Inc.*, 798 F.2d 590, 605 (2d Cir. 1986) (refusing to apply disparate impact theory) with *Zahorik v. Cornell Univ.*, 729 F.2d 85, 95-96 (2d Cir. 1984) (applying both disparate treatment and impact theories); *Grant v. Bethlehem Steel Corp.*, 635 F.2d 1007, 1010 (2d Cir. 1980) (applying disparate impact theory), *cert. denied*, 452 U.S. 940 (1981); *compare also Geisler v. Folsom*, 735 F.2d 991, 994 (6th Cir. 1984) (refusing to apply impact analysis) with *Rowe*, 690 F.2d at 92 n.9 (permitting disparate impact analysis).

304. *See, e.g., Bunch v. Bullard*, 795 F.2d 384, 393 n.10 (5th Cir. 1986).

305. Stacey B. Babson, *Evaluation of Subjective Selection Systems in Title VII Employment Discrimination Cases: A Misuse of Disparate Impact Analysis*, 7 CARDOZO L. REV. 549, 582-83; 3 ARTHUR LARSON & LEX K. LARSON, *EMPLOYMENT DISCRIMINATION: REMEDIES, RACE, RELIGION, AND NATIONAL ORIGIN* (1993 ed.), § 76.34.

tems.<sup>306</sup> These courts, and a veritable swarm of academic commentators,<sup>307</sup> argued that disparate treatment theory had proven inadequate to address the subtle bias which characterizes subjective practices discrimination.<sup>308</sup> Proponents justified the use of disparate impact theory as a tool to staunch the flow of "false negatives" generated by systemic disparate treatment analysis and to further the general policies underlying Title VII.<sup>309</sup> Furthermore, proponents argued, from the perspective of their impact on a protected class, there exists little meaningful distinction between objective and subjective practices.<sup>310</sup>

The Supreme Court resolved the split among the circuits in *Watson v. Fort Worth Bank and Trust*,<sup>311</sup> holding unanimously that disparate impact theory could be applied in cases involving subjective decisionmaking criteria. However, four of eight participating justices<sup>312</sup> were sufficiently troubled by how disparate impact analysis could practically be applied in such cases without undermining other values implicated by Title VII that they proposed an almost complete reworking of the disparate impact paradigm.<sup>313</sup>

306. See, e.g., *Atonio v. Wards Cove Packing Co.*, 810 F.2d 1477, 1479 (9th Cir.) (en banc), *reh'g granted*, 827 F.2d 439 (1987), *rev'd on other grounds*, 490 U.S. 642 (1989); *Griffin v. Carlin*, 755 F.2d 1516, 1522-25 (11th Cir. 1985); *Hawkins v. Bounds*, 752 F.2d 500, 503 (10th Cir. 1985); *Lasso v. Woodmen of World Life Ins. Co.*, 741 F.2d 1241, 1246 (10th Cir. 1984), *cert. denied*, 471 U.S. 1099 (1985); *Segar v. Smith*, 738 F.2d 1249, 1270-71 (D.C. Cir. 1984), *cert. denied*, 471 U.S. 1115 (1985); *Wilmore v. City of Wilmington*, 699 F.2d 667, 670 (3d Cir. 1983); *Coe v. Yellow Freight Sys., Inc.*, 646 F.2d 444, 450-51 (10th Cir. 1981); *Williams v. Colorado Springs Sch. Dist.*, 641 F.2d 835, 842 (10th Cir. 1981).

307. See e.g., Susan Melanie Jones, Comment, *Applying Disparate Impact Theory to Subjective Employee Selection Procedures*, 20 LOY. L.A. L. REV. 375 (1987); Julia Lamber, *Discretionary Decisionmaking: The Application of Title VII's Disparate Impact Theory*, 1985 U. ILL. L. REV. 869; Mark Maney, Comment, *Application of the Adverse Impact Analysis to Subjective Criteria in Title VII Employment Discrimination Cases*, 38 BAYLOR L. REV. 363 (1986); Jane Rigler, *Title VII and the Applicability of Disparate Impact Analysis to Subjective Selection Criteria*, 88 W. VA. L. REV. 25 (1985).

308. See, e.g., *Wards Cove*, 810 F.2d at 1484 n.5. The Court stated: "If, in fact, the subjective practices are a 'covert means' to discriminate intentionally, by definition intent will be difficult to prove. . . . Proof of intent where adverse impact can be shown may be not only unnecessary but undesirable because of the animus the process generates." *Id.* at 1484. See also Anita M. Allesandra, Comment, *When Doctrines Collide: Disparate Treatment, Disparate Impact and Watson v. Fort Worth Bank and Trust*, 137 U. PA. L. REV. 1755, 1776-1778 (1989).

309. See *Wards Cove*, 810 F.2d at 1483.

310. As the Court stated in *Wards Cove*, "[t]here is no bright line distinction between objective and subjective hiring criteria, because almost all criteria necessarily have both subjective and objective elements." *Id.* at 1485.

311. 487 U.S. 977, 999 (1988).

312. *Watson* produced three opinions, all of which concluded that disparate impact theory could be applied to challenge subjective decisionmaking practices. Justice O'Connor, joined by Justices Rehnquist, White, and Scalia, articulated the concerns and proposed the doctrinal modifications to which I refer. Justice Kennedy took no part in the consideration and decision of the case.

313. For example, the *Watson* plurality proposed lowering the business necessity burden from a burden of proof to a burden of production and endorsed a lenient interpretation of business necessity which merely required an employer to produce evidence of "legitimate business reasons" for using a particular employment practice. *Watson*, 487 U.S. at 997, 999. Additionally, the plurality rejected the imposition of an obligation on employers to engage in statistical validation of selection devices, whether subjective or objective. *Id.* at 998. The following year, in *Wards Cove*, a majority of the Court accepted the doctrinal changes proposed in *Watson*, but remained silent on the question of formal validation. *Wards Cove*, 490 U.S. at 656-61. *Wards Cove* prompted Congress to amend Title VII to restore disparate impact theory's elements of proof and defense as they had existed before *Wards Cove*. Civil Rights Act of 1991, Pub. L. 102-166, §§ 2-3, 105 Stat. 1071 (1991) (codified at 42 U.S.C. § 2000e-2(k))

All of the justices participating in *Watson* recognized that an unfettered system of subjective decisionmaking could produce the same discriminatory results as a system "pervaded by impermissible intentional discrimination."<sup>314</sup> Furthermore, all acknowledged that the disparate treatment paradigm failed to provide an adequate analytical framework for subjective practice adjudications. As Justice O'Connor observed, even if one assumed that disparate treatment theory was sufficient to address intentional discrimination, "the problem of subconscious stereotypes and prejudices would remain."<sup>315</sup>

However, the *Watson* plurality, led by Justice O'Connor, feared that applying disparate impact theory to cases involving subjective selection criteria would pose a serious threat to other values and interests at stake in Title VII cases. Specifically, the justices observed that, as a practical matter, subjective decisionmaking practices would be so difficult and expensive to validate, that, were traditional disparate impact theory applied to subjective decisionmaking, employers would be forced to adopt and enforce numerical selection quotas to avoid liability. Whatever one might think about this solution, it was unacceptable to the *Watson* plurality.<sup>316</sup>

How, theoretically, should we understand the problem presented by the subjective practices/disparate impact cases? In a sense, application of disparate treatment theory to subjective practices cases produced anomalous results which systematically violated appellate courts' sense of justice. This sense of injustice serves as the jurisprudential equivalent of Kuhn's "experimental anomalies"<sup>317</sup> which, left unaddressed, can erode a community's confidence in its most important analytical tools.

The facts of Clara Watson's case provide a compelling example. At one point in her long, unsuccessful attempt to secure a promotion to teller, Ms. Watson was told by one of her superiors "that the teller position was a big responsibility with 'a lot of money for . . . blacks to have to count.'" <sup>318</sup> Nonetheless, Ms. Watson was unable to prove that the bank had engaged in intentional discrimination. Consequently, she lost both her individual and class-based disparate treatment claims, an outcome Justice O'Connor noted with some consternation.<sup>319</sup>

The apparent injustice generated by the application of disparate treatment theory in these cases might have prompted the Court to recognize that the equation of causation and intentionality in individual and class-based disparate treatment jurisprudence is fundamentally flawed. The Court could have acknowledged that at least some forms of disparate treatment caused by unconscious stereotypes and prejudices do not necessarily manifest discriminatory

---

(1992)). Congress too remained silent not only on the validation question, but also on the question of precisely what was meant by the term "business necessity." See 42 U.S.C. § 2000e-2(k)(1)(A) (1992).

314. *Watson*, 487 U.S. at 997.

315. *Id.* at 990-91.

316. *Id.* at 990.

317. KUHN, *supra* note 251, at 52-53.

318. *Watson*, 487 U.S. at 990.

319. *Id.* at 983-85.

motive or intent. With a more nuanced understanding of disparate treatment discrimination, the Court could then have restructured disparate treatment theory itself. In particular, it could have interpreted the Act as requiring simply that a plaintiff demonstrate a causal connection between her group status and the employer's decision rather than as requiring proof of a particular kind of causation, namely specific intent to discriminate.

Unfortunately, the Court chose to do neither. Rather, it turned to disparate impact theory to solve what is essentially a disparate treatment problem. Ultimately this approach must fail. The disparate impact paradigm as currently constructed is an inappropriate analytical tool for addressing the intergroup biases inherent in subjective decisionmaking.

This claim is based on three arguments—one practical, one political, and one theoretical. From a practical standpoint, disparate impact theory is inadequate because the empirical tools on which it depends cannot effectively or economically be applied in situations involving complex, subjective assessments. Further, it is only disparate impact's grounding in empiricism which provides its political legitimacy as a civil rights theory in the face of competing normative claims. Finally, as a theoretical matter, I argue that disparate impact theory is the wrong tool to address subjective decisionmaking because it presupposes a significantly different type of bias from those at play in subjective practices cases. From a phenomenological standpoint, subjective practices discrimination is a disparate *treatment* problem, not a disparate *impact* problem, and it requires a disparate treatment solution. Let us consider each of these three arguments and their implications for our present inquiry.

Under current disparate impact theory<sup>320</sup> after a plaintiff demonstrates that a particular employment practice has a disparate impact on a protected class, the burden shifts to the employer to prove that the challenged practice is “job related for the position in question and consistent with business necessity.”<sup>321</sup> In fashioning the compromise which made passage of the 1991 Civil Rights Act possible, Congress deliberately failed to define the terms “job-related” and “business necessity.”<sup>322</sup> However, the business necessity defense had long been interpreted to require that employers formally validate disparately im-

---

320. Section 105 of the Civil Rights Act of 1991 amended Title VII to reverse the stringent burdens of proof proposed by the *Watson* plurality and ultimately adopted in *Wards Cove*. See 42 U.S.C. § 2000e-2(k)(1) (1992).

321. 42 U.S.C. § 2000e-2(k)(1)(A)(i) (1992).

322. Neither term is defined in the Act. Further, § 105(b) provides that no statements other than an Interpretive Memorandum appearing at 137 CONG. REC. S15,276 may be considered or relied upon as legislative history in construing the disparate impact provisions of the Act. The brief memorandum to which the Act refers states only that “the terms ‘business necessity’ and ‘job related’ are intended to reflect the concepts enunciated by the Supreme Court in [*Griggs*], and the other Supreme Court decisions prior to [*Wards Cove*].” 137 CONG. REC. S15,276 (daily ed. Oct. 25, 1991). Previous Supreme Court decisions defining these two terms were ambiguous and in many ways contradictory; the Act merely codified this uncertainty.

pacting<sup>323</sup> selection devices in accordance with professionally accepted statistical procedures.<sup>324</sup>

Herein lies the practical problem with applying disparate impact theory to subjective practices cases: Validating subjective decisionmaking systems in accordance with professionally acceptable standards is neither empirically nor economically feasible, especially for jobs where intangible qualities, such as interpersonal skills, creativity, and the ability to make sound judgments under conditions of uncertainty, are critical. Accordingly, if a court applies disparate impact theory in subjective practices cases, one of two undesirable outcomes will necessarily result: either the validation requirement will be weakened or eliminated entirely, as already appears to be occurring, or its imposition will place severe and ultimately unworkable burdens on small and medium-sized employers.

Scholarly treatments of the subjective practices/disparate impact problem divide sharply on the question of whether the law should require employers to validate subjective decisionmaking systems.<sup>325</sup> Those commentators who favor application of the validation requirement rely on a combination of doctrinal arguments and appeals to fairness toward groups protected by Title VII.<sup>326</sup> Although these commentators assert that validation is feasible in the subjective decisionmaking context, a rigorous analysis of validation procedures is generally missing from their work. Rather than conducting their own analysis to support their assertion that validation is technically feasible, these treatments generally rely on the highly qualified conclusions of the American Psychological Association ("APA") in its *Watson* amicus brief.<sup>327</sup>

The APA's brief in *Watson* argued that employers could validate subjective decisionmaking systems with the same methods used to validate objective crite-

323. Validation is a statistical method which can determine whether a particular selection device actually predicts successful performance in a particular activity. See Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1607.1-.7 (1994).

324. See *id.* (stating that the use of any selection procedure which has an adverse impact will be considered discriminatory unless the procedure has been validated in accordance with guidelines); *Dothard v. Rawlinson* 433 U.S. 321, 331 (1977); *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 431 (1975).

325. Compare, e.g., Allesandra, *supra* note 308 (arguing that the validation requirement should apply to subjective decisionmaking systems); Barholet, *supra* note 300; Alfred W. Blumrosen, *The Legacy of Griggs: Social Progress and Subjective Judgments*, 63 CHI.-KENT L. REV. 1 (1987); Lamber, *supra* note 307; Steven L. Willborne, *The Disparate Impact Model of Discrimination: Theory and Limits*, 34 AM. U. L. REV. 799 (1985) with Babson, *supra* note 305 (arguing against imposition of the validation requirement in subjective practices cases); Kingsley R. Browne, *The Civil Rights Act of 1991: A "Quota Bill", A Codification of Griggs, A Partial Return to Wards Cove, or All of the Above?*, 43 CASE W. RES. L. REV. 287 (1993); George Rutherglen, *Disparate Impact Under Title VII: An Objective Theory of Discrimination*, 73 VA. L. REV. 1297 (1987).

326. See, e.g., Theodore Y. Blumoff & Harold S. Lewis, *The Reagan Court and Title VII: A Common-Law Outlook on a Statutory Task*, 69 N.C. L. REV. 1, 30 (1990); Merrill D. Feldstein, *Watson v. Ft. Worth Bank & Trust: Reallocating the Burdens of Proof in Employment Discrimination*, 38 AM. U. L. REV. 919, 946-947 (1989); see generally Christopher Dee, Comment, *Disparate Impact and Subjective Employment Criteria Under Title VII*, 54 U. CHI. L. REV. 957 (1987) (arguing against limitation of plaintiff's challenges to objective criteria only); Jane Howard-Martin, *A Critical Analysis of Judicial Opinions in Professional Employment Discrimination Cases*, 26 HOW. L.J. 723 (1983).

327. Brief for American Psychological Association, *Watson v. Fort Worth Bank & Trust*, 487 U.S. 977 (1988) (No. 86-6139).

ria.<sup>328</sup> This conclusion, however, is highly qualified. Specifically, the APA states that subjective practices are amenable to validation only if they derive from a "thorough analysis of the target job,"<sup>329</sup> if they are sufficiently "behaviorally specific,"<sup>330</sup> if "all ratings can be supported by objective, observable evidence,"<sup>331</sup> if the observations are "systematically recorded,"<sup>332</sup> and if "raters are trained to reduce sources of bias."<sup>333</sup> These caveats effectively swallow the APA's conclusion whole.

The process of validation assumes that a particular selection instrument will produce a quantifiable and accurate "output" or score of a candidate's performance on a particular selection instrument. However, as the APA concedes, the output of any subjective assessment is a personal, nonquantifiable determination subject to many forms of evaluator bias.<sup>334</sup> Reducing an evaluator's subjective evaluation of a candidate's persuasiveness, judgment, or analytical ability to a reliable numerical score is a questionable endeavor. Thus, the accuracy of the very variable being validated cannot be assured.

A related problem stems from the questionable reliance of validation procedures on statistical regression analysis. Regression analysis requires that the selection device be an identified, isolated variable. But the separation of an evaluator's subjective assessment of one trait or criterion from his assessment of another related trait or criterion is often impossible. Subjective assessments of intangible qualities are products of a combinative or synthetic process involving many *interrelated* subjective and objective factors.<sup>335</sup> Thus, the independence of various predictive variables cannot be assured.

Equally serious measurement problems appear with respect to quantifying job performance. Criterion-related validity studies<sup>336</sup> require some quantified measure of job performance with which performance on the challenged selection device can be correlated. Frequently, however, it is impossible to quantify employee performance as criterion-related validity testing requires. In such situations, the only available measure of performance is itself likely to be subjective evaluations, often provided by the same supervisors who made the subjective assessments involved in employee selection.<sup>337</sup> This poses obvious

---

328. *Id.* at 22.

329. *Id.* at 17.

330. *Id.*

331. *Id.* at 21.

332. *Id.* at 4-5.

333. *Id.* at 21.

334. *Id.*

335. Babson, *supra* note 305, at 577 n.134, 577 n.142.

336. Criterion-related validity is demonstrated by establishing a statistically significant positive correlation between successful performance on the selection instrument and success on some measure of job performance. The 1970 Uniform Guidelines on Employee Selection Procedures appear to favor criterion validation over other methods. See 29 C.F.R. § 1607.5(a), 35 Fed. Reg. 12,333, 12,334-35 (1970). See also Rutherglen, *supra* note 325.

337. See Ivan T. Robertson & Andrew Kinder, *Personalities and Job Competencies: The Criterion-Related Validity of Some Personality Variables*, 66 J. OCCUPATIONAL & ORGANIZATIONAL PSYCHOL. 225, 241 (1993). Robertson and Kinder analyzed 20 individual criterion-related validity studies involving personality variables. In all 20 studies, criterion measures had been based on supervisor ratings.

problems. First, a subjective selection device, such as the assessment of a particular personality trait, may be indistinguishable from the performance criterion with which it is being compared. Even if the selection criterion and the performance criterion can be distinguished, both may be affected by the very evaluator bias that makes subjective decisionmaking problematic in the first place. Accordingly, the use of supervisor ratings in criterion-related validity studies comes coupled with profuse apologies for their inherent unreliability.<sup>338</sup> As one commentator reflects, justifying one questionable form of subjective evaluation by appealing to another holds little meaning; the entire process is subject to "a nearly inescapable circularity."<sup>339</sup>

Content validation<sup>340</sup> strategies are even less feasible in the subjective practices context. In order to distinguish content from construct validity,<sup>341</sup> the Uniform Guidelines restrict the application of content validation strategies to situations involving concrete behaviors actually performed on the job.<sup>342</sup> Subjective decisionmaking criteria relate to abstract, intangible qualities, not tangible, objectively measurable skills or abilities. Content validity is simply inapplicable in the context of jobs requiring abstract rather than concrete skills; it is designed for situations involving simple, well-defined elements with easily observable manifestations.<sup>343</sup>

Unfortunately, construct validation, while theoretically the most appropriate method for validating the intangible abilities or personality constructs at issue

338. See, e.g., Paul R. Sackett, Neal Schmitt, Mary L. Tenopyr, Jerard Kehoe, & Sheldon Zedeck, *Commentary on Forty Questions About Validity Generalization and Meta-Analysis*, 38 PERSONNEL PSYCHOL. 697, 731 (1985) ("We know supervisory ratings are flawed, but they're all we could get.").

339. Rutherglen, *supra* note 325 at 1342; see also Alessandra, *supra* note 308 at 1771, n.85 (observing that "[s]ubjective criteria . . . require qualities or skills which cannot be measured objectively").

340. Content validation requires that an employer analyze the job in question by breaking the job down into the knowledge, skills, and abilities (KSAs) and determining the relative importance of and the proficiency level required for each. The employer must demonstrate that the challenged selection device tests the applicant's proficiency on each of these KSAs in accordance with the relevant levels of importance and required proficiency. See *Guardians Ass'n of N.Y. City Police Dep't v. Civil Serv. Comm'n*, 633 F.2d 232, 242 (2d Cir. 1980), *aff'd*, 463 U.S. 582 (1983); see also Dennis Doverspike, Gerald V. Barrett & Ralph A. Alexander, *The Feasibility of Traditional Validation Procedures for Demonstrating Job Relatedness*, 9 LAW & PSYCHOL. REV. 35, 43 (1985).

341. Construct validity is demonstrated by establishing a statistically significant correlation between performance on an assessment instrument and the possession of some perceived trait, such as "intelligence," "leadership," or "motivation," deemed necessary for the successful performance of a job. Construct validation is uniformly viewed with suspicion. The Uniform Guidelines express skepticism regarding the utility of construct validation in the employment testing context, as do testing professional associations. See 29 C.F.R. § 1607.14(D)(1); see also KENNETH N. WEXLEY & GARY A. YUKI, ORGANIZATIONAL BEHAVIOR AND PERSONNEL PSYCHOLOGY 443 (Kenneth N. Wexley & Gary A. Yuki eds., 1984); Michael R. Carrier, Anthony T. Dalessio, & Steven H. Brown, *Correspondence Between Estimates of Content and Criterion-Related Validity Values*, 43 PERSONNEL PSYCHOL. 85, 85 (1990). Very rarely have employers even attempted to use a construct validation strategy to justify a challenged selection device. See generally Barbara Lindemann Schlei & Paul Grossman, EMPLOYMENT DISCRIMINATION LAW 153 (2d ed., 1982).

342. 29 C.F.R. § 1607.14(C)(1)-(4). See Rutherglen, *supra* note 325, at 1325 n.119, 1342.

343. The Uniform Guidelines state that "a content strategy is not appropriate for demonstrating the validity of selection procedures which purport to measure traits or constructs, such as intelligence, aptitude, personality, commonsense, judgment, leadership, and spatial ability." 29 C.F.R. § 1607.14(C)(1). For further discussion of this issue, see Michael R. Carrier, Anthony T. Dalessio, & Steven H. Brown, *Correspondence Between Estimates of Content and Criterion-Related Validity Values*, 43 PERSONNEL PSYCHOL. 85, 97 (1990).

in subjective decisionmaking cases, is practically and economically unfeasible. Consider the dire warning contained in the Uniform Guidelines:

Construct validity is a more complex strategy than either criterion-related or content validity. . . . The user should be aware that the effort to obtain sufficient empirical support for construct validity is both an extensive and arduous effort involving a series of research studies, which include criterion related validity studies and which may include content validity studies. Users choosing to justify use of a selection procedure by this strategy should therefore take particular care to assure that the validity study meets the standards set forth below.<sup>344</sup>

Those standards require, among other things, that criterion-related validity studies be conducted before construct-related validation is undertaken.<sup>345</sup> Consequently, all of the problems which make criterion-related analysis technically unfeasible in the subjective practices context render construct validation unfeasible as well.

Formal validation of employee selection devices are premised on an increasingly outdated Tayloristic model of production in which job-related behaviors and performance outcomes can be readily segmented and quantified. Their application to more complex jobs is open to serious question.<sup>346</sup>

Even if there were solutions to these technical problems, additional practical problems would remain. First, small businesses, or enterprises of whatever size with few workers in the same job classification, are precisely those types of employers most likely to use subjective decisionmaking.<sup>347</sup> Consequently, we can reasonably predict that, in the vast majority of cases, the challenged subjective practice will not have been applied to a sufficiently large number of candidates to provide a statistically adequate sample which to analyze the challenged subjective practice. Without sufficient data, validity studies, even if otherwise feasible, could not produce statistically significant results.

Finally, legal policymakers cannot ignore the problem of cost. Formal validation of even relatively straightforward objective selection devices is an expensive and time-consuming process, often requiring several years and hundreds of thousands of dollars in professional fees and employee time.<sup>348</sup> As the Uniform Guidelines recognize, construct validation studies are even more

---

344. 29 C.F.R. § 1607.14(D)(1).

345. *Id.* § 1607.14(D)(4).

346. See, e.g., Barry R. Nathan & Ralph A. Alexander, *A Comparison of Criteria for Test Validation: A Meta-Analytic Investigation*, 41 PERSONNEL PSYCHOL. 517, 533 (1988) (concluding that some subjective criteria in the selection of clerical employees could be subjected to traditional criterion-related validity analysis, but not necessarily in selection of more complex managerial or professional positions).

347. For a discussion of changes in the labor market leading to increased use of subjective decisionmaking methods, see generally Blumrosen, *supra* note 325.

348. See James Gwartney, Ephraim Asher, Charles Haworth, & Joan Haworth, *Statistics, the Law and Title VII: An Economist's View*, 54 NOTRE DAME L. REV. 633, 634 (an employer seeking to validate a single selection device, such as an arithmetic test for machinists, can expect to spend from \$20,000 to \$100,000); Barbara Lerner, *Employment Discrimination: Adverse Impact, Validity and Equality*, 1979 SUP. CT. REV. 17, 18 n.6 (reporting that adequate criterion-related validity studies generally cost between \$100,000 and \$400,000 and require approximately two years to complete); Rutherglen, *supra* note 325, at 1317-18 (cost of validation in cases studied exceeds \$100,000).

expensive and time consuming than criterion or content studies. Lastly, even if completed, construct validity is sufficiently controversial that such studies would provide little assurance against disparate impact liability.

One could, of course, apply disparate impact theory to subjective decision-making without imposing a duty to validate. The Uniform Guidelines provide for this possibility in section 1607.6, which provides in relevant part:

B. There are circumstances in which a user cannot or need not utilize the validation techniques contemplated by these guidelines. In such circumstances, the user should utilize selection procedures which are as job related as possible and which will minimize or eliminate adverse impact, as set forth below.

(1) When an informal or unscored selection procedure which has an adverse impact is utilized, the user should eliminate the adverse impact, or modify the procedure to one which is a formal, scored or quantified measure or combination of measures and then validate the procedure in accord with these guidelines, or otherwise justify continued use of the procedure in accord with Federal law.<sup>349</sup>

This portion of the Uniform Guidelines fails to offer a politically viable alternative to validation.

The political legitimacy of disparate impact theory rests on a delicate balance of competing normative claims. On one side of this balance rest employer claims to freedom and autonomy. According to these values, if an employer wishes to use a particular employee selection device which she believes will predict successful job performance and thereby enhance the productivity of her enterprise, courts should generally permit her to do so. But, if the chosen selection device significantly and systematically excludes members of a group protected by Title VII, equality and autonomy-related values conflict.

Given that the persuasiveness of the employer's position is based on an empirical claim—namely, that using a particular selection device will enhance the performance of her enterprise—it comports with a subjective sense of fairness to mediate the conflict between equality and freedom with an empirical standard. If, in accordance with empirically acceptable methods, the selection device is shown to be predictive of those particular elements of work behavior which comprise the position in question, employer autonomy trumps equality. If, however, the employer's empirical claim cannot be empirically justified, equality trumps autonomy.

In a sense, the validation requirement uses mathematics as a neutral arbiter of the conflict between two fundamental values—equality and freedom. In a culture that trusts empiricism, it is the empirical grounding of the disparate impact model that gives it political and philosophical legitimacy. If, as I have argued, validating subjective decisionmaking systems in accordance with professionally acceptable standards is neither empirically nor economically feasible, disparate impact theory, when applied to subjective criteria, loses its grounding in empiricism. Individual judicial decisionmakers applying vague,

---

349. 29 C.F.R. § 1607.6(B), (B)(1). The Guidelines nowhere indicate what would "justify continued use of the procedure in accord with Federal law."

subjective, and unpredictable legal standards replace relatively objective empirical methods as the arbiters of business necessity. Thus, the practical problems inherent in applying disparate impact theory to subjective decisionmaking become political problems as well.

Disparate impact theory was created to address two particular equality problems which surfaced in the wake of Title VII's initial implementation. The first originated in response to employers' substitution of ostensibly neutral requirements for the blatant, race-based classifications previously used as proxies for job-relevant characteristics.<sup>350</sup> Once judicially accepted, the model was extended to challenge gender inequities resulting from employers' reliance on the representativeness heuristic in identifying attributes associated with successful performance in stereotypically male jobs.<sup>351</sup> In each of these contexts, the defendant employers had adopted ostensibly objective, neutral selection devices which, either by accident or design, shielded them against accusations of intentional discrimination. The model met these ostensibly empirical selection tools on their own empirical terms. It was not designed to reckon with the phenomena at play in subjective, individualized interpersonal decisionmaking, and it remains ill-suited to that task.

The application of a jurisprudential model outside of its original practical context is a process fraught with peril. Such attempts, even if initially successful from a partisan point of view, can over time lead to the erosion of a legal theory's "practical logic" and ultimately backfire on the theory's proponents. This is, I suggest, precisely what happened in *Watson* and in *Wards Cove*, and it may continue as the courts grapple with the many tough issues dodged by Congress in the Civil Rights Act of 1991.

In short, if current disparate impact theory is applied to subjective practices cases, we will find ourselves in one of two positions. Either the model will be relegated to a narrowing range of cases, as already appears to be occurring, or its application will place severe and ultimately unworkable burdens on employers in the fastest-growing and most promising segments of the national labor market.<sup>352</sup>

---

350. The Supreme Court's initial exposition of disparate impact theory in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), evolved from a Title VII claim against an employer with a history of blatant racial discrimination. *Id.* at 426-27. Following the passage of the Civil Rights Act of 1964, the Duke Power Company replaced its blatant, racially exclusionary hiring practices with two facially neutral requirements—a high school education requirement, and satisfactory scores on standardized aptitude tests. These two requirements disqualified a disproportionate percentage of black applicants from consideration. *See id.* at 426.

351. *See, e.g., Dothard v. Rawlinson*, 433 U.S. 321 (1977). In *Dothard*, the Court invalidated on disparate impact grounds a requirement that all correctional counselors with the Alabama Board of Corrections meet a minimum 5'2" height requirement and weigh at least 120 pounds. *Id.* at 324-25. After *Dothard*, an employer cannot presume a relationship between job performance and a particular attribute or characteristic if such a presumption disproportionately excludes members of a group protected by Title VII. *See id.* at 330.

352. *See generally* Bartholet, *supra* note 320, at 945 (documenting judicial reticence toward applying disparate impact theory to upper level jobs).

B. *The Costs of Disjunction Between Theory and World: Effects of Incoherent Theory on Equal Employment Adjudications and the Exacerbation of Intergroup Tensions*

I have argued that the disjunction between the jurisprudential construction of disparate treatment discrimination and the real life phenomenon it purports to represent has led to an analytical unraveling of Title VII's disparate treatment and disparate impact liability models. The resulting doctrinal complexity and theoretical incoherence has a number of negative practical consequences.

Consider the situation confronting the parties to a newly filed individual disparate treatment case. They cannot know until well into the litigation which liability model will govern. Will this be a mixed-motives case, greatly improving the plaintiff's strategic position, or will it be a pretext case, advantaging the defendant? In some jurisdictions, this determination cannot be made until the case goes to the jury. In others, the rules are uncertain and the issue may not be resolved until appeal.

Should the plaintiff proceed under a disparate impact theory, equally tangled uncertainties will confront the litigants and the court. Is validation of the defendant's decisionmaking system required? If not, what does "job related for the position in question and consistent with business necessity"<sup>353</sup> mean in the context of subjective decisionmaking? More fundamentally, when does a particular subjective employment decision constitute a "particular employment practice"<sup>354</sup> amenable to a disparate impact challenge? Answers to these questions greatly affect the parties' chances of success at trial, but, short of appeal, the parties have little way of predicting what those answers might be.

Should the plaintiff proceed under traditional pretext analysis, different but equally undesirable effects follow. As described earlier, the plaintiff in a pretext case must effectively disprove the motivating significance of each legitimate, nondiscriminatory reason proffered by the defendant. After *Hicks*, he must also be able to predict what additional reasons, not offered by the defendant, might surface at trial and be deemed causally efficacious by the trier of fact. To prove that race or some other protected status was the "real," "sole" reason for the defendant's decision requires the plaintiff to construct a kind of informal multivariate analysis which controls for all potential causal factors and supports the conclusion that the plaintiff's group status is the only plausible explanation for the defendant's decision.

This requires the investment of substantial discovery and analysis costs, even if the analysis is not technically statistical, but is based on a collection of anecdotal comparisons with other "similarly situated" employees. Under the pretext model of proof, plaintiffs will have to engage in extensive discovery, thereby increasing litigation costs and delays. Indeed, failure to take the extensive discovery required to support an effective pretext analysis could easily constitute legal malpractice by plaintiff's counsel. In short, the nature of pre-

---

353. 42 U.S.C. § 2000e-2(k)(1)(A)(i) (1992).

354. *Id.* at § 2000e-2(k)(1)(B)(i).

text proof, particularly after *Hicks*, necessitates the investment of substantial information costs in disparate treatment adjudications.

Even if the parties go to the extremes required by pretext analysis, it may still be difficult, if not impossible, to distinguish group status-related judgment errors from errors based on other cognitive or motivational contaminants.<sup>355</sup> If in fact a broad class of discriminatory employment decisions results not from discriminatory motivation, but from a variety of unconscious and unintentional categorization-related judgment errors, current disparate treatment theory can be expected both to underidentify and overidentify discriminatory employment decisions.<sup>356</sup> Under the current interpretation of Title VII's disparate treatment prohibition, a plaintiff who was victimized by cognitive as opposed to motivational sources of intergroup bias will have no remedy under Title VII unless he can convince a jury that the discrimination was intentional. Thus, in many cases, plaintiffs who were in fact subjected to cognitive sources of intergroup bias will lose. In other cases, the well-intended decisionmaker whose intergroup perception and judgment has been tainted by unconscious cognitive bias will be found liable for intentional discrimination. Whether or not one has been involved in a disparate treatment adjudication, it is easy to imagine the defensiveness, bitterness, and intergroup resentment such results might engender.

We take for granted that laws should serve a normative function. They are enacted to encourage socially desirable behaviors and discourage socially undesirable ones. But a law will be normatively ineffectual if it is structured, interpreted, or enforced in a manner that does not effectively communicate compliance criteria to its target community. It is difficult to comply with a law if you do not understand what it means or requires. But Title VII has become just such a law. Mired in increasingly incoherent liability models and premised on faulty assumptions about the nature of human social judgment, Title VII disparate treatment jurisprudence is failing to explain what it means "not to discriminate."

At the root of this problem lie the faulty assumption that disparate treatment discrimination necessarily manifests discriminatory motive or intent and the mistaken belief that a proscriptive duty "not to discriminate" can eliminate intergroup bias. As I have attempted to demonstrate, intergroup discrimination can be cognitive as well as motivational in origin. Further, not only are many of the cognitive processes which cause discrimination automatic and beyond ordinary conscious self-awareness, they are adaptive, indeed, essential to effective cognitive functioning. People will divide the natural and social environment into categories; they will use stereotypes, scripts, and schemas to interpret, encode, and retrieve information relevant to social judgment. They will rely on the availability and representativeness heuristics to estimate frequency and predict the future. And, because race, ethnicity, and gender have

---

355. See, e.g., *Moze v. Jeffboat, Inc.*, 746 F.2d 365, 371 (7th Cir. 1984).

356. Given the difficulty involved in identifying individual instances of discrimination, the likelihood of error in disparate treatment adjudications is quite high. See David A. Strauss, *The Law and Economics of Racial Discrimination in Employment: The Case for Numerical Standards*, 79 GEO. L.J. 1619, 1644-46 (1991). A cognitive bias approach to discrimination plainly supports Strauss' view.

been made salient by our history and by observable patterns of economic, demographic, and political distribution, people will continue to categorize along those lines. Indeed, by proscribing discrimination on certain defined bases, Title VII makes its own protected classifications salient.

A legal duty which admonishes people simply not to consider race, national origin or gender harkens to Dostoevsky's problem of the polar bear: "Try . . . not to think of a polar bear, and you will see that the cursed thing will come to mind every minute."<sup>357</sup> For reasons this anecdote makes plain, the "colorblindness" approach to the nondiscrimination duty embodied in current disparate treatment jurisprudence cannot succeed in eliminating category-based judgment errors and thus cannot effectuate equal employment opportunity.<sup>358</sup> As Eleanor Rosch observed, people cannot be admonished out of categorical divisions so long as those divisions help them explain and function in their natural or social environment.<sup>359</sup> And so long as people categorize along lines of race, gender, or ethnicity, we can expect the resulting categorization-related distortions in social perception and judgment to bias intergroup decisionmaking.

A colorblindness-centered interpretation of the nondiscrimination principle, coupled with people's awareness that they do categorize along racial and ethnic lines, may well account for much of the intergroup anxiety and ambivalence which various social psychologists have posited as the underlying cause of "aversive racism."<sup>360</sup> As these theorists have observed, racial ambivalence, normative ambiguity, and fear of one's own potential prejudice all serve to amplify white's discrimination against blacks in the giving and requesting of assistance,<sup>361</sup> the evaluation of behavior,<sup>362</sup> physical distancing,<sup>363</sup> and the selection of sanctions for social transgressions.<sup>364</sup> Thus, not only might the prescriptive approach toward nondiscrimination embodied in disparate treatment

357. Daniel M. Wegner, *You Can't Always Think What You Want: Problems in the Suppression of Unwanted Thoughts*, 25 *ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY* 193, 193 (Mark P. Zanna ed., 1992) (quoting Fyodor Dostoevsky, *WINTER NOTES ON SUMMER IMPRESSIONS*).

358. For other critiques of the "colorblindness" perspective, see generally T. Alexander Aleinikoff, *A Case for Race-Consciousness*, 91 *COLUM. L. REV.* 1060, 1068 (1991); Alan David Freeman, *Legitimizing Racial Discrimination Through Antidiscrimination Law: A Critical Review of Supreme Court Doctrine*, 62 *MINN. L. REV.* 1049, 1066 (1978); Neil Gotanda, *A Critique of "Our Constitution is Color-Blind"*, 44 *STAN. L. REV.* 1, 68 (1991); David A. Strauss, *The Myth of Colorblindness*, 1986 *SUP. CT. REV.* 99.

359. See *Human Categorization*, *supra* note 115, at 40.

360. See, e.g., Gaertner & Dovidio, *supra* note 179, at 61-62; Donald G. Dutton & Robert A. Lake, *Threat of Own Prejudice and Reverse Discrimination in Interracial Situations*, 28 *J. PERSONALITY & SOC. PSYCHOL.* 94 (1973).

361. See, e.g., Faye Crosby, Stephanie Bromley, & Leonard Saxe, *Recent Unobtrusive Studies of Black and White Discrimination and Prejudice: A Literature Review*, 87 *PSYCHOL. BULL.* 546 (1980); Samuel L. Gaertner & John F. Dovidio, *The Subtlety of White Racism, Arousal, and Helping Behavior*, 35 *J. PERSONALITY & SOC. PSYCHOL.* 691 (1977).

362. See, e.g., Irwin Katz, Glen R. Hass, & Joyce Wackenhut, *Racial Ambivalence, Value Duality, and Behavior in Prejudice, Discrimination, and Racism*, *supra* note 39, at 35.

363. See, e.g., Carl O. Word, Mark P. Zanna, & Joel Cooper, *The Nonverbal Mediation of Self-Fulfilling Prophecies in Interracial Interaction*, 10 *J. EXPERIMENTAL SOC. PSYCHOL.* 109 (1974).

364. See, e.g., Katz, Hass, & Wackenhut, *supra* note 362.

jurisprudence be normatively ineffectual, it may actually serve to exacerbate intergroup tensions.

#### IV. PROPOSALS FOR REFORM

In the early years following Title VII's enactment, conscious, deliberate discrimination was quite prevalent, and disparate treatment theory effectively addressed it. But overt forms of discrimination, while they no doubt still exist, are increasingly rare. The conscious, deliberate desire to exclude women and minorities from the workforce has largely disappeared, but forms of intergroup bias stemming from social categorization and the cognitive distortions which inexorably flow from it remain. Our antidiscrimination jurisprudence has failed adequately to address this new type of disparate treatment discrimination, to think rigorously about it, and to fashion doctrine equipped to reckon with it. There is much that we can do, even without amending Title VII, to remedy the various problems we have examined. We now turn to these easily-implemented reforms, then to other potential approaches requiring further study and development.

##### A. *Reformulating Individual Disparate Treatment Theory*

The pretext model of individual disparate treatment proof should be eliminated entirely and replaced with a unitary "motivating factor" analysis similar to, but in certain critical respects different from the framework used in *Price Waterhouse v. Hopkins*. As I have demonstrated, pretext analysis is based on a number of insupportable assumptions and fails to account for cognitive sources of intergroup bias. Most importantly, the false dichotomy between "real reasons" and "phony reasons" embedded in pretext theory fundamentally misapprehends how cognitive processes distort intergroup judgment and decision-making and cause disparate treatment discrimination.

In the vast majority of cases now adjudicated under the pretext model of proof, the nondiscriminatory reason(s) articulated by the employer probably did play an actuating role in the employer's decision. But it does not follow from this that no discrimination occurred. The same employer might have interpreted the same event differently and made a different decision had the target employee been a member of a different social group. To require a disparate treatment plaintiff to disprove the motivating significance of every nondiscriminatory reason articulated (or, after *Hicks* not articulated) by his employer imposes on plaintiffs an almost impossible burden, and will lead to the gross and utterly unjustifiable underidentification of biased employment decisions.

Furthermore, pretext analysis should be eliminated because it logically relies on a presumption of invidiousness that the majority of judicial decisionmakers are, quite justifiably, no longer willing to accept. The fact that an employer made an irrational or inconsistent decision does not mean that the decision was motivated by conscious discriminatory animus. But, given the ubiquity and biasing effects of social stereotypes, the tendency towards schematic information processing, the salience of race, gender, and other social cat-

egories, and the apparent automaticity of ingroup favoritism, it is reasonable to presume in such situations that the employer's decisionmaking was contaminated by cognitive sources of intergroup bias. Accordingly, under the new approach I propose, evidence that an employer's decision was irrational or unjustified in light of the relevant data set would still be important in proving that discrimination occurred, but it would be interpreted differently. No longer would the trier of fact be required to conclude that the employer was lying about the reasons for its decision in order to rule for the plaintiff.

So, for example, in my case against the box manufacturer, much of the same evidence now offered to prove intentional disparate treatment would be used to prove unintentional discrimination. The fact that similarly situated Caucasian employees were treated more favorably than Miguel would be relevant to show that his ethnicity had affected his supervisor's decisionmaking, but not to show that the supervisor's proffered reasons for acting as he did were a cover-up for a "real" discriminatory motive. Similarly, the fact that Miguel was the sole Latino in an otherwise all Caucasian work group would be relevant not to show that management harbored a deliberate desire to exclude Latinos but because of the biasing effect of solo status on social perception and judgment. Comments reflecting ethnic stereotypes would, in such cases, be relevant to show that the plaintiff's group membership had biased a decisionmaker's perception and interpretation of, or memory for, decision-relevant events, but they would not necessarily be offered as proof of discriminatory intent.

Most fundamentally, under the approach I propose, courts would reformulate disparate treatment doctrine to reflect the reality that disparate treatment discrimination can result from things other than discriminatory intent. To establish liability for disparate treatment discrimination, a Title VII plaintiff would simply be required to prove that his group status *played a role* in causing the employer's action or decision. Causation would no longer be equated with intentionality. The critical inquiry would be whether the applicant or employee's group status "made a difference" in the employer's action, not whether the decisionmaker intended that it make a difference. Accordingly, discriminatory intent would constitute one among a number of actionable causes of disparate treatment discrimination.

No amendment to Title VII would be required to implement this aspect of my proposed approach. Section 703(m) of Title VII, added as part of the Civil Rights Act of 1991, already provides that "an unlawful employment practice is established when the complaining party demonstrates that race, color, religion, sex, or national origin was a motivating factor for any employment practice, even though other factors also motivated the practice."<sup>365</sup>

---

365. 42 U.S.C. § 2000e-2(m) (1994). It is important to note that § 703(m) does not limit its motivating factor standard to cases involving "direct evidence" or evidence directly tied to discrimination.

While the terms "motive" and "intent" are often used interchangeably in Title VII caselaw,<sup>366</sup> they are not in fact equivalent, as both Title VII scholars and courts interpreting the Age Discrimination in Employment Act have recognized. As the Seventh Circuit noted in *Burlew v. Eaton Corp.*, "motive is what prompts a person to act, or fail to act. Intent refers only to the state of mind with which the act is done or omitted."<sup>367</sup>

Thus, we can understand the term "motivating factor" as a synonym for "actuating factor"—something which causes a person to act or decide in a particular way. As Part II made plain, a person's group status can bias a decisionmaker's perceptions, judgments, and actions through processes that are quite independent of any invidious intention. Judicial decisions interpreting section 703(m) should not conflate motivation and intentionality; they are distinct constructs and reflect distinct mental processes. Courts should construe "motivation" as "actuation," thereby constructing a more flexible and ultimately more accurate jurisprudential model of disparate treatment discrimination.

To restore Title VII's analytical coherence, and ultimately its normative efficacy, courts should differentiate more clearly between intentional and unintentional forms of disparate treatment discrimination. Intentional discrimination should be understood as being equivalent to the ADEA's concept of "willful" discrimination, and a 2-tier liability system, similar to that currently utilized in ADEA cases, should be grafted onto Title VII.

Under such a system, first-tier disparate treatment liability would attach when the plaintiff proved, by whatever type of evidence he chose, that his protected group status played a role in causing the employer to act or decide as he did. Unless the defendant can thereafter prove that it would have taken the same action or decision even absent the biasing effect of the plaintiff's group status,<sup>368</sup> the disparate treatment plaintiff should be entitled to the same remedies available in a disparate impact case.<sup>369</sup> These include general and individualized injunctive and other equitable relief, including back and front pay and attorneys' fees. Should the plaintiff also prove that the discrimination resulted from the conscious use of his or her group status in the decisionmaking process, he should also be entitled to compensatory and punitive damages as provided by the Civil Rights Act of 1991.<sup>370</sup>

Some of my colleagues in the plaintiffs' bar will no doubt criticize my suggestion that compensatory and punitive damages be unavailable in cases involving unintentional discrimination, but I believe that such limitation is justified and will, in the long run, best serve Title VII's social purpose.

---

366. See text accompanying notes 16-38 *supra*.

367. 869 F.2d 1063, 1066 (7th Cir. 1988) (quoting BLACK'S LAW DICTIONARY 727 (5th ed. 1979)).

368. This portion of my proposal also parallels existing law. See 42 U.S.C. § 2000e-5(g)(2)(B) (1994).

369. See *id.* at § 2000e-5(g)(1).

370. See *id.* at § 1981(a).

First, remedies for disparate impact and cognitive bias-based disparate treatment should be similar because the types of bias giving rise to these two forms of discrimination are in important ways quite similar. For example, a particular selection criterion, such as a height and weight requirement for a police officer job, may be imposed because the relevant characteristic, an imposing physical presence, is representative of the decisionmaker's cognitive prototype of a successful police officer. Similar applications of the representativeness heuristic may bias a decisionmaker's evaluation of an individual candidate or employee and result in disparate treatment discrimination.<sup>371</sup> Thus, schema-driven information processing strategies and related misapplications of the representativeness heuristic can manifest themselves in either disparate impact or cognitive bias-based disparate treatment discrimination.

Additionally, social cognition research reminds us that we are more likely to notice, protest, and strive to alleviate hardships or perceived injustices when they befall ingroup as opposed to outgroup members.<sup>372</sup> In addition, we are often unconscious of our disparate levels of responsiveness in such situations. From these findings, it is reasonable to predict that decisionmakers will less quickly recognize and more readily tolerate the injustice that results from statistically invalid selection criteria when those criteria impact more harshly on outgroup than ingroup members. Certainly, the same "discriminatory empathy" which perpetuates disparate impact discrimination underlies much individual disparate treatment as well as the subtle, often unconscious ingroup preferences that determine whose hardships will be alleviated and whose will not. Again, in the interests of normative coherence, disparate impact and social cognition-based disparate treatment should be understood and remedied in substantially similar ways.

Finally, limiting compensatory and punitive damages to cases involving conscious, deliberate discrimination will best serve the goals of improving intergroup relations and minimizing both cognitive and motivational sources of discrimination. Categorization-related and other cognitive biases in intergroup judgment are unintended and, for many people, earnestly undesired byproducts of essential mental processes and strategies. Attaching moral opprobrium or the risk of substantial financial liability to cognitive intergroup judgment errors can only serve to heighten intergroup anxieties and make racial, ethnic, and gender distinctions more salient. This enhanced salience can, in turn, be expected to exacerbate categorical responding. Accordingly, providing compensatory and punitive damages in cases involving first-tier disparate treatment liability could well prove counterproductive.

---

371. See text accompanying notes 176-179 *supra*.

372. See, e.g., Crosby, Bromley, & Saxe, *supra* note 361, at 548-52; Gaertner & Dovidio, *supra* note 179, at 79; Lauren G. Wispe & Harold B. Freshley, *Race, Sex, and the Sympathetic Helping Behavior: The Broken Bag Caper*, 17 J. OF PERSONALITY & SOC. PSYCHOL. 59, 64-65 (1971).

B. *Adopting a Prescriptive Approach to Nondiscrimination*

Disparate treatment jurisprudence has long treated the nondiscrimination principle as a *proscriptive* duty not to discriminate. However, social cognition theory suggests that the nondiscrimination principle would be more effective in reducing intergroup bias were it understood as *prescriptive* duty to identify and control for errors in social perception and judgment which inevitably occur, even among the well-intended. Accordingly, I agree with David Oppenheimer's suggestions that a negligence approach to discrimination and equal employment opportunity would further Title VII's purpose.<sup>373</sup>

I am not confident, however, that we know enough about how to reduce cognition-based judgment errors to enable us to translate such a duty into workable legal rules. Cognitive psychologists have told us more about the shortcomings of human social inference cognition than about how the various biases they identify can be reduced or controlled.

Professor Oppenheimer proposes that "[w]henever an employer fails to act to prevent discrimination which it knows, or should know, is occurring, which it expects to occur, or which it should expect to occur, it should be held negligent."<sup>374</sup> However, imposing such a duty as a matter of law, it is important to ask precisely what, short of using numerical quotas, an employer should do to prevent discrimination from occurring. In other words, what would an affirmative duty to prevent discrimination entail, and what actions would satisfy it?

Professor Oppenheimer suggests two answers to these questions. First, he proposes that:

When a woman or minority job applicant is rejected, the rejection should act as a triggering device, requiring the decision maker to instantly stop and examine his or her own motives. If the decision cannot be justified with a reasonable, nondiscriminatory reason, such as bona fide qualifications, the decision may have been negligently reached.<sup>375</sup>

Given what we know about attribution bias<sup>376</sup> and the relative inability of social decisionmakers to identify accurately the causal antecedents of their actions and decisions, such an approach is unlikely to be effective. Additionally, given the manner in which stereotypes and other social expectancies bias information processing, few employment decisionmakers will perceive their decisions, whether biased or not, as lacking a reasonable, nondiscriminatory justification.

Professor Oppenheimer further suggests that where an employer has created job screening or employee evaluation procedures that fail to correct for unconscious discrimination, and such discrimination influences the process, the employer should be subject to negligence liability.<sup>377</sup> As I have stated, I agree that this is the direction in which Title VII liability theory must evolve, but I

---

373. Oppenheimer, *supra* note 14, at 900.

374. *Id.* at 969.

375. *Id.* at 970.

376. See text accompanying notes 206-211 *supra*.

377. Oppenheimer, *supra* note 14, at 970.

question whether we are ready to impose such an obligation without better understanding what specific steps employers can take to reduce cognitive bias.

Consider the problem of illusory correlation,<sup>378</sup> the erroneous association of one variable, for example athletic ability, with another, for example race, a doubtless source of bias in the evaluation of employee performance. I could find only one study, conducted in 1967, which investigated how illusory correlation could be reduced through the implementation of specific process controls.<sup>379</sup>

Other empirical studies provide a few additional clues. For example, various experiments indicate that carefully specifying evaluative criteria and providing decisionmakers with a large amount of information on those criteria can reduce intergroup bias.<sup>380</sup> Another study found that white evaluators were less likely to deliver disproportionately harsh sanctions to black targets when they believed that their actions would be subject to review and potential censure by other whites.<sup>381</sup> From this finding, one may infer that the once prevalent equal employment opportunity/affirmative action reviews of employment decisions might have done more good in reducing intergroup bias than policymakers have recognized.<sup>382</sup>

Unfortunately, the most developed body of research regarding the reduction of intergroup bias indicates that many of the conditions which reduce categorical responding lie outside the control of any individual employer. The existence of cross-cutting category boundaries, cooperative interdependence between members of different groups, the content of media and other cultural depictions of women and minority group members, and appeals to categories and values that transcend discrete social groups all appear to lessen intergroup discrimination.<sup>383</sup> But no single employer can integrate a community, alter the

378. See text accompanying notes 152-168.

379. Chapman and Chapman found that permitting subjects to review written materials from which covariation assessments could be made, to make notes, and to arrange the relevant materials in piles led to a statistically significant reduction in illusory correlation on four out of six illusory correlates. Chapman & Chapman, *Genesis of Popular But Erroneous Psycho-Diagnostic Observations*, *supra* note 154, at 202-04. Interestingly, this study also found that providing monetary incentives for accuracy had no significant effect in reducing illusory correlation. We can thus infer that the threat of potential liability alone may not reduce cognitive bias in the assessment of covariation.

380. Veronica F. Nieva & Barbara A. Gutek, *Sex Effects on Evaluation*, 5 *ACAD. OF MGMT. REV.* 267, 270-71 (1980).

381. Edward Donnerstein & Marcia Donnerstein, *Variables In Interracial Aggression: Potential Ingroup Censure*, 27 *J. OF PERSONALITY & SOC. PSYCHOL.* 143, 143-44 (1973).

382. After taking office in 1981, President Reagan dramatically reduced federal efforts to enforce Executive Order 11246, which required government contractors to develop and file written affirmative action plans. Later in the 1980s, the Reagan Administration seriously considered rescinding the Executive Order in its entirety. In response to these and other civil rights policy changes, many companies eliminated their equal employment opportunity/affirmative action departments, which had previously provided institutionalized review of hiring, promotion, and employment termination decisions. See generally THOMAS B. EDSALL & MARY P. EDSALL, *CHAIN REACTION* (1991).

383. See, e.g., Marilyn B. Brewer & Norman Miller, *Beyond the Contact Hypothesis: Theoretical Perspectives on Desegregation*, in *GROUPS IN CONTACT: THE PSYCHOLOGY OF DESEGREGATION* 281, 283-84 (Norman Miller & Marilyn B. Brewer eds., 1984) (Categories are more apt to be salient if they are based on "convergent boundaries" in which group identities based on many different distinctions such as race, economic status, residence location, political persuasion, all coincide. Categorical distinctions are rendered less salient in the presence of "cross-cutting" category boundaries.); Stuart W. Cook, *Coop-*

depiction of minorities in the media, or transform the tone or content of our cultural discourse on race. Upon careful reflection, we may conclude that improving intergroup relations and reducing cognitive sources of intergroup bias is a broad-based cultural problem requiring a broad-based cultural solution.

If our goal is to reduce race, gender, and ethnicity-based categorical responses, the imposition of a duty of care without defining what specific actions an employer should undertake to fulfill that duty could prove counterproductive. As John Calfee and Richard Craswell have suggested, uncertainty about legal standards can produce either overcompliance or undercompliance in the targeted community, depending on a variety of largely unpredictable variables.<sup>384</sup> Perhaps more fundamentally, psychological research suggests that normative uncertainty exacerbates intergroup anxieties and correlates with *increased* levels of intergroup discrimination. These considerations counsel caution: We need additional theoretical, and perhaps even empirical, investigations into how to reduce cognitive sources of bias before we move disparate treatment theory beyond the modest doctrinal reforms I have proposed. Specifying precisely what an affirmative duty of care would comprise and what specific steps an employer should take to satisfy it is a critical task, but one deserving a separate, more comprehensive treatment than it can be accorded here.

## V. CONCLUSION

*Fires in the Mirror*, Anna Deavere Smith's provocative play about racial conflict in Crown Heights, includes these reflections of a New York social worker:

I think you know the Eskimos have seventy words for snow? We probably have seventy different kinds of bias, racism, and discrimination, but it's not in our mind-set to be clear about it, so I think that we have sort of lousy language on the subject and that is a reflection of our unwillingness to deal with it honestly and to sort it out.<sup>385</sup>

It is probably no accident that legal policymakers interpreting Title VII have constructed all disparate treatment discrimination as manifesting a conscious, discriminatory purpose. This interpretation holds the problem of intergroup bias at a safe distance, something those "other people," those "bad people" do. Unfortunately, jurisprudential unwillingness to think rigorously

---

*erative Interaction in Multiethnic Contexts*, in *GROUPS IN CONTACT: THE PSYCHOLOGY OF DESEGREGATION* 15 (Norman Miller & Marilyn M. Brewer eds., 1984) (discussing the potential effectiveness of cooperative interdependence, coupled with cooperative task interaction, in reducing intergroup hostility and accompanying category-based social responding); Henri Tajfel, *Social Categorization, Social Identity and Social Comparison*, in *DIFFERENTIATION BETWEEN SOCIAL GROUPS* 61 (Henri Tajfel ed., 1978); Bernard M. Bass & George Duntzman, *Biases in the evaluation of one's own group, its allies and opponents*, 7 *J. OF CONFLICT RES.* 16 (1963) (finding perceived competition between groups increases intergroup discrimination in the evaluation of ingroup versus outgroup performance).

384. Richard Craswell & John E. Calfee, *Deterrence and Uncertain Legal Standards*, 2 *J. OF L. ECON. & ORGANIZATION* 279, 279-80 (1986).

385. ANNA DEAVERE SMITH, *FIRES IN THE MIRROR: CROWN HEIGHTS BROOKLYN AND OTHER IDENTITIES* 65-66 (1993).

about how and why discrimination occurs has had serious negative consequences. As one scholar has observed, discrimination is not one thing, but many, and the “[f]ailure to recognize this results in intellectual and moral confusion as well as bad policy.”<sup>386</sup> Disparate treatment doctrine has labored too long under a rhetoric of invidiousness that has outlived its social utility. If Title VII is to retain its effectiveness as an instrument for improving intergroup relations, this rhetoric—and its reflection in disparate treatment doctrine—will have to change.

---

386. Larry Alexander, *What Makes Wrongful Discrimination Wrong? Biases, Preferences, Stereotypes, and Proxies*, 141 U. PA. L. REV. 149, 153 (1992).