Washington Law Review

Volume 74 | Number 4

10-1-1999

Unconscious Bias and Self-Critical Analysis: The Case for a Qualified Evidentiary Equal Employment Opportunity Privilege

Deana A. Pollard

UNCONSCIOUS BIAS AND SELF-CRITICAL ANALYSIS: THE CASE FOR A QUALIFIED EVIDENTIARY EQUAL EMPLOYMENT OPPORTUNITY PRIVILEGE

Deana A. Pollard*

Abstract: Recent breakthroughs in social psychology have resulted in the ability to measure unconscious bias scientifically. Studies indicate that prejudiced responses are largely unconscious, the result of normal cognitive processing and stereotypical associations of which the prejudiced subject may be completely unaware. The studies also indicate that a subject's awareness of the discrepancy between her conscious, egalitarian value system and her unconscious prejudice is a critical step towards the convergence of her cognitive functioning and her egalitarian viewpoints. Antidiscrimination legislation requires a showing of intent to discriminate to obtain relief in all but a small percent of circumstances. The result is a legal framework that does not, and cannot, properly redress most instances of discrimination. While the use of unconscious-bias testing may be more effective than antidiscrimination legislation in identifying and redressing the cognitive phenomenon of discrimination, evidence law does not support its use because test results are not privileged against discovery in discrimination lawsuits. This Article argues that in light of the enormous potential social benefit of unconscious-bias testing, a qualified evidentiary privilege should be recognized to encourage its use.

We live in a racially fragmented society. Although the U.S. Supreme Court ruled formal segregation unconstitutional in 1954,[1] our society remains segregated in housing, education, employment, and virtually every other indicator of socioeconomic well-being and status.[2] Racial tension is high and unlikely to change as long as we fail to take seriously the ever-present threat that oppression and injustice necessarily pose to any society.[3] Americans are inundated with race-conscious news items,

* Attorney, Ivie, McNeill & Wyatt, Los Angeles. University of California, Berkeley, LL.M.; University of Southern California, J.D.; University of Washington, B.A. I am grateful to my dear friend Erwin Chemerinsky for his consistently generous advice and support. I also owe thanks to the following people, all of whom gave me input or support in relation to this Article: Mitar Vranic, Laurent Mayali, Linda Krieger, Anthony Greenwald, Gloria Pollard, and Cee Cee, Kris & Kyle Moore. This Article is dedicated to my darling son Ryan Pollard, who at the tender age of two demonstrated angelic patience for his mom's long hours at the computer—I pray that your future includes a more harmonious world.

1. *See* Brown v. Board of Educ., 347 U.S. 483 (1954).

2. *See, e.g.*, David Benjamin Oppenheimer, *Understanding Affirmative Action*, 23 Hastings Const. L.Q. 921, 958–89 (1996).

3. Although I frequently refer to race-based social issues, the same analysis applies to discrimination and attitudes based on gender, culture, religion, and sexual preference. Race-based issues have received substantial public and government attention in recent years and exemplify the problem of automatic stereotyping addressed in this Article. This Article discusses racial and other

work and school sensitivity training, and even race- and ethnicity-related voter propositions. A mentality of "taking sides" along racial lines has emerged with vigor in recent years.

The resultant racial tension finds a forum for expression in high-profile litigation. When high-profile cases involve race, they become a microcosm of society's racial rift. High-profile cases have manifested and intensified the racial divide, and polls indicate that people's perceptions and opinions of these cases are highly correlated with race.[4] Meanwhile, a conservative U.S. Supreme Court has eroded protection of minorities and women in particular, and civil rights generally, over the past twelve years.[5] No doubt inspired by the Court's "majoritarian" treatment of civil rights and deference to government action, plebiscites have begun to dismantle affirmative action programs all over the country, which promises to perpetuate, and probably worsen, racial and cultural tensions.[6]

Employment discrimination certainly contributes greatly to the disparity in socioeconomic status between white men and all other Americans. Gainful employment, after all, is the cornerstone of economic security and is the necessary precursor to adequate housing, health care, and educational and social opportunities. The proliferation of employment

---

forms of discrimination interchangeably, as social psychology research indicates that racial minorities and women suffer similar disadvantages on account of stereotyping or unconscious discrimination. *See infra* Part III.

4. For example, the 1991 Rodney King beating and the riots following the acquittal of the Los Angeles police officers exposed the depth of our country's racial divide. During the 1994–95 O.J. Simpson murder trial, public views on O.J. Simpson's guilt or innocence cut along racial lines. According to a Harris poll of February 11, 1995, 61% of whites believed O.J. Simpson was guilty, while 68% of blacks believed he was innocent. Similar race-based discrepancies emerged in a Los Angeles Times poll taken in September 1994. *See News Analysis: Simpson Jury Could Defy Conventional Wisdom Trial: Majority of Blacks on Panel May Prove to Be a Wild Card Rather Than a Plus for the Defense, Experts Say*, L.A. Times, Nov. 5, 1994, at A1.

5. *See, e.g.*, Erwin Chemerinsky, *The Vanishing Constitution*, 103 Harv. L. Rev. 43 (1989). Justice Kennedy joined the Court in February 1988, and became the fifth vote in a number of conservative opinions hostile to civil rights. *See id.* at 45.

6. The U.S. Supreme Court has allowed the majority to determine minorities' rights issues, contrary to the High Court's traditional antimajoritarian role, which is secured in part by the Justices' life tenure. For example, in Washington State and California anti-affirmative-action initiatives have received popular support, rendering state employers and educational institutions unable to consider race in hiring, contracting, or admissions. *See infra* notes 129–33 and accompanying text. Legal challenge to California's anti-affirmative-action Proposition 209 was summarily defeated in federal court, and the U.S. Supreme Court refused to grant certiorari. *See* Coalition for Econ. Equity v. Wilson, 122 F.3d 692 (9th Cir. 1997), *cert. denied*, 521 U.S. 1141 (1997). *See generally* Linda Hamilton Krieger, *The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity*, 47 Stan. L. Rev. 1161 (1995).

discrimination litigation over the past ten years parallels our country's increased racial tension. Employment discrimination cases are often as hotly contested as the typical divorce case. This intensity of emotion results both from fundamentally different perspectives on the role minority status plays in employment decisions and from the severing of the bonds that formed while the now-adverse parties worked together. Typically, the defendant's indignation at the accusation of discrimination is strikingly juxtaposed with the plaintiff's adamant belief that she was victimized by discriminatory foul play. Inevitably, the loser is shocked by the jury's verdict and left to deal with the burden of disillusionment and heartfelt injustice. Considering that equalizing employment opportunities is a necessary step toward equalizing economic and other opportunities, we cannot begin to resolve this country's racial problems without focusing on the effect of discrimination in employment.

Research on unconscious bias and employment discrimination may provide some insight into the relational dynamics of discrimination. Some legal scholars rely on empirical psychology studies to show that intentional, conscious discrimination is only a small fraction of workplace discrimination and that most discriminatory acts result from unconscious stereotyping and cultural biases that never enter into the decisionmaker's conscious mind—hence the outrage felt by defendants when accused of intentional discrimination.[7] Thus, these scholars argue that the requirement of Title VII of the Civil Rights Act of 1964 (Title VII)[8] mandating that the plaintiff must prove that the defendant acted with intent to discriminate deprives Title VII of the ability to redress the majority of situations involving discrimination. Scientific studies show that uncovering unconscious bias may explain why minorities remain disadvantaged in employment.

Thanks in part to advances in computer technology, recent breakthroughs in social psychology demonstrate that unconscious bias can be objectively measured. The most exciting research shows that unconscious bias may be reversed for subjects holding conscious egalitarian views once the unconscious bias is brought to the subject's

---

7. *See, e.g.*, Krieger, *supra* note 6; *see also* David Benjamin Oppenheimer, *Negligent Discrimination*, 141 U. Pa. L. Rev. 899 (1993). All discrimination claims other than disparate impact cases require the element of purpose or intent to discriminate. *See, e.g.*, Texas Dep't of Community Affairs v. Burdine, 450 U.S. 248 (1981); McDonnell Douglas Corp. v. Green, 411 U.S. 792 (1973); Griggs v. Duke Power Co., 401 U.S. 424 (1971).

8. 42 U.S.C. §§ 2000e–2000h (1994). Title VII is the most comprehensive and commonly litigated federal employment discrimination statute.

attention. Some social psychologists are creating unconscious-bias-testing services that soon will be available to corporate employers. Unconscious-bias testing of employees could be a substantial step toward identifying situations in which discrimination is likely to occur, setting up preventative measures, and actually reversing discriminatory attitudes. If employers were to conduct unconscious-bias testing, they could use the results to counter unconscious tendencies through education or reassignment of biased supervisors. However, because courts recognize no evidentiary privilege protecting employers' unconscious-bias test results from disclosure, employers are unlikely to utilize such tests out of fear of their use in employment discrimination litigation.

This Article argues for the recognition of a privilege for unconscious-bias testing to encourage its use in equal employment opportunity efforts. Solving the problem of unconscious bias is the necessary predicate to a more racially just society. Yet courts have abandoned the "self-critical analysis" privilege for voluntary self-critical analysis in the employment context, a privilege which probably would have protected unconscious-bias testing.[9] To encourage effectively the use of unconscious-bias testing, a new evidentiary privilege is needed to protect the test results from discovery.

Part I of this Article explains that unconscious bias is rampant and may precipitate most discriminatory actions. Part II discusses the inadequacy of current law to prevent and redress discrimination motivated by unconscious bias. Part III argues that employer testing of unconscious bias is essential to prevent the influence of unconscious bias in employment decisions. Part IV explains why an evidentiary privilege is essential to encourage employers to test for unconscious bias. Part V discusses the current law's failure to recognize a self-critical-analysis privilege in employment that would encompass unconscious-bias testing. Part VI argues that a qualified self-critical-analysis privilege for unconscious-bias testing is consistent with privilege theory and advances the policies underlying the formerly recognized self-critical-analysis privilege in the employment context. Part VI also discusses ways to discourage employer abuse of the privilege. Finally, Part VI rebuts objections to unconscious-bias testing based on the First Amendment

9. The self-critical-analysis privilege has been asserted in a number of contexts to prevent discovery of companies' candid self-assessments. The privilege has been recognized in the employment context but has fallen into disfavor in recent years. *See infra* Part V.

right to freedom of beliefs and state and federal constitutional rights to privacy.

I. THE NATURE OF UNCONSCIOUS BIAS AND AUTOMATIC STEREOTYPING

As recently as the early 1980s, psychology researchers considered explicit measures of people's conscious attitudes toward minorities, generally through self-reporting,[10] as reliable evidence of the extent of societal prejudice. Revolutionary insights in social psychology have revealed that a person's conscious replies to prejudice questionnaires are only part of the story, and indeed, how progressive a person appears on the surface may bear little or no relation to how prejudiced she is on an unconscious level.[11]

In the past ten years, researchers have devoted substantial empirical attention to the distinction between explicit and implicit mental processes in relation to racial and other forms of bias.[12] Breakthroughs in understanding how the normal mind categorizes information have allowed researchers to identify an individual's unconscious bias and automatic stereotype activation. Advanced computer technology has enabled testing for biased reactions toward stereotyped minority groups that occur very quickly on a cognitive level beyond conscious control. Although explicit prejudice is on the decline, as measured by progressive societal norms and popular awareness of discrimination, the problem of stereotyping on an implicit level is worse than researchers had imagined.[13]

Some social psychologists believe that stereotyping is a manifestation of "in-group/out-group" dynamics and the human instinct to identify with a group or clan.[14] When traditional groups such as villages and tribes broke down, people were inclined to classify themselves along race and class lines. Part of the nature of "in-group/out-group" dynamics

---

10. *See* John F. Dovidio et al., *On the Nature of Prejudice: Automatic and Controlled Processes*, 33 J. Experimental Soc. Psychol. 510, 511 (1997).

11. *See* Annie Murphy Paul, *Where Bias Begins: The Truth About Stereotypes*, Psychol. Today, May–June 1998, at 52, 53; *see also* Anthony G. Greenwald & Mahzarin R. Banaji, *Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes*, 102 Psychol. Rev. 4 (1995); Telephone Interview with Anthony G. Greenwald, Professor of Psychology, University of Washington (Feb. 14, 1999).

12. *See* Dovidio et al., *supra* note 10, at 511.

13. *See* Paul, *supra* note 11, at 53.

14. *See id.* at 53–54 (quoting New York University Professor John A. Bargh).

is the tendency to see members of one's own group as individuals, but out-group members as an undifferentiated, stereotyped mass.[15]

Stereotypical assumptions activate automatically and implicitly, often contrary to conscious beliefs. Unconscious processing controls the manifestation of "implicit attitudes" and "automatic stereotypes" in outward behavior.[16] Individuals may not be aware of the implicit attitudes, and therefore they may be unable to act consistently with their personal beliefs about how they should feel, or even how they want to feel, about minorities.

Unconscious attitudes are usually exposed with response latency measures of activation, a methodology where the subject is timed while she makes stereotypical and counter-stereotypical associations between target group members and descriptive words.[17] It takes slightly longer (in terms of milliseconds) for an unconsciously biased individual to associate positive descriptive words with a member of a disfavored group than for the individual to associate positive descriptive words with other persons.[18] This discrepancy in time is due to the differences in cognitive barriers in making the connections. Implicit attitudes discovered via latency testing may be empirically unrelated to self-reported explicit prejudice.[19] Thus, people with no conscious bias may be processing information with substantial unconscious bias. Indeed, a discrepancy between response latency measures of unconscious bias and self-report tests of conscious bias is expected to be manifest for socially sensitive issues such as racial attitudes, because despite our collectively held unconscious biases, our societal norms say we should not stereotype people.

The immediate ancestor to automatic stereotyping was the cognitive revolution of the 1970s, in which social scientists' view on how people categorized and processed information changed radically and produced "social cognition theory."[20] After decades dominated by the study of

---

15. *See id.*; *see also* Krieger, *supra* note 6, at 1194–95 & nn.149–51.

16. Implicit attitudes and automatic stereotyping are "introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects." Greenwald & Banaji, *supra* note 11, at 8.

17. For a more detailed description of response latency testing measures, see *infra* Part III.B.

18. *See* Anthony G. Greenwald et al., *Measuring Individual Differences in Implicit Cognition: The Implicit Association Test*, 74 J. Personality & Soc. Psychol. 1464, 1477 (1998).

19. *See, e.g.*, Margo J. Monteith, *Self-Regulation of Prejudiced Responses: Implications for Progress in Prejudice-Reduction Efforts*, 65 J. Personality & Soc. Psychol. 469, 469–70 (1993).

20. *See* Telephone Interview with Anthony G. Greenwald, *supra* note 11.

observable behavior and a premise that prejudice was motivational in origin, scientists began focusing on the more mysterious inner workings of the human brain's processing activity. Computers enabled scientists to measure discrepancies in reaction time down to the hundreds of milliseconds and to display data quickly. At the same time, studies on human cognitive processing were beginning to illuminate how stereotypes were processed. As a result of these advances, social cognition theory holds that (1) stereotyping is part of an essential cognitive process that all people use to help perceive their surroundings efficiently;[21] (2) once constructed, stereotypes unconsciously affect judgment about members of the stereotyped class by operating as "schemas"—expectancies that "fill in" missing information about a person or event to guide our behavior during social interaction;[22] and (3) this process is beyond the decisionmaker's conscious control.[23]

The cognitive approach found that categorization and evaluation are normal and important parts of human intelligence.[24] This rejected the conclusion of European researchers of the Nazi Holocaust who theorized that only repressed, authoritarian persons with internal conflict created by inadequate parenting were guilty of stereotyping.[25] Social cognition theory discovered that categorization allows us to simplify our complex environment and predict future events based on incomplete information—an efficient and generally effective manner of information processing. Prior research indicated that each time we categorize objects or people we progressively intensify our internal stereotypes because we reaffirm the perception that members of a certain category are more similar than they actually are and that members of different categories are more dissimilar than they actually are.[26] Stereotypes result from "categories that have gone too far," using personal characteristics such as race or gender as a proxy for personality traits such as hostility,

21. *See id.*

22. *See* Mark Chen & John A. Bargh, *Nonconscious Behavioral Confirmation Processes: The Self-Fulfilling Consequences of Automatic Stereotype Activation*, 33 J. Experimental Soc. Psychol. 541, 541 (1997).

23. *See* Krieger, *supra* note 6, at 1187–88 & nn.112–14. The Krieger article also contains more information on the cognitive origins of stereotypes. *See id.* at 1190–98.

24. *See* Telephone Interview with Margo J. Monteith, Professor of Psychology, University of Kentucky (May 11, 1999).

25. *See id.*

26. *See* Krieger, *supra* note 6, at 1189.

intelligence, or weakness.[27] This cognitive approach to bias thus established an understanding of the normal human categorization process and created the foundation on which implicit bias and automatic stereotyping theories were built.

The cognitive model holds that the conscious mind connects conspicuous characteristics with stereotypes and that those connections through repetition become unconscious. Although we may not choose to view women as intellectually incompetent sex objects or blacks as violent criminals, the media bombards us with these images, impressing stereotypes onto our cognitive processing that likely affect our behavior.[28] Children are inundated with media and cultural stereotypes before they have the cognitive abilities or life experiences to challenge these images.[29] The profundity of the problem is apparent from psychological research indicating that once these cognitive connections are formed, they are automatically triggered by the slightest interaction with a target group member. From the cognitivists' viewpoint, the cycle perpetuates itself with no conscious choice by the biased individual. Indeed, some psychologists flatly question the concept of free will in light of research on stereotypes, because behavior often follows unconscious beliefs even though conscious and unconscious beliefs may be very different; thus "free will" may be nothing more than the application of unconscious assumptions.[30]

In 1989, social psychologist Patricia Devine challenged the cognitivists' views that stereotypes (unconscious assumptions) and prejudice (consciously held beliefs) were interrelated and that prejudicial behavior was an inevitable consequence of normal cognitive processes.[31] Devine distinguished between a person's cognitive knowledge of a

27. Paul, *supra* note 11, at 53 (quoting New York University Professor John A. Bargh).

28. There is some question about whether automatic stereotyping in fact results in discriminatory behavior. However, most current studies have found that unconscious bias does in fact affect our judgment toward stereotyped group members as well as our resulting verbal and physical unconscious reactions, which in turn create "behavioral confirmation" and perpetuates the cycle of prejudice. *See infra* Part III.A.

29. Note that social psychologists' belief that unconscious bias is produced through popular media messages that create "schemas" in our unconscious minds undermines the "marketplace of ideas" concept underlying the First Amendment. Choice is the cornerstone of the marketplace concept, and if we are not free to choose what messages we internalize, the concept collapses theoretically. *See infra* Part VI.D.2.

30. *See* Paul, *supra* note 11, at 55 (quoting New York University Professor John A. Bargh).

31. *See* Patricia G. Devine, *Stereotypes and Prejudice: Their Automatic and Controlled Components*, 56 J. Personality & Soc. Psychol. 5, 5 (1989).

cultural stereotype and the person's individual beliefs, which may or may not be consistent with the stereotype.[32] She argued that although stereotypes and personal beliefs have some overlapping features, they are conceptually distinct cognitive structures.[33] Her research on human information processing distinguished between automatic responses (mostly involuntary, spontaneous activation of associations developed through repeated activation in the memory) and controlled responses (mostly voluntary, controlled, intentional responses requiring active attention).[34]

Devine posited that because stereotypes are fixed at an impressionable young age, they are more "accessible" as a cognitive function than values and personal beliefs formed later through conscious evaluation.[35] Because members of a society experience common socialization, including exposure to racial and other biases, consciously biased people ("high-prejudiced people") and those who espouse more egalitarian views ("low-prejudiced people") demonstrate equal activation of societal stereotypes under conditions when no time exists for personal beliefs to interfere with the unconscious, automatic response.[36] However, unlike prior researchers, Devine posited that any person whose automatically activated stereotypes and personal beliefs conflict can break the automatic response cycle by inhibiting the old stereotypical beliefs while intentionally activating newer, more egalitarian beliefs.[37] This process concurrently makes the egalitarian beliefs more and more accessible and the prejudiced responses less and less accessible.[38]

Devine's research found that automatic stereotype activation is equally strong for low- and high-prejudiced people and that in the absence of controlled processes both groups exhibited similar stereotype-congruent, or prejudice-like, responses.[39] However, other research, both

---

32. *See id.*

33. *See id.*

34. *See id.* at 6. Devine's dissociation of automatic and controlled responses is consistent with other current conceptions of racial prejudice, suggesting that while traditional forms of prejudice are direct and overt, contemporary forms are indirect and subtle and are therefore harder to identify—even by the racist.

35. *See id.*

36. *See id.*

37. *See id.* at 6–7.

38. *See id.* at 7.

39. *See id.* at 8–12. Devine's methodology of exposing her subjects to a race or gender trait to prime the stereotyping response has been criticized as failing to differentiate between the category of

before and after Devine's 1989 study, suggests that controlled processes can inhibit the effects of automatic processing when the implications of automatic processing conflict with a low-prejudiced person's goal of establishing or maintaining a nonprejudiced identity.[40] Devine disputed the automatic-processing theorists who credited responses that were not consciously monitored over consciously mediated responses for determining the extent of societal prejudice, and she argued that conscious processes were the key to escaping and changing the degree of societal prejudice.[41]

In contrast to skeptics who thought that prejudice and racism were intractable, Devine analogized unconscious bias to a "bad habit" that can be "broken."[42] To break the bad habit of unconscious negative and automatic discriminatory attitudes, a person must consciously activate unbiased beliefs each time a stereotype is automatically triggered.[43] The more the individual does this, the more accessible the unprejudiced belief becomes, and at some point the egalitarian belief's accessibility will "rival" the automatic response.[44] This process requires intention, attention, time, and effort, and, like breaking a bad habit, is no easy task.

Devine's 1989 study was a turning point in social cognition theory. The argument that unconscious attitudes and self-reported prejudice may not be directly related departs from earlier cognitive theories of the creation and perpetration of prejudice. Devine's research has received substantial attention in social psychology literature and has sparked numerous subsequent studies attempting to clarify the nature and relationship of unconscious bias, prejudice, and discriminatory behavior.[45] Yet while it seems intuitive that unconscious bias would affect overt judgments and acts, this is not necessarily clear.[46] Although

---

persons primed and the stereotype primed per se. *See, e.g.*, Lorella Lepore & Rupert Brown, *Category and Stereotype Activation: Is Prejudice Inevitable?*, 72 J. Personality & Soc. Psychol. 275, 283–85 (1997). This impacts whether high- and low-prejudiced people harbor and activate the same stereotypes about minority groups because, as Lepore and Brown found, priming the category of persons produced different stereotype activation than priming the stereotype per se. *See id.* at 284. By priming both, Devine may have confounded stereotype and category automatic activation.

40. *See infra* Part III.A.

41. *See* Devine, *supra* note 31, at 15.

42. *See id.*

43. *See id.* at 16.

44. *See id.*

45. *See, e.g.*, Chen & Bargh, *supra* note 22; Greenwald & Banaji, *supra* note 11; Monteith, *supra* note 19.

46. *See* Dovidio et al., *supra* note 10, at 511–12.

Devine's research suggests that prejudice is akin to a "bad habit" that is amenable to correction by a willing subject once the automatic component is recognized, other research has indicated that unconsciously activated associations and stereotypes influence but do not dictate behavior.[47] Although the link between conscious bias and prejudiced actions is not yet fully understood, Devine's theory is optimistic in positing that while negative attitudes toward racial minorities and other targeted groups nest in our unconscious minds and become activated without our consent or conscious knowledge, racial bias is not inevitable, immutable, or perennial. Exposing and understanding the unconscious component to bias may be a leap toward understanding how to produce nonprejudiced attitudes both explicitly and implicitly.

There is growing inconsistency in the psychology literature about the relationship between conscious attitudes and unconscious, automatic stereotype activation.[48] When given enough time to control their responses, high-prejudiced people will exhibit more stereotyping than low-prejudiced people. The controversy relates to circumstances where high- and low-prejudiced people are thought to be unable to control their exhibited behavior—when purely unconscious, automatic stereotyping manifests itself because the subject has no time to modify her responses based on her personal beliefs. Some researchers have found a significant relationship between conscious attitudes and stereotype activation, and some have not.[49] While high- and low-prejudiced people demonstrate equal knowledge of cultural stereotypes,[50] the groups may differ in the strength of their associations between stereotypic traits and particular cultures.[51] High-prejudiced people are more likely to use stereotypes consistently and repeatedly activate negative stereotypes, resulting in more developed associations that are highly accessible and of sufficient

47. *See infra* note 49.

48. *See, e.g.*, Kerry Kawakami et al., *Racial Prejudice and Stereotype Activation*, 24 Personality & Soc. Psychol. Bull. 407, 414 (1998).

49. The researchers (Devine, Greenwald, Benaji, and others) found that although low- and high-prejudiced people showed significant stereotypic associations for blacks, the effect was stronger for high-prejudiced people, which may reflect a high-prejudiced individual's higher frequency of activating stereotypes as well as greater strength and endorsement (extremity) of stereotypes. *See* Russel H. Fazio & Bridget C. Dunton, *Categorization by Race: The Impact of Automatic and Controlled Components of Racial Prejudice*, 33 J. Experimental Psychol. 451, 468 (1997); Lepore & Brown, *supra* note 39, at 283–84.

50. Both groups are exposed to the same societal stereotypical images via, *inter alia*, the media.

51. *See* Kawakami et al., *supra* note 48, at 413–14.

strength to produce automatic activation.[52] Conversely, low-prejudiced people engage in less stereotyping and develop weaker associations that are cognitively less accessible, resulting in a lesser likelihood of automatically activating cultural stereotypes.[53]

In any event, whether or not consciously held views have a direct affect on cognitive processing, conscious knowledge of unconscious bias unquestionably can initiate conscious and unconscious mental processes[54] that mitigate prejudice, based on the person's knowledge, guilt, and motivation to change her discovered unconscious bias. This motivational approach to lessening the effects of societal, collectively held biases has led a number of social psychologists to believe that a critical first step toward eradicating discriminatory belief systems may be making people aware of discrepancies between their conscious ideals and unconsciously held negative stereotypes.[55] A person who truly wants to be fair and just in relations with others will be motivated to behave in more egalitarian ways upon discovering the discrepancy between her conscious beliefs and her unconscious bias. Social psychologists have developed a number of theories to describe the way in which a person's emotional reaction upon conscious recognition of unconscious bias can motivate that person to control subsequent spontaneous stereotypical responses and behave in less-prejudiced ways in the future.[56] Indeed, conscious efforts to suppress stereotypically biased reactions may per se inhibit automatic activation of stereotypes over time.[57] The greatest potential for creating a just society may lie in our individual willingness to recognize and work toward eradicating our own unconscious biases.

Unconscious-bias testing promises at least to help eliminate unjust discriminatory viewpoints and unlawful discriminatory practices in the workplace. Although we have little or no control over our unconscious bias, company personnel can at least take it into account when making

52. *See id.* at 414.

53. *See id.*; *see also* Lepore & Brown, *supra* note 39, at 283–85.

54. Indeed, the idea that motivation to be a truly fair-minded person may be spurred by awareness of one's unconscious bias relative to one's conscious beliefs, and that the process by which bias and prejudice are reversed may also be a mix of conscious and unconscious mental activity, would predict complicated and divergent empirical results when psychologists attempt to categorize these highly intertwined mental processes as "unconscious and automatic" or "conscious and controlled."

55. *See, e.g.*, Devine, *supra* note 31, at 6; Monteith, *supra* note 19, at 470–71.

56. *See* Monteith, *supra* note 19, at 470.

57. *See* Irene V. Blair & Mahzarin R. Banaji, *Automatic and Controlled Processes in Stereotype Priming*, 70 J. Personality & Soc. Psychol. 1142, 1159 (1996).

important employment decisions. If knowing about unconscious bias disturbs us and challenges our assumptions about ourselves, that negativity may motivate us to change the way we process information on a cognitive level and reverse our cognitive bias over time.[58] As a result, we progress towards more egalitarian treatment of stereotyped groups. We also likely make substantial, albeit painful, progress towards creating justice from where it must originate—in our hearts and minds. Although progress toward achieving racial justice is especially urgent now with the apparent demise of affirmative action, the law presently lacks the tools to resolve the problem of unconscious bias.

II. CURRENT LAW FAILS TO ADDRESS ADEQUATELY THE PROBLEM OF UNCONSCIOUS BIAS

Social scientists and legal commentators agree that existing antidiscrimination legislation, such as Title VII, is ineffective to redress most instances of discrimination.[59] Social psychologists generally agree that many acts of discrimination are perpetrated without conscious knowledge of bias at any point in the decisionmaking process.[60] Exhorting people to cease discrimination is not enough because it fails to affect the organized bedrock of cognitive bias.[61] Equal employment training, political protest, public scorn, and even civil liability work on a conscious level only. Some social psychologists believe that the only way to get rid of stereotypes is to strike at their roots in the unconscious mind.[62] To the extent that disparate treatment liability under Title VII requires proof of intent to discriminate, existing federal law fails to redress the most prevalent form of discrimination—the discrimination that results from unconscious, unintentional bias.[63]

States across the country are dismantling one traditional remedy for discrimination that bypasses cognitive processes—affirmative action.[64] Affirmative action programs designate opportunities to minorities

58. *See infra* Part III.A.

59. *See* Greenwald et al., *supra* note 18, at 1464–65; *see also* Krieger, *supra* note 6; Oppenheimer, *supra* note 7.

60. *See supra* Part I.

61. *See* Paul, *supra* note 11, at 52 (quoting Yale Professor Mahzarin R. Banaji).

62. *See id.* (quoting Yale Professor Mahzarin R. Banaji); Telephone Interview with Anthony G. Greenwald, *supra* note 11.

63. Almost all Title VII cases are disparate treatment cases. *See infra* note 70.

64. *See* Paul, *supra* note 11, at 52.

without judging them in relation to whites and are thus insulated from the perverse influence of unconscious bias. Some studies demonstrate that affirmative action may foster a balanced workforce and interaction with minority group members, both of which are the necessary predicates to eliminating stereotypes.[65] However, because the majority of people (primarily whites) do not want affirmative action,[66] other empirical studies demonstrate that affirmative action may exacerbate intergroup tensions.[67] In any case, debate over the efficacy of affirmative action may be obsolete, considering that it appears to have lost the support of the courts and the public.

A. Title VII

The jurisprudential construction of discrimination, by omitting any recognition of unconscious bias, bears little in common with the actual phenomenon it purports to represent.[68] In most cases, Title VII and state statutes modeled after it[69] require proof of conscious, discriminatory intent to state a claim and obtain relief for employment discrimination.[70]

---

65. *See* Linda Hamilton Krieger, *Civil Rights Perestroika: Intergroup Relations After Affirmative Action*, 86 Cal. L. Rev. 1251, 1274–76 (1998).

66. *See infra* notes 129–30, 151–54, 156–60, and accompanying text.

67. *See infra* notes 146–51 and accompanying text.

68. *See* Krieger, *supra* note 6, at 1217. In *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 801 (1973), the Court appeared to recognize the breadth of Title VII's coverage by stating that Title VII "tolerates no racial discrimination, subtle or otherwise." Nonetheless, in subsequent cases, courts have generally been tolerant of all but the most egregious and overt forms of discrimination.

69. For example, the California antidiscrimination statute is modeled after Title VII, and California courts use federal case law interpreting Title VII to construe the state statute. *Compare* Cal. Gov't Code § 12940 (West 1998), *with* 42 U.S.C. §§ 2000e–2000h (1994). *See* Greene v. Pomona Unified Sch. Dist., 38 Cal. Rptr. 2d 770 (1995) (holding that because state and federal antidiscrimination statutes are identical in their objectives, California courts look to federal law to interpret analogous provisions of state statutes); *see also* Pereira v. Schlage Elec., 902 F. Supp. 1095 (N.D. Cal. 1995). Washington State has a similar approach. *Compare* Wash. Rev. Code § 49.60.180 (1997), *with* 42 U.S.C. §§ 2000e–2000h. *See* Hollingsworth v. Washington Mutual Savings Bank, 37 Wash. App 386, 390, 681 P.2d 845, 848 (1984) (holding that Washington law against discrimination substantially parallels federal law against discrimination embodied in Title VII and thus in construing Washington statute Washington courts look to interpretation of federal law); *see also* Burnside v. Simpson Paper Co., 66 Wash. App. 510, 521 n.9, 832 P.2d 537, 545 n.9 (1992), *aff'd*, 123 Wash. 2d 93, 864 P.2d 937 (1994).

70. The exception is disparate impact cases, which are often prohibitively expensive to bring and constitute less than two percent of Title VII cases, and, in any event, are impossible to bring under most factual employment settings because of the practical need for a substantial group of employees suffering from the same discrimination. *See* John J. Donohue III & Peter Siegelman, *The Changing Nature of Employment Discrimination Litigation*, 43 Stan. L. Rev. 983, 998 & n.57 (1991).

The result is that Title VII is a useless mechanism for redressing what some psychologists deem the most insidious and prevalent form of discrimination: unconscious stereotyping. Title VII's current analysis fails both our minority citizens, who are often unable to prove discrimination because it is defined inappropriately, as well as our corporate employers, who face discrimination lawsuits routinely and are often found blameworthy even when they are not guilty of intentional wrongdoing. Worse yet, society at large takes sides on the issue of employment discrimination and racism. This creates more tension, more attention to the issue, and ultimately more racism.[71] Current law cannot effectively address the problem of unconscious bias.

Traditional constitutional case law and perhaps an unconscious bias among white justices explain the central role of proof of intent in Title VII jurisprudence. Professor Krieger argues that the reason for the antiquated model of discrimination under Title VII is that psychological understanding of the causes of discrimination has evolved radically since the time of the promulgation of Title VII and its state law analogues.[72] However, the U.S. Supreme Court's approach to Title VII may have more to do with its historical approach to discrimination under the Constitution than the state of the art of psychology. In *Washington v. Davis*,[73] the Court clarified that a discrimination claim brought pursuant to the Equal Protection Clause requires proof of intent or purpose to discriminate.[74] Reviewing cases dating back to 1880, the Court demonstrated that the Equal Protection Clause has never protected against disparate impact where discriminatory purpose was not shown.[75] The Court then struck down an extension of the Title VII disparate impact analysis found in *Griggs v. Duke Power Co.*[76] to discrimination cases brought under the Equal Protection Clause.[77] While underscoring the fact that discrimination analysis under Title VII and the Constitution are different, the *Washington v. Davis* Court also emphasized the historical requirement of showing intent in discrimination cases—a requirement

71. The more this controversy is raised in our consciousness, the more we can expect people to take sides and reaffirm positions, exacerbating already-prevalent stereotypes and making them more accessible in the future, which will likely produce even greater racial division and tension.

72. *See* Krieger, *supra* note 6, at 1173–77.

73. 426 U.S. 229 (1976).

74. *See id.* at 239.

75. *See id.* at 239–41.

76. 401 U.S. 424 (1971).

77. *See Washington*, 426 U.S. at 238–41.

that had been entrenched in discrimination analysis for at least eighty-four years prior to the enactment of Title VII.[78] This historical intent requirement, coupled with the fact that before Title VII overt racism was legal and rampant, may explain the Court's unwillingness to recognize little other than intentional discrimination in Title VII analysis. This is in spite of its recognition of a theory of unintentional Title VII liability in *Griggs* and the *Griggs* statement that "Congress directed the thrust of the Act to the *consequences* of employment practices, not simply the motivation."[79] Indeed, at the time Title VII went into effect, most discrimination was overt; it is partly because of Title VII that discrimination has moved underground, making intent harder to prove.[80]

White Americans with political power historically have felt that unjust racial outcomes are acceptable if unintended. The *Washington v. Davis* rule embodies a distinctively white way of thinking about racial discrimination that holds intent to be an essential element of racial injustice.[81] Whites tend to trust in race neutrality more than nonwhites, and, because of their stake in maintaining the racial status quo, they tend

78. *See id.* at 239 (citing Strunder v. West Virginia, 100 U.S. 303 (1880)).

79. *Griggs*, 401 U.S. at 432. Note also that even after *Griggs*, the Court began to retreat from the theory of unintentional liability under Title VII. Justice O'Connor's plurality opinion in *Watson v. Fort Worth Bank & Trust*, 487 U.S. 977, 987 (1988), stated:

> The distinguishing features of the factual issues that typically dominate in disparate impact cases do not imply that the ultimate legal issue is different than in cases where disparate treatment analysis is used. . . . [T]he necessary premise of the disparate impact approach is that some employment practices, adopted without a deliberately discriminatory motive, may in operation be functionally equivalent to intentional discrimination.

(citation omitted). This language seems to indicate a retreat from the *Griggs* decision, which had recognized disparate impact discrimination unrelated to intent and indicates that the disparate impact test was simply an *alternative* way of showing discriminatory intent. In *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642, 654–56 (1989), the Court adopted O'Connor's position in *Watson* and attempted to retreat from the unintentional standard set forth in *Griggs*. However, Congress responded with the Civil Rights Act of 1991, restoring disparate impact analysis to a strictly unintentional standard. *See* Oppenheimer, *supra* note 7, at 935–36.

80. In 1944, 55% of white adults polled felt that whites deserved the first shot at employment opportunities over blacks. *See id.* at 904–05. By 1963, 85% purportedly favored equal employment opportunity, and by 1972, seven years after Title VII took effect, 97% supported equal employment opportunity. *See id.* This shows both that intentional discrimination was a big concern prior to Title VII and that now covert or unconscious discrimination is a bigger problem—if 97% truly believed in equal employment opportunity, discrimination would be much less prevalent than it is.

81. *See* Barbara J. Flagg, *"Was Blind, but Now I See": White Race Consciousness and the Requirement of Discriminatory Intent*, 91 Mich. L. Rev. 953, 968 (1993).

to favor retaining the intent requirement in spite of its demonstrated failure to effectuate substantive racial justice.[82]

Although it is unlikely that social psychology literature has had a meaningful influence on the Court's intent requirement in interpreting Title VII, Professor Krieger's research demonstrates that Title VII analysis has been remarkably consistent with the psychological understanding of discrimination at the time Title VII was promulgated;[83] however, Title VII analysis fails to correspond with more contemporary theories. Recognition of the growing inconsistencies between social psychologists' explanation of how discrimination occurs and current Title VII analysis warrants modifying Title VII's models of liability to redress discrimination more effectively.

Before the promulgation of Title VII in 1963, social scientists had a different view of how discrimination takes place than social psychologists generally accept today. Title VII analysis is consistent with social psychologists' earlier belief that discriminatory acts were a direct result of "prejudice," which was understood as consisting of three components: (1) beliefs about the attitude object, such as a person (the cognitive component); (2) feelings toward the attitude object (the affective component); and (3) behavioral dispositions toward the attitude object (the behavioral component).[84] During the 1970s and 1980s, a two-component model was advanced, in which 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)."[85] Under this two-component approach, the behavior component was viewed as an independent construct called the "behavioral intention," consisting of a consciously formed intent to act toward the attitude object in a particular way.[86] Thus, discrimination was believed to be the intentional behavioral manifestation of prejudice.[87]

Title VII jurisprudence incorporates three assumptions about human inference and judgment that reflect the now-obsolete understanding of

82. *See id.* at 968–69.

83. *See* Krieger, *supra* note 6, at 1173–77.

84. *See id.*

85. *Id.* at 1176. (quoting Jack Levin & William C. Levin, *The Functions of Discrimination and Prejudice* 66 (1982)).

86. *See id.* at 1176.

87. *See id.*

discriminatory processing.[88] First, the U.S. Supreme Court's interpretation of Title VII severely limits its usefulness to redress discrimination because it fails to support claims for discriminatory actions based on unconscious bias. This assumes that intergroup discrimination results from motive or intent to discriminate; by equating intent and causation, stereotype discrimination is considered a product of discriminatory motivation.[89] Thus, federal courts have held Title VII's section 703[90] to require proof of disparate treatment caused by purposeful or intentional discrimination under a disparate treatment theory of liability.[91] The discriminatory intent requirement applies both in "pretext" cases brought under the paradigm of *McDonnell Douglas Corp. v. Green* and *Texas Department of Community Affairs v. Burdine*,[92] and in "mixed motive" cases brought under the 1989 U.S. Supreme Court decision in *Price-Waterhouse v. Hopkins*.[93] Although the mixed-motive paradigm was ostensibly created to allow Title VII plaintiffs to state a claim upon a finding that stereotypes "infected the employer's decisionmaking process," it failed to move away from equating causation and intentionality—requiring conscious, discriminatory animus, and conflating motive, intent, and causation.[94] Thus, Justice Brennan's

88. *See id.* at 1166–67.

89. *See id.* at 1166.

90. Section 703(a) provides that "[i]t shall be an unlawful employment practice for an employer—(1) to fail or refuse to hire or to discharge an 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." 17 U.S.C. § 2000e (1994). Note that this language could reasonably be interpreted as requiring proof of causation without intent, but this is not how the courts have interpreted it. *See* Krieger, *supra* note 6, at 1168.

91. *See, e.g.*, St. Mary's Honor Ctr. v. Hicks, 509 U.S. 502, 516–17 (1993). Note that this interpretation under Title VII is inconsistent with the court-imposed intent requirement of the Age Discrimination in Employment Act (ADEA), in which unconscious application of stereotyped assumptions of age-based inability or performance problems are actionable. *See* Krieger, *supra* note 6, at 1168 (quoting Syvock v. Milwaukee Boiler Mfg. Co., 665 F.2d 149, 155 (7th Cir. 1981)). In contrast to Title VII, the ADEA's language, which is nearly identical to that of Title VII, has been interpreted to differentiate between unconscious bias and conscious discrimination, although this difference may be supported in part by the fact that the ADEA establishes two tiers of liability, one for a violation per se and the other for liquidated (double) damages for "willful" violation. *See* 29 U.S.C. § 626(b) (1994). However, there are a few Title VII cases in which courts have recognized unconscious bias and plaintiffs have won. *See* Krieger, *supra* note 6, at 1169.

92. 450 U.S. 248 (1981).

93. 490 U.S. 228 (1989).

94. *See* Krieger, *supra* note 6, at 1171. In *Price-Waterhouse*, the plaintiff, a senior manager, was denied partnership in spite of a very successful record because of her "macho" personality. *See Price-Waterhouse*, 490 U.S. at 234–35. Her superiors had told her to take a "charm school" course, to "walk more femininely, talk more femininely, dress more femininely, wear make-up, have her

plurality opinion and Chief Justice Rehnquist's dissenting opinion in *Price-Waterhouse* allow plaintiffs to use proof of stereotypical statements to show intent but do not allow plaintiffs to rely on stereotyping per se to support a cause of action.[95]

The second problem of Title VII jurisprudence is that the disparate treatment paradigm limits disparate treatment analysis to an all-or-nothing finding of intentional discrimination. After the plaintiff presents the elements of a prima facie case,[96] the defendant has the burden of simply producing a legitimate, nondiscriminatory reason for the adverse employment decision.[97] The burden then shifts back to the plaintiff to prove by a preponderance of evidence that discrimination motivated the employer to take the adverse action.[98] The plaintiff can meet her burden of proof directly by presenting evidence (such as overt slurs) that discrimination was more likely than not the reason for the adverse action, or the plaintiff can meet her burden indirectly by showing that the legitimate, nondiscriminatory reason proffered is unworthy of credence, or a "pretext" for discrimination.[99] The pretext basis for establishing liability is based on the notion that people generally, and particularly in business settings, do not act arbitrarily. Thus, when an employer's proffered reason is not believable, it is presumed that the decision was based on illegitimate reasons, such as race or gender.[100] This "presumption of invidiousness" permits plaintiff to win without any direct evidence of discrimination.

---

hair styled, and wear jewelry." *Id.* Plaintiff was clearly a casualty of discrimination based on stereotyping, but the Court focused on the employer's conscious state of mind—intent rather than motive. This is in spite of circuit court opinions distinguishing motive (what prompts a person to act or fail to act, which could include stereotypes) and intent (the state of mind with which the act is done or omitted). *See, e.g.*, Burlew v. Eaton Corp., 869 F.2d 1063, 1066 (7th Cir. 1989) (quoting *Black's Law Dictionary* 727 (5th ed. 1979)). For more analysis of the *Price-Waterhouse* plurality opinion, see Kreiger, *supra* note 6, at 1172–73.

95. *See* Krieger, *supra* note 6, at 1172.

96. The elements of a prima facie case are: (1) plaintiff is a member of a protected class (such as female or a racial minority); (2) plaintiff was qualified for the position she held; (3) plaintiff suffered an adverse employment decision (for example, termination); and (4) plaintiff's position remained open or was later filled by someone with similar qualifications. *See, e.g.*, *Burdine*, 450 U.S. at 253 & n.6; McDonnell Douglas Corp. v. Green, 411 U.S. 792, 802 (1973). Of course, the elements are modified depending on the factual circumstances of the adverse employment action.

97. *See Burdine*, 450 U.S. at 248.

98. *See id.*

99. *See id.* at 256.

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

Because overt racism and sexism are on the decline and employers have become sophisticated about discrimination law, direct evidence of discriminatory animus is rare. Many plaintiffs must resort to proving discrimination indirectly by presenting evidence that the defendant fabricated pretextual reasons for the adverse decision in order to hide its true discriminatory purpose. The plaintiff must prove that defendant is not only bigoted, but is lying to the court and jury—a difficult burden to meet when the defendant is unaware of the unconscious bias and a jury can see the real shock and indignation at being accused of bigotry and perjury.

The current analytical paradigm for disparate treatment leaves no room for the reality in most employment discrimination cases in which the employer believes in its legitimate, nondiscriminatory reason for the adverse decision, but in fact operated under the unconscious direction of cognitive bias and stereotyping. The current analysis shortchanges both parties because it creates an absurdly difficult burden of proof for the plaintiff, who will often lose in spite of the discriminatory treatment, and unfairly casts the defendant as villain, often imposing punitive damages to "punish" behavior not engaged in intentionally.[101]

A third fundamental problem with the current jurisprudential model in disparate treatment cases is its assumption that decisionmakers are aware of their biases and consciously consider them when making an allegedly discriminatory decision. For example, in *Price-Waterhouse*, the plurality opinion held that the words "fail or refuse" in Title VII's section 703(a)(1) mean that the protected status must have been a factor in the decision at the actual moment it was made.[102] Justice O'Connor's concurrence would require even more of the plaintiff—a showing of "*direct* evidence of discriminatory animus *in the decisional process*."[103] For O'Connor, "stray remarks" demonstrating stereotyped beliefs made outside of the timeframe of the decisionmaking process are insufficient per se because they do not show that the discriminatory attitude actually motivated the decisionmaker to make the decision at the actual time the decision was made.[104] It is rare for a person affected by unconscious bias to be aware of the bias at all, let alone to manifest it in unambiguous

101. *See* Krieger, *supra* note 6, at 1177–80.

102. *See* Price-Waterhouse v. Hopkins, 490 U.S. 228, 240–41 (1989); Krieger, *supra* note 6, at 1182–84.

103. *Price-Waterhouse*, 490 U.S. at 278 (O'Connor, J., concurring) (emphasis added).

104. *See id.* at 277; *see also* Krieger, *supra* note 6, at 1183–84 & n.88.

terms at the time of the decision. Indeed, greater awareness of one's unconscious bias lowers the risk that the bias will manifest in discriminatory words or acts.[105] The proof requirements for showing decisionmaker bias in mixed-motive cases are nearly impossible for plaintiffs to meet, and they reflect a fundamental misunderstanding of the human cognitive process.

Not only is Title VII ineffective to redress the real life phenomenon of discrimination, but current disparate treatment theory may actually exacerbate discriminatory animus and intergroup tension. Due to the intent requirement and all that entails, the plaintiff's necessary pleadings generate defensiveness and resentment.[106] People dislike accusations of intentional wrongdoing when they think they have done nothing wrong. This normal human response to an accusation of intentional discrimination, when no such intent was actually involved and when coupled with litigation costs, engenders distrust of minorities claiming discrimination. Title VII's convoluted analytical history (also a result of the "lack of fit" between Title VII's models of liability and the discriminatory process it attempts to redress)[107] has contributed to resentment toward Title VII plaintiffs because it is confusing and fails to instruct defendants properly on how to avoid liability.[108] This resentment causes people to "notice" race and other protected statuses more, making protected statuses more salient and thereby exacerbating the vicious cycle on both conscious and unconscious levels. In sum, Title VII is failing both practically and normatively.

Professor Krieger suggests a number of reforms to Title VII analysis to deal more appropriately with the more common form of employment discrimination—cognitive bias. First, she advocates eliminating the pretext model of individual disparate treatment and replacing it with a "motivating factor" analysis similar to that set forth in *Price-Waterhouse*.[109] Evidence of stereotyping, the biasing effect of solo status (for example, being the only Latino in a company), and better treatment

105. *See supra* Part III.A.

106. That is, the plaintiff must produce evidence of race- or gender-inappropriate language and deeds to demonstrate intent. This produces feelings of betrayal on the part of the defendant, who often honestly does not believe her off-color remarks are relevant.

107. For a thorough analysis of the incoherence of Title VII's liability theories originating from their failure to recognize the unconscious aspect of most instances of discrimination, see Krieger, *supra* note 6, at 1218–37.

108. *See id.* at 1239–40.

109. *See id.* at 1241–43.

of similarly situated Caucasian employees could result in liability without proof of intentional wrongdoing or proof that the employer is lying about its reasons for the adverse employment decision. Second, courts should differentiate between intentional and unintentional forms of disparate treatment by setting up a two-tiered liability scheme for "willful" versus unintentional discrimination, similar to that set forth in the Age Discrimination in Employment Act (ADEA).[110] These are possible ways to redress unconscious bias immediately because they depart analytically from the all-or-nothing proof-of-intent scheme currently in place but do not require legislative amendment to the statute.

David Benjamin Oppenheimer has similarly addressed Title VII's ineffectiveness and the need for a different model of liability. Professor Oppenheimer has suggested a negligence model of discrimination liability under Title VII.[111] Professor Oppenheimer argues that under both the *Griggs* disparate impact theory and the *Albemarle Paper Co. v. Moody*[112] less-discriminatory-alternative theory, employers' Title VII liability is already implicitly based on a negligence standard.[113] Thus, Title VII liability can arise either as a result of the negligent adoption of an employee selection device that has a discriminatory impact not required by business necessity, or Title VII liability can arise as a result of the negligent adoption of a discriminatory device that may be justified by business necessity but is not the least-discriminatory alternative.[114] By explicitly reaffirming these standards of liability in the Civil Rights Act of 1991, Congress clarified that intentional discrimination was only part of the conduct Title VII seeks to redress and that Title VII is concerned with the consequences of discrimination, which are the same regardless of motive or intent.[115] This lends support to the theory that Title VII liability could be explicitly based on a negligence theory.[116]

In addition, cases establishing an employer's duty to accommodate protected class members and to prevent workplace harassment demonstrate that liability need not be premised on intent, or even disparate impact, but rather on a negligence-type failure to take

110. *See id.* at 1243; *supra* note 91.

111. *See* Oppenheimer, *supra* note 7, at 900.

112. 422 U.S. 405 (1975).

113. *See* Oppenheimer, *supra* note 7, at 931–36.

114. *See id.* at 931.

115. *See id.* at 935–36.

116. *See id.*

affirmative steps to prevent harassment.[117] Oppenheimer concludes that employers who treat minorities and women less favorably than white men breach their duty not to discriminate and should be held liable under a negligence theory.[118] Oppenheimer thus proposes that when an employer knows or should know that discrimination is occurring, the employer should be held negligent if it fails to act to prevent the discrimination.[119] Negligence liability should similarly attach when an employer breaches its duty to avoid making employment decisions by means that have a discriminatory effect—for example, by failing to find less-discriminatory alternatives and failing to examine its own motives for evidence of stereotyping.

Oppenheimer presents two examples of an employer's affirmative duty to prevent negligent discrimination. First, when a protected class member is denied employment, the rejection should act as a "triggering device" requiring the employer to review the decision for possible discriminatory motives.[120] However, research has shown that people are unable to define the reasons for their unconscious-bias decisions, based upon "ultimate attribution error," "aversive racism," and other ways of denying our own prejudice, part conscious and part unconscious.[121] Social psychology research indicates that most people operating under unconscious bias will fail to identify their bias. However, objective unconscious-bias testing exposes the existence and probable impact of unconscious bias on personnel decisionmaking, helping the employer to recognize discriminatory attitudes that may result in discriminatory conduct. Unconscious-bias testing thus enhances and makes legitimate the application of a negligence standard of liability under Title VII by giving the employer objective data which in turn gives rise to an affirmative duty to act to prevent discrimination. In other words, while

117. *See id.* at 936–69.

118. *See id.* at 967.

119. *See id.* at 969.

120. *Id.* at 970.

121. People are often completely unaware of their own motivation, and when confronted with reasons for their actions they will state reasons consistent with their actions, although those stated reasons may be completely divorced from their real unconscious motivation. *See* Telephone Interview with Margo J. Monteith, *supra* note 24; *see also* Krieger, *supra* note 6, at 1206–07 & nn.206–11, 1245 & n.376. The theory of "Aversive racism" explains the phenomenon whereby people create unprejudiced, rational reasons for their actions because they are unwilling to believe their actions were motivated by prejudice. *See, e.g.*, Samuel L. Gaertner & John F. Dovidio, *The Aversive Form of Racism*, in *Prejudice, Discrimination, and Racism* 61–69 (John F. Dovidio & Samuel L. Gaertner eds., 1986).

aversive racism and other tendencies to disbelieve discrimination is happening might prevent a finding that an employer "knew or should have known," unconscious-bias-testing results establish employer knowledge, and if combined with employer failure to act, negligence. Second, Oppenheimer suggests that when an employer has created procedures that fail to correct for unconscious discrimination, the employer should be subject to liability for negligent discrimination if discrimination results.[122] However, inadequate guidance exists on what specific steps employers can take to avoid the effects of unconscious bias. At present, without an objective measure of unconscious bias and an understanding of how to mitigate its effects, employers would have to "guess" at both the source of bias and the means to mitigate it effectively. Thus an imposition of an affirmative duty is unworkable and could even exacerbate intergroup anxiety and discrimination.[123] Unconscious-bias testing, however, would provide employers with objective data on which to justify removing decisionmaking authority over minority employees and applicants from specific biased persons, thus lessening the potential for unfair evaluations or discipline. An employer's failure to take appropriate action based on test results would constitute negligence.

As proposed herein, the privilege for unconscious-bias testing is qualified and would be lost by failure to take corrective action in response to test results. Thus, although the privilege intends to encourage unconscious-bias testing with the expectation of confidentiality, its higher purpose is to encourage more egalitarian employment practices. So to the extent an employer fails to act reasonably in response to test results, it will lose the privilege's protection and the test results could then be used as proof of knowledge and failure to act, that is, negligence.

Unconscious-bias testing may thus provide notice to employers that discrimination is likely to occur without intervention, creating a duty to act to prevent discrimination and supporting a negligence theory of liability. In addition, if and when unconscious-bias testing becomes widely used, the relationship between test results and discrimination will become more clear, which will help to define further employers' duty of care and perhaps also lead to additional ways to identify discrimination objectively. Of course, while unconscious-bias testing may provide a solid basis on which to base negligence liability, simply recognizing a

122. *See* Oppenheimer, *supra* note 7, at 970.

123. *See* Krieger, *supra* note 6, at 1246–47.

negligence theory of liability under Title VII will encourage unconscious-bias testing, as it is one way in which employers can exercise care to avoid discrimination.

B. Affirmative Action

Other than employment discrimination lawsuits, affirmative action has been the primary tool for redressing and controlling employment discrimination. This antidiscrimination device may be particularly effective to combat unconscious bias because, as social psychologist Mahzarin R. Banaji puts it, it "bypasses our unconsciously compromised judgment."[124] That is, affirmative action is more effective than other means of preventing discrimination because it does not depend upon the employer's attempting not to consider race. As the research demonstrates, race is constantly considered, if only unconsciously. Because attempting to "ignore" race equates to disadvantaging racial minorities, it is necessary to consider race simply to counter the ubiquitous effect of racial bias and to create equal opportunity.[125] Theoretically, affirmative action should remain in place until research demonstrates that racial bias, both conscious and unconscious, has been virtually eradicated from our society—a far cry from the current empirical statistics. Affirmative action works to create more equal opportunities; the loss of affirmative action will mean perpetuation of discrimination and an unjust status quo.

However, the empirical evidence regarding the effect of affirmative action on intergroup relations is conflicting. Compelling research indicates that, at least in certain situations, affirmative action exacerbates intergroup tensions, increases stereotyping, and creates a sense of injustice among whites.[126] On the other hand, other empirical research has found that a racially balanced workforce involving cooperative, individuating contact is necessary to eliminate intergroup bias—a racial balance that can be achieved quickly through affirmative action.[127] In any case, the debate over affirmative action may be on the brink of obsolescence, as affirmative action is currently threatened with extinction, leaving the questionable efficacy of Title VII with its

---

124. Paul, *supra* note 11, at 52 (quoting Yale Professor Mahzarin R. Banaji).

125. *See* Krieger, *supra* note 65, at 1287–88 (arguing that because we are unaware of our unconscious bias, we must consider race to avoid discriminating).

126. *See infra* notes 146–51 and accompanying text; *see also* Krieger, *supra* note 65, at 1258–76.

127. *See* Krieger, *supra* note 65, at 1275–76.

"colorblind" approach to discrimination as the only mechanism left to fight discrimination in employment.[128] The reasons for affirmative action's recent institutionalized rejection underscore the need for alternative proactive measures to control discrimination in general and the need for unconscious-bias testing in particular.

That affirmative action is on its way out is apparent; it has lost the support of both the courts and the majority of Americans. For example, in 1996, California voters passed Proposition 209 and amended their state constitution to preclude affirmative action in public education, contracting, and employment.[129] Washington State voters approved virtually identical statewide legislation in 1998, making affirmative action illegal in state and local government employment, contracting, and education.[130] Other states have already begun to follow suit, encouraged no doubt by the fact that California's Proposition 209 survived legal challenge, culminating with the U.S. Supreme Court's refusal to grant certiorari.[131] Proponents of similar federal legislation will also be encouraged,[132] and in conjunction with the Court's recent "unprecedented assault" on affirmative action, it appears at this time to be destined for extinction.[133]

---

128. *See infra* note 142.

129. On November 5, 1996, Proposition 209 won California voter approval by a narrow margin, amending the California Constitution. *See* Dave Lescher, *Battle over Prop. 209 Moves to the Courts*, L.A. Times, Nov. 7, 1996, at A1. The amended section now reads: "The state shall not discriminate against, or grant preferential treatment to, any individual or group on the basis of race, sex, color, ethnicity, or national origin in the operation of public employment, public education, or public contracting." Cal. Const. art. I, § 31.

130. Virtually indentical to California's Proposition 209, Washington State's Initiative 200, which provided for a ban on preferential treatment based on race, ethnicity, and gender in state and local government employment, contracting, and education, went to the Washington voters on November 3, 1998, and passed by a margin of 58% to 42%. *See Initiative 200—New Battle Begins: Interpreting Law*, Seattle Times, Nov. 4, 1998, at B1.

131. *See* Krieger, *supra* note 65, at 1254–55 & nn.3–7.

132. *See, e.g., id.* at 1255 & n.8.

133. Erwin Chemerinsky, *What Would Be the Impact of Eliminating Affirmative Action?*, Symposium on Race Relations in America, 27 Golden Gate U. L. Rev. 313, 314 (1997). In 1978, the U.S. Supreme Court first ruled on affirmative action in *Regents of the University of California v. Bakke*, 438 U.S. 265 (1978), and the Court held without a majority opinion that "set asides" for minority students were impermissible. In 1989, the Court held that strict scrutiny was the proper test for evaluating a city's race-based affirmative action program. *See* Richmond v. J.A. Croson Co., 488 U.S. 469 (1989). The Court extended that holding to federal government affirmative action programs in 1995. *See* Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 213–37 (1995) (overruling Metro Broad., Inc. v. FCC, 497 U.S. 547 (1990), which held that intermediate scrutiny was the proper test for "benign" classifications, that is, race-based affirmative action). Race-based affirmative action programs thus face a presumption of unconstitutionality and must meet the strict scrutiny test of

The goals of affirmative action have not been met. The statistics that originally gave rise to affirmative action remain compelling. Affirmative action was born out of the factual reality that minorities historically have faced and continue to face considerably inferior public education, underrepresentation in programs of higher education, limited access to housing and lending (both mortgage and business), vastly different treatment by the health care community (resulting in increased risk of a variety of dangerous and terminal medical conditions), and a disproportionate likelihood of living in poverty, or at least, with substantially less economic power than white men.[134] Similarly, women continue to earn substantially less than men, yet pay more for products and services and face objectively proven discrimination in health care and economic opportunities.[135] Indeed, women of all races earn substantially less than men. White women earn less than black men, and minority women face a "synergistic" discrimination based on the intersection of their minority and gender statuses, causing them generally to earn substantially less than either minority men or white women.[136] There is simply no question that gross disparities continue to exist between white males and all other groups in terms of employment opportunities, economic advantage, political power, and almost every other indicator of social status.

Civil rights activism of the 1960s increased awareness of the depth and breadth of racial and gender injustice and led to efforts to combat it, including affirmative action. The arguments for affirmative action include remedying past discrimination,[137] enhancing diversity,[138]

---

being narrowly tailored and necessary to achieve a compelling government interest, the same test to determine the constitutionality of allegedly race-based discrimination. Note, however, that consistent with the intermediate scrutiny applied to claims of gender-based discrimination, *see* Craig v. Boren, 429 U.S. 190 (1976), the intermediate-scrutiny standard is the test for gender-based affirmative action plans. *See, e.g.*, Orr v. Orr, 440 U.S. 268 (1979); Califano v. Webster, 430 U.S. 313 (1977).

134. *See, e.g.*, Chemerinsky, *supra* note 133, at 313–14; Oppenheimer, *supra* note 2, at 958–96.

135. *See* Oppenheimer, *supra* note 2, at 966–73, 978–89.

136. *See id.* at 966–73; *see also* Evelyn Nakano Glenn, *Cleaning Up/Kept Down: A Historical Perspective on Racial Inequality in "Women's Work,"* 43 Stan. L. Rev. 1333, 1333–34 n.5 (1991).

137. *See, e.g.*, United States v. Paradise, 480 U.S. 149 (1987) (upholding court order for one-for-one hiring and promoting of blacks and whites until effects of past discrimination were eradicated as appropriate measure for past discrimination); Fullilove v. Klutznick, 448 U.S. 448 (1980) (upholding federal law setting aside public-works money for minority businesses based on congressional finding of history of discrimination in construction industry).

138. In *Bakke*, 438 U.S. at 314, Justice Powell argued that colleges and universities have a compelling interest in having a diverse student body. Applying intermediate scrutiny, the Court has also held that government preferences to minority businesses in licensing of broadcast stations are permissible because they enhance diversity of viewpoints. *See* Metro Broad., Inc. v. FCC, 497 U.S.

providing role models,[139] increasing the political power of minorities,[140] and generally enhancing wealth, services, and opportunities to minorities.[141] Arguments against affirmative action include claims that it is unconstitutional, as the constitution requires "colorblind" application of the laws;[142] it is unfair because it is not based on merit;[143] and it harms innocent white persons, stigmatizes racial minorities, and increases racial tension.[144]

While these arguments may prove obsolete in light of the U.S. Supreme Court's tacit support of the majority's ability to vote out

---

547 (1990), *overruled by Adarand*, 515 U.S. 200. In *Adarand*, the Court overruled *Metro Broadcasting* and held that strict scrutiny is the proper test for federal-government-imposed affirmative action. *See Adarand*, 515 U.S. at 227. While the Court has not ruled that diversity may never constitute a compelling interest, at least one circuit has found that diversity cannot be a compelling interest in education based on *Adarand* and *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469 (1989). *See* Hopwood v. Texas, 78 F.3d 932, 945–46 (5th Cir. 1996).

139. The Court, however, has held that providing role models is not a valid basis for making race-based employment decisions. In *Wygant v. Jackson Board of Education*, 476 U.S. 267, 284 (1986), the Court declared unconstitutional a lay-off system in which white teachers with more seniority were laid off ahead of black teachers based upon the assertion that black teachers were needed as role models for the students.

140. This argument has been unsuccessful in justifying re-drawing election districts to provide racial minorities with more political power. *See, e.g.*, Bush v. Vera, 517 U.S. 952 (1996); Shaw v. Hunt, 517 U.S. 899 (1996); Miller v. Johnson, 515 U.S. 900 (1995); Shaw v. Reno, 509 U.S. 630 (1993); *see also* Erwin Chemerinsky, *Making Sense of the Affirmative Action Debate*, 22 Ohio N.U. L. Rev. 1159, 1164–65 (1996).

141. *See* Chemerinsky, *supra* note 140, at 1166–67.

142. *Id.* at 1171–72. The "colorblind" concept that race should not be used to benefit minorities any more than it should be used to discriminate against minorities finds support in Justice Harlan's famous dissent in *Plessy v. Ferguson*, 163 U.S. 537 (1896). Justice Harlan stated that "[o]ur Constitution is color-blind, and neither knows nor tolerates classes among citizens.... The law regards man as man, and takes no account of his surroundings or of his color ...." *Id.* at 559. Of course, the *Plessy* majority's "separate but equal" doctrine was ruled unconstitutional in *Brown v. Board of Education*, 347 U.S. 483 (1954). Justice Harlan's quote has been interpreted to mean that "benign" or "reverse" discrimination based on race, that is, *any* race-based classification including affirmative action plans, is unconstitutional as contrary to the "color-blind" mandate unless it meets the test of strict scrutiny. *See, e.g., Richmond*, 488 U.S. 469. For an interesting discussion on why "colorblind" analysis cannot work in light of empirical research indicating that normal human cognitive processing is far from colorblind, and in fact, a colorblind approach to intergroup relations is dangerous and will aggravate intergroup relations, see Krieger, *supra* note 65, at 1276–91. *See also* Chemerinsky, *supra* note 133, at 318–21 (arguing that at times Constitution *requires* that we consider color); Neil Gotanda, *A Critique of "Our Constitution is Colorblind,"* 44 Stan. L. Rev. 1 (1991).

143. *See* Chemerinsky, *supra* note 140, at 1172–73. For an analysis of the construct of "merit" from a social-psychological perspective, see Krieger, *supra* note 65, at 1291–302.

144. *See* Chemerinsky, *supra* note 140, at 1171–75; *see also* Krieger, *supra* note 65, at 1258–76.

affirmative action,[145] they are relevant to understanding how affirmative action has come under attack and how affirmative action's legacy supports the need for alternatives such as unconscious-bias testing. The most compelling argument against affirmative action is that it is fundamentally inefficacious; it harms minorities by damaging intergroup relations more than it helps, with the attendant increased discriminatory attitudes towards minorities which is manifested, *inter alia*, in voting out affirmative action. The irony is that the reason affirmative action has taken so much heat is the same reason it is so desperately needed. The practical reality, however, is that as long as most people harbor negative feelings about affirmative action and its beneficiaries, the net result may very well be that it hurts minorities (and society in general) more than it helps. Understanding why affirmative action has come under so much attack is thus particularly important to creating and evaluating alternatives.

Research regarding the potential effectiveness of affirmative action is contradictory and unclear. Some empirical research is fairly straightforward in supporting the view that using minority preferences exacerbates intergroup bias: majority group subjects who believe employers use preferences in selecting minority and women employees view the employees (that is, affirmative action beneficiaries) as less qualified and less capable.[146] This perception of affirmative action beneficiaries tends to entrench and confirm out-group stereotypes, which are very difficult to change even in the face of disconfirming data.[147]

---

145. That is, the Supreme Court's refusal to grant certiorari regarding Proposition 209 essentially gives voters the ability to decide whether affirmative action is acceptable, which means majority vote controls whether minorities are given equalizing opportunities through affirmative action. This is consistent with the Rehnquist Court's majoritarian paradigm but is philosophically contrary to the very purpose of the Court as a nonmajoritarian entity to protect fundamental rights embodied in the Constitution against majority rule. The Framers explicitly distrusted majority rule and created the Constitution primarily to shield certain rights from majority control. *See* Chemerinsky, *supra* note 5, at 74–77.

> The very purpose of a Bill of Rights was to withdraw certain subjects from the vicissitudes of political controversy, to place them beyond the reach of majorities and officials and to establish them as legal principles to be applied by the courts. One's right to life, liberty, and property . . . and other fundamental rights may not be submitted to vote; they depend on the outcome of no elections.

West Va. Bd. of Educ. v. Barnette, 319 U.S. 624, 638 (1943). Affirmative action's purpose is to counteract the unfair deprivation of employment and other opportunities, and it should not be subject to majority rule.

146. *See* Krieger, *supra* note 65, at 1264–65.

147. *See infra* Part III.B; *see also* Krieger, *supra* note 125, at 1267–70.

Perhaps even worse, the use of preferences likely increases the salience of the out-group characteristic preferred, which in turn increases attention to the characteristic, increases the number of negative attributions, and ultimately reinforces stereotypes and intergroup bias.[148] Basically, the extent to which affirmative action beneficiaries succeed is often attributed to preferential selection, not merit, and negative stereotypes are thus perpetuated. However, other research shows that where relatively equal numbers of members of different groups exist, group membership distinctions become less salient, and cooperative interdependence tasks and social interaction can reduce stereotyping.[149] Because affirmative action can result in a more equally balanced workforce quickly, it can create environments in which out-group characteristics are rendered less salient and stereotyping is reduced.[150] Nonetheless, because the majority of Americans believe that affirmative action is unfair, unnecessary, or counterproductive, this belief probably trumps affirmative action's potential. In spite of affirmative action's potential to create more harmony among various societal groups, the fact that most Americans do not want it renders it considerably less effective and possibly even counterproductive.

Ironically, the reasons why most Americans do not want affirmative action are based on the same type of intergroup dynamics that gave rise to the need for affirmative action in the first place. In-group/out-group psychology, coupled with factual ignorance about minorities' plight in America, has resulted in disdain for affirmative action, which has led to its demise.

Social psychology research indicates that people's perception of fairness turns on whose treatment is judged as fair or unfair.[151] When members of one's own group are rewarded for superior performance, subjects tend to favor merit-based rewards.[152] However, when an out-group member performs better, subjects tend to want to equalize rewards between the in-group and out-group members.[153] The result is that one's own group member benefits as much as possible under either set of

148. *See* Krieger, *supra* note 65, at 1267, 1274–75.

149. *See id.* at 1274–76.

150. *See id.* at 1276.

151. *See id.* at 1297.

152. *See id.*

153. *See id.*

circumstances.[154] The logical extrapolation of this research is that the fact that affirmative action's very purpose is to benefit out-group members probably creates a sense of injustice on the part of the majority in-group. It is thus not surprising that the majority of Americans feel this country has "gone too far" in pushing equal rights. So, in addition to all of the ways in which unconscious bias, not to mention just plain racism, affect the majority's perception of minorities, in-group/out-group dynamics militate against the majority's viewing racial justice as "fair."

In conjunction with the human tendency to harbor biases against minorities and to favor in-group members at the expense of out-group members, Americans are woefully ignorant about racial statistics. Interestingly, popular conceptions about affirmative action reflect a dissociation between Americans' subjective beliefs and factual reality strikingly similar to individuals' dissociation between consciously held racial beliefs and unconscious bias. The striking similarity lies in the oblivion in which people in our society form opinions on racial issues and form judgments about out-group members. In other words, society at large has views about affirmative action and discrimination that are grossly divergent from reality, a divergence reminiscent of the dissociation between conscious beliefs and unconscious bias first empirically proven by Patricia Devine.[155] This divergence has created impatience and hostility towards minorities in general and resistance to affirmative action in particular.

A few statistics relating to the divergence between people's views on American treatment of minorities and the reality of the American minority experience are illustrative of the ignorance with which people form opinions and cast votes on racial issues. For example, minorities are considerably lacking in medical insurance compared with whites, are less likely to receive expensive medical treatment for the same illnesses, and generally receive inferior health care services even controlling for ability to pay. Yet, sixty-four percent of whites surveyed in 1995 believed that blacks' access to health care is as good or better than whites' access to health care (only thirty percent responding that blacks' access is worse), and fifty-one percent believed access to health care for Hispanics is better than for white Americans.[156] In the employment context, blacks make considerably less money than whites as a whole (at least twenty to

154. *See id.* at 1297–98 & n.160.

155. *See supra* notes 31–47 and accompanying text.

156. *See* Oppenheimer, *supra* note 2, at 981–84.

thirty percent less), are more than four times less likely to receive job offers than white applicants, are disproportionately laid off in times of economic recession, and had an unemployment rate two to three times higher than whites between 1970 and 1990 (with the rate generally increasing over the twenty-year period).[157] Nonetheless, white perception is quite the contrary. Polls taken in 1995 indicated that a majority (fifty-eight percent) of whites believe that the average American black is as well off as the average American white in terms of jobs, and seventy percent of whites believe that blacks are better off than whites in terms of risk of losing their jobs.[158]

Fundamentally, Americans understand neither why affirmative action is necessary nor the fact that it simply counters discrimination, equalizing employment opportunities as opposed to giving minorities an unfair and undeserved advantage. A Newsweek poll taken in 1995 found that twice as many respondents believed that whites were losing out because of affirmative action than believed blacks were losing out because of discrimination.[159] Other polls found that fifty-four percent of whites believe that America has "gone too far" in pushing equal rights; thirty-seven percent of whites believe that whites are losing out to minorities in the workplace due to unfair preferences and that this is a bigger problem than minorities' facing discrimination and lack of opportunity for advancement.[160] But these statistics are more a reflection of unconscious bias and in-group/out-group dynamics than a cause for not supporting affirmative action. That is, simply educating people about statistics cannot change people's attitudes towards minorities per se. This is because even with corrected factual understanding, the psychological processes behind the attitudes will not be affected, and people will tend to apply stereotypes (such as that black people are lazy) to account for the statistical disparity—one form of "aversive racism."[161]

Underlying this white perspective that affirmative action "benefits" minorities to the detriment of whites is the "transparency

157. *See id.* at 966–73; *see also* Chemerinsky, *supra* note 133, at 315–16; Oppenheimer, *supra* note 7, at 914–15. Note that one study on high-priced restaurants in Philadelphia found male applicants five times more likely to receive job offers than female applicants. *See* Krieger, *supra* note 65, at 1303 & n.187. There is simply no question that both minorities and women face extensive employment discrimination.

158. *See* Oppenheimer, *supra* note 2, at 972 & nn.325–26.

159. *See id.* at 958–59 & n.217.

160. *Id.*

161. *See supra* note 121.

phenomenon":[162] whites do not "notice" their own race or the "property interest"[163] being white automatically confers. Whites are constantly advantaged on account of their skin color but are so accustomed to this advantage that it is transparent; whiteness is invisible to whites. Whites' unawareness of their own race causes them to overlook their automatic head start in all social realms. By failing to see how whiteness puts them ahead already, affirmative action is viewed as upsetting the equal playing field rather than leveling it. Whites view themselves as unfairly losing out due to affirmative action, instead of unfairly winning without it. They therefore mistakenly view the noble concept of "colorblindness" as a truly equalizing paradigm rather than the perpetuation of an unjust status quo.

As unjustified as most Americans' views may be, they help explain whites' lack of support for affirmative action and buttress the argument that affirmative action is bad for society and exacerbates intergroup relations. Regardless of affirmative action's potential, if the majority feel that it is unfair, the majority will target affirmative action beneficiaries with increased negative stereotypes, minorities will become more aware of their minority status, and the majority will focus more on individual minorities—all of which entrenches and exacerbates the social phenomena that create discriminatory attitudes and actions.[164] Thus, considering most Americans' opinions about affirmative action, it may increase stereotyping and discrimination. Paradoxically, the loss of affirmative action will result in fewer minorities and women in the workplace, making their "token" presence more noticeable and their minority status more salient, thus increasing the risk of stereotyping. Furthermore, because studies indicate that juries have a more difficult time discerning discrimination in individual cases as opposed to in the aggregate, employment discrimination litigation is even less likely to be effective.[165]

Not coincidentally, the key to understanding both the need for affirmative action (in addition to other proactive measures to combat racial injustice) and our individual biases is reconciling perception with

---

162. *See generally* Flagg, *supra* note 81.

163. *See generally* Cheryl I. Harris, *Whiteness as Property*, 106 Harv. L. Rev. 1709 (1993).

164. *See* Krieger, *supra* note 65, at 1267, 1274–75.

165. *See id.* at 1305–09 (describing research indicating that when subjects viewed aggregate data of compensation differentials between male and female employees, subjects were significantly more able to discern discrimination than when presented with specific case information only).

reality. The first step in changing our perceptions is to recognize that they are not reality. This recognition is necessary to begin the process of creating more egalitarian beliefs from within,[166] which in turn will become manifest in less stereotyping, better intergroup relations, and more support for affirmative action or other proactive measures of combating discrimination. This recognition can be achieved through unconscious-bias testing.

In sum, regardless of the success of some affirmative action programs[167] and affirmative action's potential for improving intergroup relations, as long as the public perceives affirmative action as unfair or unnecessary and the courts allow the public to determine its applicability by popular vote, affirmative action cannot and will not be an effective means for redressing and controlling discrimination and may even exacerbate intergroup tension and discrimination. Unconscious-bias testing is an alternative that potentially not only helps control biased decisionmaking in employment settings but also educates individuals about their own and others' bias. Recognition of our own unconscious bias precipitates changes in attitudes toward minorities and ultimately should drastically change the public's perception of the need for affirmative action and other types of proactive programs, while concurrently contributing to their efficacy.[168]

166. *See infra* Part III.A.

167. For example, in 1975, the year after Governor Reagan made California's affirmative action program official, high-salaried California public employees were 90% white. The minority classifications of African American, Asian American, and Hispanic each constituted less than 3% of the high-salaried positions. By 1993, less than 70% of these positions were held by whites, with African Americans occupying 9.3% of the positions, Asian Americans 9.9%, and Hispanics 11.7%. *See* Rebecca LaVally et al., California Office of Research, *The Status of Affirmative Action in California* 35 (1995).

168. One other way in which individual unconscious-bias testing positively contributes to people's attitudes is that it creates a sense of personal responsibility, precluding the human tendency to "diffuse" responsibility. Psychologists have long known that when in groups, people tend to assume someone else will take responsibility for a necessary action (such as reporting a fire or crime), a concept known as "diffusion of responsibility." Such diffusion is particularly likely to occur among people who do not believe they contributed to the need for action. In discussing affirmative action, most of us have heard the argument: "Affirmative action is unfair . . . *I* never owned slaves . . . why should I be punished?!" By exposing the subjects' individual bias, and hopefully concurrently educating the subjects that they can change if they want to, unconscious-bias testing makes diffusion of responsibility less likely to occur and forces people to confront the role they play in the perpetuation of racial injustice.

III. EMPLOYER TESTING OF UNCONSCIOUS BIAS IS ESSENTIAL TO ERADICATING EMPLOYMENT DISCRIMINATION

A. The Benefits of Unconscious-Bias Testing

The benefits of unconscious-bias testing are many. Our minds are programmed to create and perpetuate stereotypes, and unconscious-bias testing is a way of combating unfair, negative stereotypes in a number of important ways. The potential societal benefit of unconscious-bias testing cannot be overstated.

Our minds naturally stereotype. Successfully interacting in a society requires that we categorize information to retrieve efficiently the right information at the right time, such as, for example, when discerning from a distance a dog from a bicycle. Such categorization is considered a normal and necessary component of efficient cognitive functioning.[169] Stereotypes result from categorizing and constitute schematic knowledge structures that facilitate the encoding of expectancy-congruent rather than incongruent information in memory.[170] This means that stereotypes "fill in" missing information and are self-perpetuating, in that data supporting the stereotype is encoded while data not supporting the stereotype is discarded. For example, when people learn something about another person that is not consistent with their preconceived beliefs (such as a soccer player with a towering IQ), people tend to employ strategies to undermine the impact of information that "disconfirms" the stereotype. People may ignore discrepant information or consider it based on situational causes. The result is that stereotypes are maintained in the face of disconfirming data.[171] Furthermore, "forgetting" stereotypes is a "resource demanding mental operation," counterproductive to the efficiency for which they were created in the first place, and therefore stereotypes are not amenable to change without intervention.[172]

---

169. *See* C. Neil Macrae et al., *On the Regulation of Recollection: The Intentional Forgetting of Stereotypical Memories*, 72 J. Personality & Soc. Psychol. 709, 711 (1997).

170. *See id.* "Schemas" fill in missing information about a person or event based on information presented to generate expectancies about what is going to happen next. These expectancies guide our behavior during social interaction so we can respond appropriately to social situations. *See* Chen & Bargh, *supra* note 22, at 541.

171. *See* Macrae et al., *supra* note 169, at 717.

172. *Id.* at 713–14.

The way the mind processes memory also discourages stereotype change. Stereotypes are easy to remember and, in part based on the ease of retrieval, perniciously hard to forget. Indeed, attempting to "forget" stereotypes may lead to a "rebound" effect in which the stereotype one tries to avoid is actually promoted.[173] This is because stereotypes are so easy to call up, easy to reinforce, and are called up frequently, and, as a result, take tremendous mental resources to "forget." In sum, stereotypes are created as part of our brain's normal operation and are programmed in such a way that they are particularly resistant to change.[174]

Tests uncovering the degree to which individuals engage in unconscious stereotyping benefit society in two general ways. First, simply identifying persons with extreme tendencies to stereotype arms employers with information necessary to prevent this unconscious stereotyping from manifesting in discriminatory actions in the workplace. Job responsibilities calling for subjective analysis of employees' attitude or performance could be taken away from persons demonstrating extreme unconscious bias. Alternatively, employers could create a "check" on personnel decisions and actions taken by these people to ensure that unconscious bias is not occurring in the decisionmaking process. While such modifications to the employment decisionmaking process may be far from perfect, at least the pernicious problem of stereotyping is exposed and considered rather than hidden from all but those who feel its painful effects.

Second, and probably more important in the long run, recognition of unconscious bias may initiate mental processes that actually stop stereotyping at its origin.[175] In spite of pessimism concerning controlling and ameliorating societal stereotypes, recent studies investigating the degree to which a person's motivations and knowledge of her own stereotypes can cause her to overcome automatically produced behavioral effects are encouraging. Scientists are beginning to develop theories about how stereotypes can be "deautomatized."[176] In addition to the obvious and immediate benefit of being able to use unconscious-bias testing to achieve more fair personnel practices, several theories exist

173. *See id.* at 717.

174. *See id.*

175. *See infra* notes 181–220 and accompanying text.

176. *See infra* notes 181–220 and accompanying text.

that support unconscious-bias testing as a step in actually eradicating unconscious bias from where it originates.[177]

Thus, fundamental reasons why unconscious-bias testing is beneficial to society exist apart from the instant ability to assure that personnel decisionmakers are not harboring unfair bias and manifesting it against certain employees. These reasons fall into two major related categories: (1) actually reversing unconscious stereotyping and (2) interfering with a well-known interpersonal dynamic called "behavioral confirmation" through feedback about how unconscious bias affects our outward behavior. Regarding the first category, recent research has indicated that even unconscious bias can be reversed once it is identified. For example, a series of psychology studies imply that when people exhibit more prejudiced responses than they deem appropriate based upon their conscious beliefs, they begin a process of conforming their exhibited behavior to their feelings about appropriate behavior, at least when they experience guilt as a result of discovering their unconscious bias.[178] Psychologists call subjects' unconscious responses that are more prejudiced than the subjects' conscious standards "prejudice-related discrepancies," and few if any researchers would disagree with the conclusion that these discrepancies are common.[179]

Both high-prejudiced and low-prejudiced subjects are prone to these discrepancies. The main difference between the two groups is that low-prejudiced people view their low-prejudice standards as highly important to their self-concept, experience strong negative consequences of transgressing their own standards (such as guilt and self-criticism), and are committed to responding consistently with their standards.[180] High-prejudiced people do not have well-internalized personal standards regarding prejudice. Thus, they are not motivated in the same way as low-prejudiced people are to bring their actions into conformity with their belief system. This is in spite of the fact that they showed even greater prejudice in unconscious testing than they admitted to in conscious testing, and thus also have prejudice-related discrepancies.

According to Patricia Devine's research, prejudiced responses persist among many low-prejudiced people because of spontaneous,

177. *See infra* notes 181–220 and accompanying text.

178. *See, e.g.*, Devine, *supra* note 31; Kawakami et al., *supra* note 48; Lepore & Brown, *supra* note 39; Monteith, *supra* note 19.

179. *See* Monteith, *supra* note 19, at 469.

180. *See infra* notes 193–99.

unintentional stereotypes that are highly accessible knowledge structures that can be automatically activated even if they are not actually endorsed by the person.[181] Conscious beliefs or attitudes about stereotyped groups are less accessible and often inconsistent with the stereotypical associations stored in memory. In order to conform one's behavior with one's low-prejudiced belief system, one must inhibit spontaneous stereotype-based responses and deliberately replace them with belief-based responses.[182] Devine theorizes that achieving control over prejudiced responses after internalizing low-prejudiced beliefs[183] is no easy task and appears to be a gradual process much like breaking a habit.[184]

Devine's theory relies on prior research about the way we stop behavior we dislike. In the early 1980s, researcher J.A. Gray created the "behavioral inhibition system" (BIS) neuropsychological model of motivation and learning, to explain the mechanisms involved in learning to inhibit discrepant responses that have resulted in aversive past events, such as guilt.[185] This model holds that when an unexpected or aversive event occurs (a "mismatch" or "discrepancy" that causes guilt or self-criticism), the subject's arousal is heightened, and an automatic, momentary pausing or interruption of the behavior takes place, which is similar to an orienting response.[186] Then, the sequence of responses occurring when the discrepancy was detected is "tagged" with a "faulty, needs checking" indicator and is given enhanced attention.[187] In addition, the subject engages in "exploratory-investigative behavior," searching for indications of the discrepant response.[188] Thus, Gray argued that the enhanced attention coupled with the exploratory-investigative process enabled the subject to identify stimuli and responses that predict the aversive event.[189] Developing response-contingent punishment "cues" is

181. *See* Devine, *supra* note 31, at 15–16; *see also supra* notes 31–47 and accompanying text.

182. *See* Monteith, *supra* note 19, at 469.

183. The lack of motivation for high-prejudiced people means they are unlikely to engage in the steps necessary to achieve low-prejudice responses, because the first step to making change is internalizing less-prejudiced conscious beliefs.

184. *See* Devine, *supra* note 31, at 15–16.

185. *See* Monteith, *supra* note 19, at 470.

186. *See id.*

187. *Id.*

188. *Id.*

189. *See id.*

crucial for acquiring the ability to inhibit future discrepant responses.[190] The BIS is activated whenever cues that were previously associated with response-contingent punishment are present, so that the individual responds with greater restraint (for example, more slowly) for the purpose of executing a more desirable response.[191] In this way, people can self-monitor prejudiced responses provided they are aware of them and are motivated to do so and eventually break the habit of prejudice and conform their responses to their conscious belief systems and self-image.[192]

Based upon this theoretical framework about the human ability to change automatic responses to conform with more-ideal behavior, Devine and other researchers believe that the first step in breaking the habit of prejudice is to make the subject aware of a prejudice-related discrepancy, an awareness that can be achieved through unconscious-bias testing.[193] Social psychology researcher Margo Monteith postulates that prejudice-related discrepancies should facilitate the prejudice-reduction process among low-prejudiced individuals.[194] She studied low- and high-prejudiced subjects who were led to believe that they had engaged in prejudice-related discrepant behavior (in this case judging a gay law school applicant more harshly than a heterosexual applicant) and determined that, indeed, the results provided "clear, converging evidence that the discrepancy experience did engage these self-regulatory mechanisms."[195] In other words, Monteith's research results were consistent with the theory that once people believe they are engaging in unfair, discriminatory behavior, they will begin the process of modifying their own behavior—provided of course that they believe such behavior is wrong.

In a second experiment, Monteith found that the experience of a prejudice-related discrepancy (again regarding subjects' self-assessed inappropriate bias against gay men) improved low-prejudiced subjects' ability to inhibit prejudiced responses and to respond consistently with their personal standards.[196] The results showed that the discrepancy

190. *See id.* at 470–71.

191. *See id.* at 471.

192. *See id.*

193. *See* Devine, *supra* note 31, at 15; *infra* note 201 and accompanying text.

194. *See* Monteith, *supra* note 19, at 469–70.

195. *Id.* at 477.

196. *See id.* at 482.

experience produced negative self-directed affects among low- (but not high-) prejudiced subjects, who then engaged in heightened self-focus and a preoccupation with their personal prejudice-related discrepancy. Theoretically, this would result in the subjects' eventually gaining control over their discrepant responses.[197] Monteith has explained this process by the following real life example: Suppose you are at a party and someone makes a racist joke and you laugh.[198] Then you feel guilty about having laughed at the joke, and you become focused on your thought processes. All sorts of cues then become associated with laughing at the racist joke, including the person who told it, the circumstances of being at a party, drinking, and other factors. The next time you encounter these cues, a warning signal sounds so that you remember your guilt from before and therefore slow down your responses and use greater restraint later in a similar situation. Monteith calls this process "recruiting," which is gathering conscious, deliberate choices about one's own behavior to supplant the offensive automatic response—a process that requires self-awareness and dedication to egalitarian values.[199]

According to this model, the initiation of the self-regulatory mechanisms should produce slower, more controlled, and more careful responses in future situations when prejudiced responses are possible. Thus, it is possible for people to inhibit prejudiced responses that are based on spontaneous stereotype activation and replace such responses with belief-based responses.[200]

It is critical to this process that subjects recognize that their responses are discrepant from their personal beliefs. As Monteith stated:

> [A] potential first step in promoting change among these individuals may entail heightening their awareness of the discrepancy between their prejudiced tendencies and their egalitarian self-images. Such discrepancy experiences may then encourage subjects to ascribe greater importance to their personal standards for responding to stereotyped groups and to embark eventually on the stages of change described herein.[201]

197. *See id.*

198. *See* Telephone Interview with Margo J. Monteith, *supra* note 24.

199. *Id.*

200. *See* Monteith, *supra* note 19, at 477.

201. *Id.* at 483 (citation omitted).

Thus, unconscious-bias testing may be a first step in helping people to conform their behavior to the type of unbiased person they want to be.[202] The ultimate impact to society from people's behaving in conformity with their righteous beliefs has metamorphic potential for creating social justice.

Other psychological research suggests that there may be scientific methods for "deautomizing" stereotypical responses. Recent analyses of stereotype processing have proposed that if a subject deliberately processes nonstereotypical information, she may be able to overcome automatic processes that create stereotypic responses.[203] In other words, increasing a subject's attention to nonstereotypical information may decrease stereotypic responses, which could make it possible to create an intentional strategy that directs attention toward counterstereotypic information and thereby moderates stereotypic priming.[204] That is, the use of counterstereotypic expectancies may operate to disconfirm a stereotype by facilitating the processing of counterstereotypic information.[205]

In one experiment, researchers sought to discover whether it was possible to overcome and moderate stereotyped responses by testing subjects' response time and accuracy in pairing up words with masculine or feminine target names.[206] Each subject was given either a "stereotype strategy" or a "counterstereotype strategy."[207] For the stereotype strategy, the researchers advised the subjects that if the initial word was stereotypically masculine (for example, ambitious) they should expect the target to be a male name (for example, Patrick).[208] If the first word was stereotypically feminine (for example, perfume), they should expect the target to be female (for example, Lisa).[209] They were told that, most

202. Of course, reducing prejudice through the use of careful self-regulation would be impossible if the subject chose to rationalize nonprejudiced justifications for her responses, which is an "easier" way to deal with the cognitive dissonance created by behaving in ways that are inconsistent with one's belief system. John F. Dovidio and Samuel L. Gaertner's theory of "aversive racism" states that many people are unwilling to recognize their prejudiced responses, so they generate nonprejudiced rationalizations for the responses. *See* Gaertner & Dovidio, *supra* note 121.

203. *See generally* Blair & Banaji, *supra* note 57; Monteith, *supra* note 19.

204. *See* Blair & Banaji, *supra* note 57, at 1149.

205. *See id.*

206. *See id.* at 1150.

207. *Id.*

208. *See id.*

209. *See id.*

of the time, the first word and target name would match in their gender association, so expecting stereotypes would help them predict upcoming events and improve their speed and accuracy.[210] The researchers presented a 5:3 ratio of stereotypic to counterstereotypic trials.[211] For the counterstereotypic strategy, the subjects were told that if the first word was stereotypically masculine, they should expect the target name to be female and vice versa.[212] The same standards for speed and accuracy were used, and the subjects received a 3:5 ratio of stereotypic to nonstereotypic trials.[213] Thus, these subjects were motivated to maintain a counterstereotypic intention because expecting counterstereotypes improved their performance.

The researchers found that when cognitive constraints were low (meaning the subjects had time to control their responses), the subjects with stereotype strategy produced strong stereotype priming, and those with counterstereotype priming produced a complete reversal of stereotype priming.[214] This data suggests that the subjects were able to eliminate stereotypic responses with an intentional strategy when cognitive constraints were relatively low. This result is consistent with current theories regarding conditions in which controlled processes can override automatic processes to determine the outcome. This is significant because the automatic operation of semantic and stereotype associations is generally believed to be based on long-term learning and therefore not vulnerable to intervention.[215] Yet, the researchers were able to demonstrate that when people have an intentional strategy to expect counterstereotype associations, the basic automatic stereotype priming was completely reversed.[216]

Even more surprising were the results when subjects operated under high cognitive constraints, meaning the amount of time given to complete the association tasks is generally believed to be too short to allow for controlled processes and thus the response is considered

210. *See id.*
211. *See id.*
212. *See id.* at 1150–51.
213. *See id.*
214. *See id.* at 1154.
215. *See id.*
216. *See id.*

unconscious or "automatic."[217] Even under such constraints, subjects with an intentional counterstereotype strategy were able to moderate stereotype priming significantly.[218]

The results powerfully support the concept that stereotype priming can be controlled through use of intentional strategies. These findings are significant in exposing the role that other task components may play in revealing an automatic versus controlled response—for example, the perceiver's intention, the strength of the underlying association, and the perceiver's motivation to maintain the strategy.[219]

While this research is new and not yet well understood, it potentially opens the door to using scientific means to fight racism and prejudice. Understanding the relationship between the unconscious and conscious mind is key, considering the various studies indicating that most prejudice is not conscious. This type of research provides some hope that persons who are interested in reversing their own unconsciously held discriminatory attitudes may soon turn to scientists, who may be able to provide cognitive exercise therapy to reverse the prejudice that inundates us all from childhood.[220]

The second major category in which unconscious-bias testing may benefit society lies in its potential to moderate another well-documented problem with stereotypes: they change the perceiver's actions, which in turn affects the target group member's actions. This concept is known as the "self-fulfilling prophecy" or the "behavioral confirmation" effect, and it is the most widely studied expectancy effect in social psychology because of its enormous importance in real-life settings such as the work place, and because its effects are a compelling example of the impact of

---

217. Previous research has indicated that a stimulus onset asynchrony (SOA)—which is the time participants are given to engage, focus, and commit attention to the prime before the onset of the target—of less than 500 milliseconds sufficiently constrains cognitive resources to reveal an automatic process. *See id.* The subjects in this study were given 2000 milliseconds SOA and 350 milliseconds SOA for the low cognitive constraints and high constraints, respectively. *See id.* In another experiment, the same researchers lowered the SOA to 250 milliseconds and again found that the subjects may be able to significantly moderate gender stereotyping under such high cognitive constraints. *See id.*

218. *See id.*

219. *See id.* at 1157.

220. As Margo Monteith puts it, children do not have a choice about accepting or rejecting societal stereotypes about minorities and women because they are acquired well before children have the cognitive abilities or life experiences to form their own beliefs. These stereotypes come from cultural influences, media images, and people. *See* Telephone Interview with Margo J. Monteith, *supra* note 24.

cognitive/perceptual processes in social interaction.[221] As explained herein, we use schemas to anticipate what a stereotyped group member will be like, how he or she will react, and how we should behave.[222] Our anticipatory behavior toward the group member influences how the group member responds; in essence, the anticipatory behavior by the person operating under the stereotype influences and causes the target group member to engage in the "expected" behavior, thus confirming the stereotype regardless of the stereotype's accuracy.[223] These behavior confirmation processes, which often begin with an implicit association, provide a powerful explanation for how stereotypes and discriminatory actions are justified and propagated. While the expectancy may be conscious, both the source of the expectation (for example, subconscious stereotyping) and the perceiver's role in producing the confirmatory behavior are not conscious.[224] This makes it "'particularly difficult to convince the perceiver that his or her stereotypic beliefs are wrong'" because the respondent's actions reaffirm the original stereotype.[225]

The mental processing steps involved in behavioral confirmation are as follows: First, the group stereotype is the source of expectancies or "provisional hypotheses" about individual members of that group.[226] The perceiver then behaves toward the target as though these beliefs were true.[227] These usually negative expectancies then affect the perceiver's behavior toward the target person in a variety of ways.[228] As social

221. *See* Chen & Bargh, *supra* note 22, at 541–42.

222. *See supra* note 170.

223. *See* Chen & Bargh, *supra* note 22, at 542.

224. *See id.* at 543.

225. *Id.* at 544 (quoting D.L. Hamilton & T.K. Troiler, *Stereotypes and Stereotyping: An Overview of the Cognitive Approach,* in *Prejudice, Discrimination, and Racism* 150 (John F. Dovidio & Samuel L. Gaertner eds., 1985)). Note also that although the phenomenon of behavioral confirmation is widely accepted, some researchers have questioned its validity. *See id.* at 544–45. The criticism is twofold: first, that the passive social-perceptual activities of the perceiver regarding the target can be modified with the proper motivation (for example, to be accurate); and second, that experimental expectancies are assumed to be false whereas in real-world settings, the perceiver's expectancies may be usually accurate. *See id.* Laboratory studies may thus maximize the confirmatory effect and not disconfirmatory effects. *See id.* However, as noted by Chen and Bargh, these criticisms do not apply to the nonconscious model of behavioral confirmation because the behavioral consequences of social perception can produce behavioral confirmation automatically as a result of stereotype activation, in the absence of any false information to experimental participants about their target-partners. *See id.* at 545.

226. *Id.* at 542.

227. *See id.*

228. *See id.*

psychology researchers Mark Chen and John A. Bargh state, "The target responds to the perceiver's behavior in kind (for example, with hostility and coldness begetting hostility and coldness) or even actively conforms to the perceiver's apparent opinion so as not to disrupt the interaction (for example, playing the 'stupid foreigner' in order to get one's visa approved)."[229] Finally, according to Chen and Bargh, "the perceiver interprets the target's behavior in line with the expectancy and encodes yet another instance of stereotype-consistent behavior."[230] These processes thus provide a powerful mechanism for maintaining, propagating, and justifying stereotypes and prejudiced behavior.[231]

The concept of behavioral confirmation has been accepted by the social psychology community for more than twenty-five years. For example, a 1974 study found that Caucasians who interviewed African Americans and Caucasians displayed different interview styles toward the two groups: when interviewing African Americans, the interviewers took less time, made more speech errors, and treated the interviewees with less urgency.[232] An application of the two objective interview styles to a group of all Caucasian interviewees demonstrated that Caucasians subjected to African-American-style interviews performed more poorly than in Caucasian-style interviews.[233] Race initially affected the interview style, but once the style was produced, anyone subjected to it would respond less positively—that is to say, anyone would conform behaviorally to the expected behavior as manifested in the interview style.[234] This illustrates the cycle of racism and its tangible and objective effects.

More recent studies have shown that the original expectancies flow from unconscious bias that the interviewer cannot control absent understanding and taking steps to correct for the bias.[235] Unconscious bias is thus a powerful mechanism for perpetuating and continually reaffirming prejudice and exerts a ubiquitous impact on social interaction.[236] Exposing unconscious bias, therefore, has the potential to

229. *Id.*

230. *Id.*

231. *See id.* at 543.

232. *See id.* These objective differences were coded by "blind" raters. *See id.*

233. *See id.*

234. *See id.*

235. *See id.*

236. *See id.* at 543 & n.2.

undermine powerfully unhealthy and unfair societal practices and to break down the machinery that reproduces discrimination continually.

In 1997, Chen and Bargh took the behavioral confirmation model a step further, positing that the link between the perceptual and behavioral representations may be entirely unconscious.[237] Thus, in contrast to the traditional model in which behavioral confirmation is mediated by biased information-gathering and perceptual processes, such as strategically adopted behavioral strategies, Chen and Bargh's approach hinges on an automatic, implicit link between perception and behavior.[238] That is, the stereotype activation that results from perceiving a target group member causes unconscious, automatic behavior consistent with the content of the activated stereotype.[239] Hence, the entire sequence from the first cue (such as skin color) that activates a stereotype to its final confirmation is automatic and unconscious.

Chen and Bargh conducted experiments that supported all aspects of their unconscious model of behavioral confirmation. They tested whether the automatic activation of African-American stereotypes directly produced behavioral confirmation effects as to Caucasian participants.[240] In sum, the African-American stereotypes were activated unconsciously, through subliminal exposure to young male African-American faces during a computer task.[241] However, because the participants were all Caucasian (NYU students), there was nothing in the experimental situation which could have produced conscious stereotype activation and expectancy-confirmation processes, such as actually interacting with African Americans in the experiment.[242] When Caucasian targets interacted with Caucasian perceivers who had been subliminally exposed to African-American faces (with the attendant "hostility" stereotype), their level of hostility increased due to an increase in the perceiver's hostility resulting from the subliminal priming.[243] These findings extend the psychological understanding of the effects of unconscious stereotype activation, their effects on interracial relations, and their resistance to change.

---

237. *See id.* at 554–56.

238. *See id.* at 545.

239. *See id.*

240. *See id.* at 548.

241. *See id.* at 555.

242. *See id.* at 555–56.

243. *See id.*

In real life, Caucasians' unconscious activation of a false stereotype may create the hostile response by African Americans who accurately perceive the hostile expression or other indications of the Caucasian's unconscious stereotype activation and react to it, which in turn reaffirms the "validity" of the stereotype. Ironically, the perceiver creates the "evidence" that "confirms" the stereotype.[244] Because the behavioral confirmation model relies on an initial stereotype or "expectancy," uncovering unconscious expectancies theoretically should critically undermine the entire dynamic, considering that the expectancy constitutes the first part of the process and also considering the effect of exposing people to their own prejudice-related discrepancies.[245]

In sum, in addition to providing better information about our own unconscious bias, which enables us to make conscious efforts to prevent it from manifesting in illegal and discriminatory actions, testing for unconscious bias may be the initial step toward reversing unconscious bias and the interpersonal dynamics that result from it. The potential benefits are thus varied and many.

Employers concerned about workplace discrimination should be allowed to conduct tests for unconscious bias for the purpose of analyzing its relationship to their employment practices (that is, self-critical analysis) and for taking steps to produce a discrimination-free work environment. Social psychologists have just begun to unravel the mystery of societal prejudice and to identify where prejudice is born and how and why it has survived, and indeed flourished, throughout history. If there is any merit to the current psychological understanding that bias is "locked" in our unconscious minds and exposing it is "key" to beginning the healing process, then the benefit of encouraging unconscious-bias testing is potentially socially transformative.

B. Methods of Testing Unconscious Bias

Professors Anthony G. Greenwald of the University of Washington and Mahzarin R. Banaji of Yale University are leaders in the field of unconscious-bias testing. They have created a method that is potentially useful in diagnosing a wide range of socially significant associative structures, which they refer to as "implicit association testing," or

244. *See id.* at 555.

245. For a discussion of additional studies relating to intergroup bias, see generally Oppenheimer, *supra* note 7.

"IAT."[246] They have conducted research to appraise the IAT method's usefulness in measuring evaluative associations that underlie implicit attitudes and have found that it is effective to measure subjects' unconscious bias regarding, *inter alia*, race, gender, and age.[247]

The tests used by Greenwald and Banaji are typical of the tests used by automatic stereotype researchers and utilize latency response methods to measure subjects' response time in pairing up series of attributes (associative attribute discrimination) with categories to which people often attach negative or positive associations, that is, stereotypes (initial target-concept discrimination).[248] The proliferation of computer use and the attendant ability to measure response times in milliseconds allows scientists to obtain data reflecting the extent of the cognitive barriers to making associations between the attributes and categories of persons.[249] Subjects are told to categorize the associations as quickly as possible, and results that indicate excessive errors or that are otherwise unreliable are not included in the test results. When an association is more easily made (such as the association between "flower" and "pleasant"), the response time is significantly faster than when the association is harder (such as the association between "insect" and "pleasant").[250] The researchers describe this as "superior performance for combinations that were evaluatively compatible than for noncompatible combinations."[251] Apparently, additional and more time-consuming cognitive processes are required to bridge the gap between "insect" and "pleasant." That is, it takes additional cognitive processes to overcome cognitive barriers to associate those two items. The unsurprising research results are that people generally have a more positive attitude toward flowers than insects, and they consistently make the connection between "flowers" and "pleasant" faster than "insects" and "pleasant."

---

246. Professors Anthony G. Greenwald and Mahzarin R. Banaji have set up a web site in which users can take the tests in the privacy of their homes to determine their unconscious bias in relation to gender, race, and age discrimination, among other categories. The web site also describes ongoing research and contains a comprehensive bibliography of research to date. *See* Anthony G. Greenwald & Mahzarin R. Banaji, *Implicit Association Test* (last modified Oct. 18, 1998) <http://depts.washington.edu/iat>.

247. *See* Greenwald et al., *supra* note 18, at 1464. The Implicit Association Test, or IAT, is similar in intent to cognitive priming procedures for measuring automatic affect or attitude. *See id.*

248. *See id.* at 1464–66.

249. *See id.* at 1467.

250. *See id.* at 1468–69.

251. *Id.* at 1466.

After determining that the IAT method can effectively measure implicit attitudes, the researchers applied the same methodology to racial categories. For example, in one experiment, subjects were tested for unconscious bias between target-concept categories of "black" names, such as Lamar, Malik, Ebony, Latisha and Tawanda, versus "white" names, such as Brandon, Ian, Nancy, Katie and Betsy.[252] The researchers used unpleasant and pleasant association words such as evil, war, love, and paradise. The data clearly revealed patterns consistent with the expectation that white subjects would reveal an implicit attitude difference between the black and white racial categories.[253] Specifically, whites generally showed an implicit attitudinal preference for white over black, which was manifested by their faster response times in combining white names and pleasant words as opposed to black names and pleasant words.[254]

The IAT researchers concurrently conducted explicit attitude tests to compare with the implicit measures. The IAT measure indicated considerably stronger relative preference for white than any of the explicit tests performed, indicating a divergence between the constructs assessed by the implicit and explicit measures.[255] An important purpose of the white-black experiment was to determine whether the IAT would reveal an implicit white preference among subjects who explicitly disavowed any black-white evaluative preference.

A review of the test results tended to show that the IAT may indeed implicitly reveal explicitly disavowed prejudice. While a majority of the white subjects (nineteen of twenty-six) explicitly endorsed a position of either black-white indifference or black preference, all but one had an IAT score indicating white preference.[256] These findings buttress the researchers' theory and expectation that people often have no conscious bias even though they are making implicit associations unconsciously, and the IAT can measure this implicit bias, also known as "in-group preference."[257] The tests reveal that unconscious forms of prejudice are

252. *See id.* at 1473–74.

253. *See id.*

254. In February 1999, when the author took the test over the Internet, 54% of Internet test takers had shown a "strong preference" for white over black, while only about 10% had "little or no" preference either way, and another 24% had shown varying degrees of bias against blacks. For current statistics, see Greenwald & Banaji, *supra* note 246.

255. *See* Greenwald et al., *supra* note 18, at 1477.

256. *See id.* at 1474–75.

257. *See id.* at 1476.

indeed pervasive even though as a society we do not explicitly endorse prejudice.[258]

Prior to the IAT, the most widely accepted method used for assessing automatic evaluative associations was evaluative semantic priming.[259] The similarity between evaluative semantic priming and the IAT supports IAT's acceptance as a trustworthy and accurate test method. In evaluative semantic priming, subjects classify series of target words based on the words' evaluative meaning, and a prime word, which was to be ignored, preceded each target word.[260] Prime-target evaluative congruencies facilitates responding to the target, thus producing variations in response latencies that can be used to measure automatic evaluation of the prime category. The more a category of words speeds judgments of positive evaluative targets or hinders judgments of negative evaluative targets, the more evaluative positivity is indicated for that category.[261] Evaluative priming studies have used prime stimulus categories very similar to the target-concept categories used in the IAT.

IAT measures share some important properties with the semantic priming method. Both procedures measure attitudes as the evaluative difference between two categories, such as the racial categories of black and white (known as "target concepts" in the IAT and "priming item categories" in semantic priming). For example, researchers using semantic priming methods contrast automatic evaluations evoked by "in-group" words such as "we" and "us," with "out-group" words such as "they" and "them," and with prime categories such as "young" and "old" or racial categories.[262] Furthermore, both procedures juxtapose items from categories for which an attribute is to be measured ("target concepts" in the IAT and "priming categories" in semantic priming), with items that have well-established attribute values, such as hostile,

258. One possible alternative to the implicit racism interpretation of the IAT score is that white subjects were more familiar with the white names than the black names, and this familiarity differential could explain the IAT results. However, because this explanation could not apply to the flower-insect experiment because the negative categories (insects and weapons) had substantially *higher* frequency in our language than the words used for positive categories (flowers and musical instruments), and therefore the researchers concluded that even if familiarity played a role, it could not fully explain the sets of finding for the studies. *See id.* at 1477.

259. *See id.* at 1477–78.

260. *See id.* at 1477.

261. *See id.*

262. *See id.*

pleasant, love, and war ("attribute categories" in the IAT and "target items" in semantic priming.)

Thus, the IAT builds upon well-established cognitive bias research methodology and, although relatively new, should be regarded as a fundamentally sound research tool for uncovering unconscious bias. Furthermore, in comparing the research results of the priming method and the IAT, researchers found that the IAT method has about twice the priming method's sensitivity to evaluative differences.[263]

Researchers are also creating other tests to measure unconscious bias. For example, scientists at Emory University have created a "lie detector" tool that they claim can uncover racial prejudice.[264] These researchers believe that measuring unconscious bias by observing affective indicators is more likely to be a consistent and strong predictor of racial and ethnic attitudes.[265] This belief is in accord with that of other researchers who posit that emotions predict some behavior better than cognitive-based measures of attitude because the affective system operates more crudely and processes information more rapidly than the rational system.[266] Thus, involuntary affective measures are more likely to reflect uncontrolled, automatic reactions to out-group members.[267]

This lie-detector-type process relies on electromyography (EMG), which detects tiny muscle movements in the face (affective indicators, such as cheek and eyebrow activity) that indicate bias just as an electrocardiogram (EKG) can detect heart murmurs.[268] Researchers found that subjects who passed rigorous oral and written tests indicating they did not hold prejudiced views held unconscious biases, as measured by invisible muscle reactions to photographs of men and women of different races.[269] Similar to the IAT test results, these tests produced discrepancies between subjects' self-reports on bias and facial EMG measures of unconscious bias toward members of other racial groups.[270]

---

263. *See id.* at 1477–78.

264. *See* Eric J. Vanman et al., *The Modern Face of Prejudice and Structural Features That Moderate the Effect of Cooperation on Affect*, 73 J. Personality & Soc. Psychol. 941, 944–45 (1997).

265. *See id.* at 944.

266. *See id.* at 943.

267. *See id.*

268. *See id.*

269. *See id.* at 947.

270. *See id.* at 941.

For purposes of this Article, it is not necessary to accept unconditionally that either the IAT or the EMG test is presently accurate enough to be a reliable indicator of unconscious bias. What is clear is that many tests have found discrepancies between people's conscious beliefs and unconscious biases against minority groups, that currently a proliferation of social psychology research exists concerning cognitive bias, and that it is only a matter of time before a generally accepted scientific testing method is available for common use. This Article thus proceeds on the assumption that unconscious-bias testing is now, or soon will be, available for use in testing employees for purposes of uncovering racial and other bias that, if left unexposed, could lead to unfair and illegal employment actions. [271]

This analysis proceeds on the theory that unconscious bias may be predictive of discriminatory behavior with the attendant harm to minorities, an assumed link that is supported by most social psychologists at this time, but remains somewhat controversial.[272] But even if the link between unconscious bias and overt behavior is unclear, what is clear is that most "unprejudiced" people are unaware of their own unconscious bias and upon recognizing it are likely to begin the process of converging their egalitarian views with their mental processes.[273] Further, the proliferation of unconscious-bias testing may accelerate current psychological understanding of human cognition and prejudice by generating data on the subject. Fundamentally, the more we understand about ourselves and other segments of our society, the better equipped we are to make strides toward creating a more just society for all of its members. Accordingly, this Article argues that unconscious-bias testing should be accorded an evidentiary privilege to encourage its use among employers who are interested in creating a truly discrimination-free environment.[274]

---

271. Importantly, Greenwald and Banaji specifically list such use of the IAT on their web page, and at least one company is presently being created to offer unconscious-bias testing to employers. *See* Telephone Interview with Anthony G. Greenwald, *supra* note 11; *see also* Greenwald & Banaji, *supra* note 246.

272. *See supra* notes 48–57 and accompanying text.

273. *See supra* notes 48–57 and accompanying text.

274. For a discussion about free speech and privacy implications of employers' use of such testing, see *infra* Part VI.D.

IV. EMPLOYERS ARE UNLIKELY TO CONDUCT UNCONSCIOUS-BIAS TESTING IN THE ABSENCE OF AN EVIDENTIARY PRIVILEGE

Presently, most employers are probably unaware of the fact that unconscious-bias testing is available or the benefits of its use. But unconscious bias and discrimination are areas in which social psychology research is proliferating quickly. As professional services for conducting unconscious-bias testing are organized and marketed, employer awareness will increase.

But employers may forgo unconscious-bias testing and sacrifice its benefits based on risk-management concerns because unconscious-bias test results have an obvious potential to be extremely damaging in discrimination lawsuits brought against employers. A factfinder's access to test results indicating a high level of racist or sexist implicit associations likely would precipitate a negative emotional response to the employer and a greater likelihood of finding liability and assessing punitive damages. No matter how committed an employer is to equal employment opportunity, the benefits of unconscious-bias testing may be outweighed by the risk of plaintiffs' accessing this potentially inflammatory and incriminating data to support an inference of discrimination.

This Article argues that unconscious bias is technically not relevant to Title VII disparate treatment analysis and mixed-motive analysis (because it does not prove intent) and is irrelevant to disparate impact analysis (which does not consider intent, but rather, relies on proof of a facially neutral employment test or policy that results in an adverse impact to a protected class). However, relevance is probably not a valid discovery objection because test results could be deemed reasonably calculated to lead to the discovery of other admissible evidence. Thus, concerns about misuse of the data in litigation still exist. Furthermore, judges may not understand the social psychology behind the relationship, or lack thereof, between unconscious bias and intentionally discriminatory action and may conflate the separate phenomena, finding unconscious bias probative and admissible on the issue of intent. Thus, without some clear protection from having unconscious-bias test results used against it, an employer has substantial reason not to conduct this testing because of the risk it could pose in litigation.

An evidentiary privilege protecting unconscious-bias testing from disclosure is necessary to encourage employers to conduct testing. As

explained more fully herein, making the privilege qualified will prevent employers from abusing the privilege.

V. CURRENT LAW FAILS TO RECOGNIZE A PRIVILEGE PROTECTING UNCONSCIOUS-BIAS TESTING

A. Introduction

At present, no privilege exists that would protect unconscious-bias testing from disclosure. However, some federal courts previously recognized a "self-critical analysis" privilege for corporate equal employment opportunity self-assessment. Although the self-critical-analysis privilege is no longer recognized in the employment context, it is useful to analyze the cases that supported the privilege to determine the policies that courts once thought justified such a privilege and which would similarly justify a privilege for unconscious-bias testing.

To understand the policies that have persuaded some federal courts to recognize the self-critical-analysis privilege in the employment context, it is necessary to preview the development of the privilege and its recognition in employment cases. The privilege originated to protect a hospital's peer-review notes in a medical malpractice case but was soon expanded into the employment context. It received mixed reactions from courts from the beginning, with some recognizing the privilege, while others refused to recognize it or simply failed to address it specifically. The case law demonstrates the way in which analysis of the privilege shifted over time until it was rendered essentially null in the employment context.

Although originally the privilege was grounded in traditional, utilitarian justifications, the courts quickly began to rely instead on factors derived from the factual circumstances of prior courts' holdings, even though the factual patterns that emerged in the case law were arbitrary and unrelated to the privilege's theoretical bases. In addition to the tendencies of many courts to look to the facts of prior cases rather than analyzing the privilege on policy grounds, the cases reflect fundamental differences of opinion about whether privilege law affects employers' behavior, whether recognition of the privilege contributes to Title VII's enforcement objectives, and whether employers are even interested in achieving a discrimination-free work environment. Many of the later opinions rely on Title VII's assumed efficacy as a mechanism for eradicating discrimination by focusing on the plaintiff's need for the information covered by the privilege to prove motivation and intent.

Thus, these later opinions many times find that allowing plaintiffs access to the documents advances Title VII's goals better than does encouraging employers to address their EEO deficiencies candidly by protecting self-criticism from discovery. These later opinions express more concern about supporting Title VII plaintiffs than allowing companies discretion to self-police in EEO matters, probably in hopes of pushing companies to eradicate discrimination by threat of litigation. Ironically, the courts' reliance on Title VII litigation as a means for eradicating discrimination is unjustified in light of Title VII's analytical framework and proof requirements, which do not comport with how discrimination really occurs in most cases.

All this created a confusing and analytically incoherent body of law relating to the self-critical-analysis privilege. Various federal courts produced a myriad of decisions representing different models of privilege analysis. Various "tests" emerged for when the privilege should apply. A multitude of factual distinctions also emerged to justify different analyses under different factual scenarios. Ultimately, factually indistinguishable cases produced directly contrary decisions based upon different basic assumptions about human nature and different levels of trust in litigation as a means to redress employment discrimination.[275] Some scholars attempted to delineate "criteria" for when the privilege was accepted:

> [F]irst, the information must result from a critical self analysis undertaken by the party seeking protection; second, the public must have a strong interest in preserving the free flow of the type of information sought; finally, the information must be of the type whose flow would be curtailed if discovery were allowed.[276]

One scholar added an additional criterion: the general proviso that no document will be accorded a privilege unless it was prepared with the expectation that it would be kept confidential.[277] However, no single set

---

275. *See, e.g.*, James A. Flanagan, *Rejecting a General Privilege for Self-Critical Analysis*, 51 Geo. Wash. L. Rev. 551, 554 n.18 (1993).

276. Note, *The Privilege of Self-Critical Analysis*, 96 Harv. L. Rev. 1083, 1086 (1983).

277. *See* Flanagan, *supra* note 275, at 574–76. For other scholarly comment on the self-critical-analysis privilege, see, e.g., Ronald J. Allen & Cynthia M. Hazelwood, *Preserving the Confidentiality of Internal Corporate Investigations*, 12 J. Corp. L. 355 (1987); Robert J. Bush, *Stimulating Corporate Self Regulation—The Corporate Self-Evaluative Privilege: Paradigmatic Preferentialism or Pragmatic Panacea*, 87 Nw. U. L. Rev. 597 (1993); Ellen Deason, *The Self-Critical Analysis Privilege and Discovery of Affirmative Action Plans in Title VII Suits*, 83 Mich. L. Rev. 405 (1984); David P. Leonard, *Codifying a Privilege for Self-Critical Analysis*, 25 Harv. J. on

of criteria or court-created test can explain the privilege's life and death. Some basic principles, however, can be derived by reviewing courts' analysis from the privilege's inception to its conclusion, and these principles are helpful in justifying a self-critical analysis kind of privilege for unconscious-bias testing.

B. Origins of the Self-Critical-Analysis Privilege: Social Benefits

The first case that recognized the self-critical-analysis privilege, *Bredice v. Doctors Hospital, Inc.*,[278] involved a hospital's peer review of doctors' procedures and judgment in a medical malpractice action. In 1970, Judge Corcoran of the U.S. District Court for the District of Columbia first recognized a "qualified privilege" for medical staff review minutes and reports.[279] The plaintiff, whose husband had died in the hospital's care, moved to compel production of the peer-review documents pursuant to former Federal Rule of Civil Procedure (FRCP) 34, which provided that courts could compel the production of documents upon a showing of good cause.[280] The judge refused to order production, finding no extraordinary circumstances preventing application of the qualified privilege.[281]

The court explained that the medical staff reviews were performed pursuant to the requirements of the Joint Commission on Accreditation of Hospitals, a prestigious organization created to effectuate standardized hospital practices nationwide, which confers accreditation only to hospitals that follow its recommendations.[282] The court found that the "sole objective" of the staff meetings was to improve the treatment of

---

Legis. 113 (1988); Stephen C. Simpson, *The Self-Critical Analysis Privilege in Employment Law*, 21 J. Corp. L. 577 (1996); Comment, *Civil Procedure: Self-Evaluative Reports—A Qualified Privilege in Discovery?*, 57 Minn. L. Rev. 807 (1973); Note, *Making Sense of Rules of Privilege Under the Structural (Il)Logic of the Federal Rules of Evidence*, 105 Harv. L. Rev. 1339 (1992); S. Kay McNab, Note, *Criticizing the Self-Criticism Privilege*, 1987 U. Ill. L. Rev. 675.

278. 50 F.R.D. 249 (D.D.C. 1970), *aff'd*, 479 F.2d 920 (D.C. Cir. 1973).

279. *See id.* at 251.

280. *See id.* at 249. Note that the federal rules relating to discovery have changed since *Bredice*, making discovery much more accessible. Because good cause is no longer required, the *Bredice* holding arguably is inapposite.

281. *See id.* at 251.

282. *See id.* at 250. The Commission is a nongovernment entity but serves a quasi-governmental role by granting accreditation only to hospitals that follow its recommendations. *See* Allen & Hazelwood, *supra* note 277, at 374 n.168.

hospital patients through critical self-analysis.[283] Furthermore, the review committee's work was done with the expectation of confidentiality, which is essential to the candid and conscientious evaluation necessary to continued improvement in patient treatment: "[P]rofessional criticism cannot occur in an atmosphere of apprehension that one doctor's suggestion will be used as a denunciation of a colleague's conduct in a malpractice suit."[284] The "overwhelming" public interest in continued advancement in medical care was the stated basis for the qualified privilege, which protects records of the reviews absent "evidence of extraordinary circumstances" needed to constitute "good cause" under FRCP 34 for this type of material.[285] Thus, the *Bredice* court based its denial of discovery on the social benefit of better health care through candid analysis of medical judgment and procedures (which can be achieved only if peer reviews were conducted without fear of disclosure), as well as the fairness in not forcing disclosure of documents created with an expectation of privacy.[286] Although the court does not mention Wigmore or the traditional justification for privileges, the bases for its holding clearly reflect the utilitarian rationale.[287]

After *Bredice*, federal courts began recognizing similar privileges to protect certain self-critical documents based on the policy concerns expressed in *Bredice*. The privilege was generally limited to three types of investigations: (1) confidential evaluations of hospital peer reviews; (2) internal corporate investigations (such as compliance with environmental laws); and (3) affirmative action/EEO reports.[288] Furthermore, a few courts recognized the privilege in other contexts, such as police department investigations of arrests and shootings.[289] This Article focuses on the third type of documents protected by federal case law as most analogous to an unconscious-bias testing privilege for EEO purposes. Because the 1975 adoption of Federal Rule of Evidence (FRE) 501 did not limit privilege analysis to then-existing law, but rather provided that federal courts must analyze privilege application "in light

283. *Bredice*, 50 F.R.D. at 250.

284. *Id.*

285. *Id.* at 251.

286. *See id.* at 250–51.

287. For a general discussion of traditional theories of privilege law analysis, see *infra* Part VI.A.

288. *See* Flanagan, *supra* note 275, at 552; *see also* Note, *supra* note 276, at 1090. "EEO" as used herein means equal employment opportunity.

289. *See* Leonard, *supra* note 277, at 118 & nn.19, 22 & 23.

of reason and experience," this Article will consider employment cases discussing the self-critical-analysis privilege both before and after 1975. This way it is possible to trace the privilege's analytical precedent and theoretical justifications for the purpose of formulating a workable and justifiable privilege rule for unconscious-bias testing that is consistent with both theoretical privilege law and the courts' prior recognition of this type of privilege.

C. *Self-Critical Analysis in Employment Cases*

Early employment cases followed the *Bredice* court's reasoning by analyzing how a self-critical-analysis privilege would serve the traditional, utilitarian justification of encouraging candor, based on the underlying assumption that recognizing a privilege would impact employers' conduct—that is to say, the cases used a social benefit theory. The idea was that recognition of a self-critical-analysis privilege would encourage frank self-criticism of EEO and AA[290] programs, resulting in a positive effect on equal employment opportunities because frank, uninhibited analysis was more likely to pinpoint areas in which antidiscrimination efforts were not working. The early cases distinguished between "facts" or "objective data" that are not the result of critical evaluation per se and subjective evaluations or portions of reports that contained subjective analysis of the company's EEO efficacy, ordering production of the former group only. During the early years, some courts operated under the belief that regardless of the fact that the reports were government mandated, employers had great latitude in how honestly they prepared the subjective, analytical portions of the reports. So to encourage honesty, with the expected insight that employers would gain from fully exploring their own potential wrongdoing and their attendant ability to identify and correct it, many courts felt that it was necessary to recognize the privilege.

In *Banks v. Lockheed-Georgia Co.*,[291] a Georgia district court in 1971 relied on *Bredice* to apply the self-critical-analysis privilege in the employment context for the first time.[292] In *Banks*, the plaintiffs sought all documents prepared by the company's EEO team, which was established in 1970 to study the company's equal employment

290. "AA" as used herein means affirmative action.

291. 53 F.R.D. 283 (N.D. Ga. 1971).

292. *See id.* at 285.

opportunities and to determine the efficacy of the company's affirmative action compliance program.[293] Formal written reports were created by the team pursuant to Executive Order 11,246,[294] and Lockheed did not object to producing these government-mandated reports.[295] However, Lockheed objected when the plaintiffs also sought to discover the team's "actual report," which included interim reports, other documents, and a candid self-analysis and evaluation of the company's EEO actions.[296]

After noting that the information requested was likely protected by the attorney work product doctrine, the *Banks* court opined that the "most critical issue" was whether the plaintiffs should have access to the candid self-analysis which was prepared "in an attempt to affirmatively strengthen the company's policy of compliance with Title VII and Executive Order 11,246."[297] Relying on the "analogous" *Bredice* case, the court held that it would be contrary to public policy to discourage frank self-criticism and evaluation of affirmative action programs of this kind.[298] Allowing the discovery would "discourage companies such as Lockheed from making investigations which are calculated to have a positive effect on equalizing employment opportunities."[299] The *Banks* court then constructed the dichotomy between "facts" and "objective data" on the one hand and subjective analysis on the other, and ordered Lockheed to produce only the factual or statistical information that was available to the team at the time it conducted the study.[300] The team's

293. *See id.* at 284.

294. Executive Order 11,246 requires nonexempt federal contractors to place nondiscrimination and affirmative action provisions in contracts with government agencies. *See* Exec. Order No. 11,246, § 202(1), 3 C.F.R. 339 (1964–1965), *reprinted as amended in* 42 U.S.C. § 2000e (1994); *see also* Exec. Order No. 11,375, 3 C.F.R. 684 (1967) (adding "sex discrimination" to text of Executive Order No. 11,246).

295. *See Banks*, 53 F.R.D. at 284.

296. *See id.*

297. *Id.* at 285. Note that since the court had apparently already decided that the documents were privileged under the work product doctrine, its self-critical-analysis decision is technically dicta. *See* Tharp v. Sivyer Steel Corp., 149 F.R.D. 177, 180 n.7 (S.D. Iowa 1993).

298. *Banks*, 53 F.R.D. at 285.

299. *Id.* at 255.

300. *See id.* An objective-subjective distinction for self-critical-analysis protection was recognized in other contexts at about the same time. *See, e.g.*, Gillman v. United States, 53 F.R.D. 316, 319 (S.D.N.Y. 1971) (distinguishing between factual testimony regarding suicide incident at mental hospital and "suggestions for future procedures" in Federal Tort Claims Act case).

analysis of its EEO program was protected by the "self-critical analysis" privilege.[301]

Several federal courts soon followed suit, declining to order production of similar EEO documents containing "evaluative" information. For example, in *Dickerson v. U.S. Steel Corp.*,[302] a Pennsylvania federal district court relied on the critical-self-evaluation privilege to deny the plaintiff's request for affirmative action and consent decree materials that contained the defendant's proposed goals, timetables, and evaluations of its progress.[303] The court based its decision on the public policy of encouraging employers to make candid internal evaluations to meet the objectives of the affirmative action program.[304] The *Dickerson* court noted that although affirmative action plans are required for all government contractors, the government cannot review in detail the massive amount of documents received under an affirmative action program.[305] The quality of these documents depends to a great extent on the good faith of employers in evaluating their programs and establishing affirmative action goals; if these documents are subject to discovery by plaintiffs, employers will not be candid and will set goals at minimum levels.[306] Thus, "a decrease in *voluntary* cooperation could seriously impair the equal employment opportunity policy."[307] The fact that the reports were government mandated did not change the court's view that because the government relied on companies' willingness to expose their mistakes and shortcomings voluntarily and candidly, the contents of the reports would be greatly affected by the privilege decision.[308]

---

301. *See Banks*, 53 F.R.D. at 285.

302. 14 Fair Empl. Prac. Cas. (BNA) 1448 (E.D. Pa. July 16, 1976).

303. *See id.* at 1449–50.

304. *See id.* at 1449.

305. *See id.*

306. *See id.*

307. *Id.* (emphasis added).

308. *See id.* For other cases upholding the privilege, see Rosario v. New York Times Co., 84 F.R.D. 626 (S.D.N.Y. 1979); Brown v. Ford Motor Co., 25 Fair Empl. Prac. Cas. (BNA) 708 (N.D. Ga. May 15, 1979) (holding internal EEO-related investigative reports that have not been disclosed by agency or were otherwise matters of public record not discoverable); Stevenson v. General Electric Co., 18 Fair Empl. Prac. Cas. (BNA) 746 (S.D. Ohio Oct. 4, 1978); Droughn v. FMC Corp., 74 F.R.D. 639 (E.D. Pa. 1977); Wehr v. Burroughs Corp., 20 Fair Empl. Prac. Cas. (BNA) 526 (E.D. Pa. June 17, 1977); Sanday v. Carnegie Mellon Univ., 12 Fair Empl. Prac. Cas. (BNA) 101 (W.D. Pa. Dec. 19, 1975) Rodgers v. U.S. Steel Corp., 12 Fair Empl. Prac. Cas. (BNA) 100 (W.D. Pa. Apr. 20, 1975).

However, during the same time period following *Banks*, other courts summarily dismissed the policy concerns undergirding the privilege—or failed to discuss them altogether. These courts apparently felt that the traditional privilege rationale of encouraging candor either did not in fact affect employers' behavior or was undermined by the fact that other communications (such as internal corporate memoranda) discussing EEO problems were discoverable; thus these courts eviscerated the intent of the privilege and circumvented its purpose. For example, in *Ligon v. Frito-Lay, Inc.*[309] the court ruled that the plaintiff in an employment discrimination action was entitled to discover the defendant's AA plans subject to the district court's initial *in camera* review for the purpose of eliminating any privileged matters.[310] The court declined to follow the *Banks* court's analysis of the need for encouraging candid evaluations because "disclosure in this lawsuit of affirmative action plans will no more discourage frank evaluations of companies than will the disclosure of admittedly discoverable inter-office communications between company officers."[311]

D. The Split of Authority and Courts' Reliance on Factual Distinctions

By 1978, enough employment cases had discussed the self-critical-analysis privilege for courts to begin addressing the "split" of authority and begin creating "tests" and "factors" to determine when the privilege applied, purportedly based primarily on the factual distinctions between cases applying the privilege and those declining to apply it. Thus, in *Webb v. Westinghouse Electric Corp.*,[312] a Pennsylvania federal district

309. 19 Fair Empl. Prac. Cas. (BNA) 722 (N.D. Tex. Jan. 23, 1978).

310. *See id.* at 723.

311. *Id.* (citing Palma v. Lake Waukomis Dev. Co., 48 F.R.D. 366 (D.C. Mo. 1970)). The *Ligon* court also noted the conflicting Freedom of Information Act (FOIA) cases in which the AA documents were sought, but found FOIA standards for production unrelated to discovery analysis under the Federal Rules of Civil Procedure. *See id.* at 722. For additional cases during this period in which documents potentially covered by the self-critical analysis were held discoverable, see *In re* Burlington Northern, 679 F.2d 762 (8th Cir. 1982) (noting that issue was whether to issue writ of mandamus, not whether to recognize privilege per se); Reynolds Metals Co. v. Rumsfield, 564 F.2d 663 (4th Cir. 1977) (holding that no self-critical-analysis privilege protects documents from disclosure to EEOC); Thompson v. Sun Oil Co., 523 F.2d 647 (8th Cir. 1975); EEOC v. Georgia Highway Express, Inc., 19 Fair Empl. Prac. Cas. (BNA) 101 (N.D. Ga. Oct. 31, 1978); Ylla v. Delta Air Lines, Inc., 25 Fair Empl. Prac. Cas. (BNA) 754 (N.D. Ga. Sept. 23, 1977); EEOC v. ISC, 16 Fair Empl. Prac. Cas. (BNA) 174 (W.D. Mo. June 28, 1977); and EEOC v. Quick Shop Markets, Inc., 396 F. Supp. 133 (E.D. Mo. 1975).

312. 81 F.R.D. 431 (E.D. Pa. 1978).

court first attempted to identify the parameters of the self-critical-analysis privilege in the employment context, recognizing that it had not been applied consistently.[313] Importantly, much of the fundamental policy analysis justifying the privilege was "lost" when courts looked to facts of prior cases rather than properly addressing privilege theory.

In *Webb,* a Title VII race discrimination class action case, the plaintiffs brought a motion to compel the disclosure of documents containing self-critical analysis of Westinghouse's employment policies and AA programs.[314] The court noted that the self-critical-analysis privilege was generally raised to shield from discovery reports required to be filed with the government and that Westinghouse's attempt to utilize the privilege in relation to the non-government-mandated documents was essentially a plea to expand the privilege beyond its past factual applications.[315] The court felt it was "required to examine the source and extent of the defense . . . and to define some guidelines for determining the circumstances which would justify shielding from discovery items constituting 'self-critical analysis'."[316]

The court first acknowledged that the theoretical basis for the self-critical-analysis defense to discovery is rooted in the public policy recognizing that employers' voluntary compliance with federal EEO laws is essential to meet EEO policy.[317] Employers must be encouraged to be candid and forthright, and disclosure of self-critical evaluations would tend to have a "chilling effect" on an employer's voluntary compliance with EEO laws.[318] On the other hand, the federal laws prohibiting employment discrimination manifest a strong policy of eradicating discrimination, and plaintiffs must be allowed to obtain the information to meet their burden of proof in discrimination lawsuits.[319] Thus, the public policy of encouraging employer candor directly conflicts with the public policy of eradicating discrimination through litigation.[320] As a result, the *Webb* court found it necessary to limit

313. *See id.* at 433.

314. *See id.*

315. *See id.* Note that in *Banks* the documents sought were not government mandated. *See also* Brown v. Ford Motor Co., 25 Fair Empl. Prac. Cas. (BNA) 708 (N.D. Ga. May 15, 1979).

316. *Webb,* 81 F.R.D. at 433.

317. *See id.*

318. *Id.*

319. *See id.*

320. *See id.*

carefully the situations in which encouraging self-critical analysis justified nondisclosure.[321]

After reviewing the case law that upheld the privilege for affirmative action plans submitted to the government pursuant to Executive Order 11,246 and for EEO-1 reports, the *Webb* court stated that the cases' "major justification" for excluding such materials was that because the reports are mandatory, public policy requires that employers be encouraged to be candid and complete in preparing the reports.[322] However, the court agreed with the *Dickerson* court's analysis that if the subjective and goal-setting portions of the mandatory reports were discoverable by the plaintiffs in Title VII suits, employers would not be candid and would set goals at minimum levels.[323]

The *Webb* court next observed that in all of the cited cases, the objective information (such as statistical data) was available to plaintiffs via other discovery channels and only subjective material was privileged because the public policy of encouraging candor does not apply to objective facts.[324] Thus, the court found that as long as the plaintiff is provided with the objective data, denial of production of the reports themselves does not really harm the plaintiff's case.[325] The *Webb* court also found that in cases where subjective materials prepared in the course of developing mandatory reports were shielded from discovery, the subjective information was either available through other means or was privileged on other grounds, which suggested that subjective, evaluative materials may not always be precluded from discovery.[326]

Based on the factual circumstances of earlier cases, the *Webb* court found three factors as "potential guideposts" for applying the self-critical-analysis privilege: (1) the materials protected have generally been prepared for mandatory government reports; (2) only subjective evaluative materials have been protected, while the objective data in the same reports have not been protected; and (3) courts have balanced plaintiffs' need for discovery and denied discovery only when the policy

321. *See id.* Note that the court assumes that Title VII is an effective means for redressing discrimination.

322. *Id.* at 434.

323. *See id.*

324. *See id.*

325. *See id.*

326. *See id.* (citing Banks v. Lockheed-Georgia Co., 53 F.R.D. 283 (N.D. Ga. 1971); Wehr v. Burroughs, 20 Fair Empl. Prac. Cas. (BNA) 526 (E.D. Pa. May 6, 1977)).

favoring exclusion clearly outweighed plaintiffs' need for the information.[327] Although the *Webb* court stated that "[i]t is not possible to draw a bright line between those situations where disclosure of 'self-critical analysis' should be compelled, and those where it should be shielded," the court held that the discovery requested was either not related to mandatory government reporting or was "objective" information and therefore "*cannot* fall within the 'critical self analysis' exemption as *defined* in this memorandum."[328]

E. Policy-Based Critique of Factors

With the pronouncement of its first guidepost based on whether the self-critical analysis is part of a mandatory government reporting, the *Webb* court wasted no time in unmooring the application of privileges from the theoretical justification for privileges. The privilege is justified to the extent that candid self-critical analysis will illuminate and enable the dismantling of discriminatory employment practices. The mere fact

---

327. *See id.* Note that the third guidepost clarifies the qualified nature of the privilege and that it can be overcome by a particular showing of necessity in any given case. One area in which this balance generally favored plaintiffs was with regard to tenure votes/peer review when discrimination in the tenure decision was alleged. *See, e.g.*, Gray v. Board of Higher Educ., 692 F.2d 901, 904–09 (2d Cir. 1983) (§§ 1981, 1983, and 1985 action); *In re* Dinnan, 661 F.2d 426, 432 (5th Cir. 1981); Lynn v. Regents of the Univ. of Cal., 656 F.2d 1337, 1347 (9th Cir. 1981) (Title VII sex discrimination case); Jespen v. Florida Bd. of Regents, 610 F.2d 1379, 1382–83 (5th Cir. 1980); Keyes v. Lenoir Rhyne College, 552 F.2d 579, 581 (4th Cir. 1977); *see also* Flanagan, *supra* note 275, at 557 n.44. Some academic institutions voluntarily adopted open proceedings for tenure decisions, based on the belief that this encourages review committees to evaluate candidates on objective criteria that can justify the decision. *See Developments in the Law—Privileged Communications*, 98 Harv. L. Rev. 1450, 1627–28 & n.203 (1985). Some states passed laws specifically excluding peer reviews from protection in discrimination suits. *See id.* at 1627 n.202; *see also* Cal. Evid. Code § 1158 (West 1985); Wash. Rev. Code § 4.24.250 (1985). Of course, in 1990 the U.S. Supreme Court held there is no peer-review privilege in Title VII suits alleging discrimination in a tenure decision. *See* University of Pa. v. EEOC, 493 U.S. 182, 189–92 (1990).

328. *Webb*, 81 F.R.D. at 434–35 (emphasis added). Specifically, the court held that the information requested in the case, such as the names of employees who had made race discrimination complaints to Westinghouse's management, and information relating to what EEO training courses were attended by Westinghouse's employees, was "clearly objective data." *Id.* The privilege did not apply to objective information about meetings (such as dates, participants, and subject matter) in which AA, race discrimination, or the litigation were discussed; because "these discussions do not relate to compliance with mandatory governmental reports, we believe that the defense of 'self-critical analysis' cannot be asserted." *Id.* at 435. This information may be necessary for plaintiffs to prove intent to discriminate. *See id.* The court upheld its prior order requiring production of "all studies, whether statistical or otherwise, demonstrating or tending to demonstrate racial discrimination in any phase of employment at the [defendant's facility]" because they consisted of "statistical and . . . objective" information. *Id.*

that the self-critical analysis is voluntary and not government mandated undercuts its promise to eradicate discriminatory practices. Indeed, the quality of the voluntary assessments, just as with the government-mandated reports, is increased when employers are free to conduct candid assessment. One may expect voluntary self-critical analysis to be even more responsive to the candid atmosphere created by a privilege because the voluntary nature of the inquiry evidences a greater commitment to EEO. If candor leads to insight and better EEO policy, then self-assessment unrestricted as to content or form and done by free choice is likely to analyze more vigorously EEO policy, resulting in more exacting assessment. Moreover, voluntary self-analysis represents the very basis of the privilege.[329] As a practical matter, the primary value that the government-mandated versus voluntary distinction serves is to ensure that the investigations are conducted in the public's best interest.[330] From a policy perspective, as long as voluntary self-assessment serves the public interest, the privilege should apply.

The fact that the privilege had generally not been applied to non-government-mandated reports[331] before *Webb* does not justify *Webb*'s first guidepost. Ironically, *Banks*, the first case to invoke the self-critical-analysis privilege in the employment context, applied the privilege to the team's reports, not the government reports per se.[332]

---

329. The peer review notes in *Bredice* were not government mandated, and neither were the team's reports in *Banks*. *See supra* note, 278–89, 298, and accompanying text.

330. *See* Allen & Hazelwood, *supra* note 277, at 378.

331. *See Webb*, 81 F.R.D. at 434. Note, however, that *Banks* and *Brown v. Ford Motor Co.*, 25 Fair Empl. Prac. Cas. (BNA) 708 (N.D. Ga. Sept. 29, 1978) both concerned documents that were not government mandated. In *Brown*, the court held that results of internal EEO investigations that had not been disclosed by a government agency or were otherwise a matter of public record were not discoverable because disclosure "may tend to discourage voluntary compliance with the self-investigation procedures carried out under the antidiscrimination laws." *Id.* at 710. Later cases held that the fact that government reports were mandated meant that companies had no choice about whether to conduct self-assessments, and therefore the privilege was unnecessary to encourage companies to cooperate. Yet, by this time, the privilege's application only to government reports had become so entrenched in the case law that companies had stopped arguing for a self-critical-analysis privilege for non-government-mandated reports or voluntary self-assessment. *See infra* Part V.E.

332. *See* Banks v. Lockheed-Georgia Co., 53 F.R.D. 283, 284 (1971). In that case, Lockheed had agreed to turn over the reports. Although it could be argued that the Lockheed team's reports were done as part of an effort to comply with the government reporting requirements, this does not change *Webb*'s flawed analysis distinguishing mandatory from voluntary reports. One court found that when a company shows that its analysis "was performed *in preparation* for a report required by the government[,] . . . [t]he 'self evaluation' privilege should apply to it" and declined to embrace the plaintiff's argument that the privilege applied only to government-mandated reports because "[s]uch an approach may be overly narrow and thus controvert the purpose behind the privilege to encourage

Webb's third guidepost—balancing plaintiffs' need for discovery against policies favoring exclusion—is similarly troublesome.[333] The cornerstone of the traditional, utilitarian privilege justification is that recognition of the privilege creates a "safe space" in which people can communicate freely without fear of the communications later being used against them. To create this safe space and realize the policies of the traditional justification, the privilege must appear absolute and certain—otherwise, the communication may be chilled by the prospect that a court might decide not to recognize the privilege in a specific case.[334] By using a formal "balancing" approach,[335] courts undermine trust in the privilege, thereby chilling the very communication sought to be encouraged.[336] In fact, some courts used the greater uncertainty in the privilege's application caused by the balancing test as a basis for holding that no expectation of confidentiality is possible because the split of authority provides "notice" that confidentiality is not certain.[337]

The *Webb* guideposts gained widespread adherence, despite *Webb*'s confused privilege analysis (mixing utilitarian-based policy with privacy-based application) and reliance on prior cases' facts. *Webb*'s "guideposts" were easy to apply. Thus, in spite of *Webb*'s lack of solid privilege law analysis, which made its guideposts theoretically untenable, philosophically inconsistent, and destined for failure, most later courts relied heavily on *Webb*'s "potential guideposts" in deciding whether to apply the self-critical-analysis privilege and sometimes even referred to them as "necessary" standards.[338] *Webb*'s weak privilege analysis

---

voluntary compliance with the equal employment opportunity laws." Hoffman v. United Telecomm. Inc., 117 F.R.D. 440, 443 (D. Kan. 1987) (emphasis added). The issue in this case arose in the context of deposition questions regarding the evaluator's analysis, not the documents. *See id.* at 442.

333. For criticism of the second guidepost, which this author does not find particularly troublesome, see Bush, *supra* note 277, at 609–10.

334. *See infra* Part VI.A.; *see also* Note, *supra* note 276, at 1097 (arguing for absolute privilege and stating that "to apply the privilege of self-critical analysis in such a [case-by-case] manner is to risk its evisceration").

335. Balancing is the preferred method for analyzing privileges grounded in *privacy*, not the traditional justification. *See infra* Part VI.A.

336. *See infra* Part VI.A.

337. Although FRE 501 directs courts to *create or develop* privileges on a case-by-case basis, a privilege should not be *applied* on a case-by-case basis once a privilege is established on utilitarian grounds. *See* Bush, *supra* note 277, at 608; Note, *supra* note 276, at 1097.

338. *See* Witten v. A.H. Smith & Co., 100 F.R.D. 446, 451 (D. Md. 1984), *aff'd*, 785 F.2d 306 (4th Cir. 1986) (referring to *Webb*'s factors as "necessary" standards); Roberts v. National Detroit Corp., 87 F.R.D. 30, 32 (E.D. Mich. 1980) (noting *Webb*'s "approach appears to strike an appropriate balance and is in keeping with the public policy arguments"). *But see, e.g.*, Woods v.

became entrenched in self-critical-analysis case law and no doubt contributed greatly to the privilege's ultimate evisceration.

F. O'Connor *Factors and Analysis*

Two years after *Webb*, a Massachusetts federal district court decision attempted to clarify further the privilege and, in conjunction with *Webb*, became the nearly uniform "standard" for determining the applicability of the self-critical-analysis privilege. In *O'Connor v. Chrysler Corp.*,[339] the issue was whether self-evaluation portions of Chrysler's affirmative action plans were discoverable.[340] Chrysler raised the privilege of self-critical analysis.[341]

The *O'Connor* court first acknowledged that authorities were divided on whether the self-evaluative portions of AA plans are discoverable.[342] According to the *O'Connor* court, the cases protecting evaluative portions identified three factors weighing against disclosure: (1) the materials are protected attorney work product prepared in anticipation of litigation, (2) disclosure would discourage employees from frankly evaluating their companies, and (3) disclosure would discourage employers from candidly reporting the results.[343]

The *O'Connor* court found the underlying question to be whether the privilege should enjoy continued recognition and asked whether confidentiality in critical self-evaluations or disclosure of material potentially probative of discriminatory intent would better serve equal employment opportunity policy.[344] The court concluded that subjecting to discovery affirmative action plan evaluative conclusions would not necessarily deter or substantially detract from the thoroughness of future self-evaluations.[345] This conclusion was supported by two considerations.

Coca-Cola Co., 35 Fair Empl. Prac. Cas. (BNA) 151, 154–55 (N.D. Ga. June 10, 1982) (applying privilege to protect self-analysis and evaluative portions of AA plans with no reference to *Webb* factors).

339. 86 F.R.D. 211 (D. Mass. 1980).

340. *See id.* at 213–14. Chrysler's AA plan was prepared for the federal government in accordance with Executive Order 11,246 as amended by Executive Order 11,375. *See id.* at 214.

341. *See id.* at 213–14.

342. *See id.* at 216.

343. *See id.* It is not clear how the second and third factors differ analytically. Both appear to be based on the traditional justification.

344. *See id.* at 217.

345. *See id.*

First, because the employers must prepare the reports regardless of the privilege,[346] the reports will occur even if the threat of discovery is present.[347] Second, employers may have significant deterrents to candid self-evaluations regardless of the possibility of discovery.[348] For example, the employee charged with writing the report may be more concerned about protecting himself or co-workers from discipline than about exposing the employer to discrimination liability.[349] Thus, the "additional deterrence of investigation occasioned by the possibility of discovery may be minute."[350]

Nonetheless, the *O'Connor* court recognized that a lack of confidentiality will cramp the investigative process and ruled that neither an unqualified requirement of disclosure nor an unqualified privilege of nondisclosure is justified. The court characterized the situation as the "clash between highly valued interests" of discovering all facts probative in a civil claim, of being fair to persons legally required to engage in self-evaluation, and of encouraging candid government-mandated self-evaluations.[351] Thus, the *O'Connor* court determined that: (1) facts and data are discoverable; (2) sheer evaluation of facts formulated in response to the government's requirement of critical self-evaluation is not discoverable; and (3) where the report combines facts and self-critical evaluations, the defendant must prepare and disclose a substitute statement that "includes all of the express or implied recitations of fact, while omitting part or all of the self-evaluative statements."[352] The court

346. This analysis is directly contrary to the analysis in *Dickerson*, which reasoned that, although mandatory, how candid the employer is affects the contents of the mandatory reports, and a decrease in the degree of employer candor in these reports could "seriously impair" EEO policy. *See* Dickerson v. U.S. Steel Corp., 14 Fair Empl. Prac. Cas. (BNA) 1448, 1449 (E.D. Pa. July 16, 1976).

347. *See O'Connor*, 86 F.R.D. at 217.

348. *See id.*

349. *See id.*

350. *Id.* Note the similarity of this "additional deterrence" argument to the Independent Counsel's argument that there is "no harm in one more exception" to the attorney-client privilege in *Swidler & Berlin v. United States*, 524 U.S. 399, 409–10 (1998). In *Swidler & Berlin*, the Supreme Court held that this is an improper basis for privilege law analysis under FRE 501, but the *Swidler & Berlin* decision did not come down until 18 years after the *O'Connor* decision. *See id.*; *infra* notes 495–500 and accompanying text.

351. *O'Connor*, 86 F.R.D. at 218. This effectiveness argument is contrary to the court's other statements that confidentiality will not create more candor (and therefore more effective self-analysis) because of other "deterrents."

352. *Id.*

detailed redaction procedures, including an *in camera* inspection of selected samples of the withheld portions.[353]

O'Connor's "highly valued interest" in fairness to persons legally required to engage in self-evaluation assumes greater importance in later cases as a basis for recognizing the privilege. These later courts consider it unfair to force an employer to create a report that could be turned against it by a private litigant, a "fairness" argument akin to the Fifth Amendment privilege against self-incrimination.[354] At the same time, these courts are not concerned about the effect of the privilege on the content of the self-evaluations, stating that because they are mandatory they simply must be done with or without the threat of discovery. Furthermore, to the extent that the contents are affected by the fear of disclosure, these courts found other reasons why employers may be inclined to "fudge" their self-evaluations, such as a report drafter's concern about his own or his fellow employees' potential discipline.[355]

G. *Critique of* O'Connor*'s Empirical Assumptions, Factual Distinctions, and Fairness Analysis*

The *O'Connor* court's analysis does not square with the traditional justification for privileges. First, its assumption that government-mandated reports will be completed in the same manner regardless of the existence of a privilege departs from the reasoning in *Dickerson* and privilege theory generally. In *Dickerson*, the court recognized that the quality of the government-mandated reports depended on the employer's "voluntary" compliance and execution of the report in good faith and with candor.[356] The premise of the traditional justification is that communication is hampered when people are afraid that their communications will be disclosed and that recognition of a privilege will affect privilege holders' communications. The force of law alone is not a reliable means of extracting truth; as other courts have noted, a person

353. *See id.* Judge Keeton recognized the difficulty of conducting an *in camera* inspection in a nonjury case, but apparently felt an *in camera* inspection of samples was the fairest way to protect both parties' interests. *See id.*

354. *See id.* The regulations require, *inter alia*, a detailed analysis of deficiencies in utilizing minority groups and timetables for correction of the deficiencies. *See* 41 C.F.R. §§ 60-1.40(a), 60-2.10, 60-2.12 (1998).

355. *See, e.g.*, *supra* notes 349–52 and accompanying text.

356. Dickerson v. U.S. Steel Corp., 14 Fair Empl. Prac. Cas. (BNA) 1448, 1449 (E.D. Pa. July 16, 1976).

forced to testify will likely give false testimony.[357] Thus, the *O'Connor* court's assumption that the mandatory reports will be done regardless of the privilege shows a lack of trust in the traditional rationale, as well as a disregard for prior courts' determination of the need to encourage accurate and candid government reporting.

Second, the traditional justification invalidates the *O'Connor* court's argument that significant deterrents to candid reporting would exist even after recognition of a privilege. The mere possibility that clients lie to their attorneys because of pride or fear or embarrassment hardly delegitimizes the attorney-client privilege. Traditionally, privileges are justified only if the benefit gained by protecting the information outweighs the impediment to the search for truth. If the report drafters are likely lying, then recognition of the privilege creates little or no loss of credible evidence because virtually any potential benefit gained would outweigh the impediment to truth.

The U.S. Supreme Court's holding in *Swidler & Berlin v. United States*[358] has proven that the existence of other deterrents is not a proper basis for assuming the privilege would not work. In *Swidler & Berlin*, the Independent Counsel argued that one more exception to the attorney-client privilege would have only a "minimal" impact because other exceptions already existed.[359] The Court responded that such uncertainty in a privilege's application could create client fear resulting in inhibited disclosure and that fear of uncertainty is precisely the reason the Court had previously rejected a balancing approach to the attorney-client privilege.[360] The Court explained that an assumption that people may already question a privilege's trustworthiness, based on existing exceptions, is not an appropriate basis for making a privilege ruling and could contribute to the general erosion of the privilege.[361]

Finally, the "unfairness" argument ignores the unfair predicament of companies that engage in honest and voluntary self-assessment that would be embarrassing and damaging if publicly exposed. The

357. Indeed, under the "Image Theory" of evidentiary privileges, the legal system bestows privileges to the groups most likely to lie or refuse to testify to preserve the image and legitimacy of the legal system. *See Developments in the Law—Privileged Communications*, *supra* note 327, at 1498–99; *see also In re* Agosto, 553 F. Supp. 1298, 1310, 1327 (D. Nev. 1983).

358. 524 U.S. 399 (1998).

359. *Id.* at 409–10.

360. *See id.* at 409 (citing Upjohn Co. v. United States, 449 U.S. 383, 393 (1981)).

361. *See id.* at 410.

O'Connor court deemed it unfair for the government to force companies to create reports that can be used against them by plaintiffs. However, it is not less unfair to force companies who engage in voluntary critical self-assessment to disclose the assessments simply because they chose to create the potentially adverse evidence. Companies who exceed the legal requirements to uncover discrimination probably do so to avoid litigation or to create a fair work environment for philosophical reasons. In light of Title VII's purposes, these voluntary efforts should be encouraged by protecting the self-critical reports from disclosure. These companies are perhaps more committed to meeting Title VII's purposes and more likely to take additional voluntary action based on their self-assessments. It would thus be more unfair to subject these companies to disclosure of private assessments because they are taking greater efforts to effectuate the purposes of Title VII. This voluntary effort to eradicate discrimination is precisely the type of corporate behavior a self-critical-analysis privilege originally sought to foster.[362]

The *O'Connor* opinion thus entrenches the *Webb* court's dichotomy between government-mandated reports and voluntary self-assessments, while apparently failing to understand, or possibly disagreeing with, the basic reasoning supporting the traditional justification for privileges. *O'Connor* nonetheless became one of the most quoted opinions in self-critical-analysis jurisprudence and carried a precedential force unjustified by its reasoning.

H. Post-O'Connor

Most courts after *O'Connor* continued to follow the "objective" facts versus "subjective" analysis distinction—the distinction debated in *Banks* and made a "factor" for application of the self-critical-analysis privilege under *Webb* and *O'Connor*—by ordering production of statistics and other facts contained in affirmative action reports.[363] However, courts tended to adopt the various arguments put forth in *O'Connor* and apply the three-prong test with no independent analysis,

362. *See* Bredice v. Doctors Hospital, Inc., 50 F.R.D. 249, 251 (D.D.C. 1970), *aff'd*, 479 F.2d 920 (D.C. Cir. 1973); *see also* Banks v. Lockheed-Georgia Co., 53 F.R.D. 283 (N.D. Ga. 1971).

363. *See, e.g.*, Woods v. Coca-Cola Co., 35 Fair Empl. Prac. Cas. (BNA) 151, 154–55 (N.D. Ga. June 10, 1982); Penk v. Oregon State Bd. of Higher Educ., 99 F.R.D. 506, 507 (D. Or. 1982); Jamison v. Storer Broad., 511 F. Supp. 1286, 1296–97 (E.D. Mich. 1981), *aff'd in part and rev'd in part on other grounds*, 830 F.2d 194 (6th Cir. 1987).

so that the flawed reasoning became masked by the sheer number of court decisions upholding the precedential "standards."[364]

Resnick v. American Dental Ass'n[365] is an excellent example of a court applying the *O'Connor* factors with virtually no policy analysis. The American Dental Association (ADA) attempted to use the self-critical-analysis privilege in an employment discrimination suit to withhold (1) a "personnel practices study" conducted by a private management consulting firm and (2) documents from employee relations committee meetings.[366] The ADA made no claim that its consultants' work involved government-required reports—unlike the defendants in *Banks*, *O'Connor*, and *Webb*.[367] The court declared that it "need not decide whether to embrace" the privilege, stating that to the extent it has been applied, the standards set forth in *O'Connor* had been met.[368] Because the "ADA *alone chose* to undertake the [self-critical assessment] activities . . . . [n]either that fairness rationale nor that effective enforcement rationale [set forth in *O'Connor*] operates here."[369] By

364. *See, e.g.*, Spencer Sav. Bank v. Excell Mortgage Corp., 960 F. Supp. 835, 843 (D.N.J. 1997) ("Where self-evaluation has been voluntarily '[u]ndertaken, neither that fairness rationale nor [an] effective enforcement rationale operates'. Accordingly, the justifications in support of applying the privilege to government-mandated reports bear no relevance to [the] ruling today.") (citations omitted); Tharp v. Sivyer Steel Corp., 149 F.R.D. 177, 182 (S.D. Iowa 1993) (discussing *O'Connor* factors and stating that "[t]hose courts which have adopted application of the privilege, however, have placed *four limitations* on its application") (emphasis added and citations omitted); Steinle v. Boeing Co., 62 Fair Empl. Prac. Cas. (BNA) 272, 278–79 (D. Kan. Feb. 3, 1992) (holding that privilege is limited to government-mandated reports); Vanek v. Nutrasweet Co., 59 Empl. Prac. Dec. (CCH) ¶ 41,600, 71,455–56 (N.D. Ill. 1992) ("Since any self-critical analysis that was performed was a voluntary decision on the defendant's part, other authorities support an argument that the information requested by plaintiff is not privileged.") (citations omitted); Hardy v. New York News, Inc., 114 F.R.D. 633, 641 (S.D.N.Y. 1987) ("Virtually every court has limited the privilege to information or reports that are mandated by statute.") (citations omitted); Witten v. A.H. Smith & Co., 100 F.R.D. 446, 450–51 (D. Md. 1984); Roberts v. Carrier Corp., 107 F.R.D. 678, 684 (N.D. Ind. 1985) ("When [the privilege] has been adopted, the courts have *consistently* applied [the *O'Connor*] standards.") (emphasis added and citations omitted); *see also* Dowling v. American Hawaii Cruises, Inc., 971 F.2d 423, 427 (9th Cir. 1992) (holding no privilege for voluntary routine safety reviews on vessel in Jones Act lawsuit). Some courts, however, took note of *O'Connor* but did not specifically apply the factors. *See, e.g.*, Zahorik v. Cornell Univ., 98 F.R.D. 27 (S.D.N.Y. 1983).

365. 95 F.R.D. 372 (1982).

366. *Id.* at 374.

367. *See id.*

368. *See id.*

369. *Id.* at 374–75 (emphasis added). Again, the court failed to recognize or discuss the latitude employers have in making the required reports, or the idea that, whether required or voluntary, more accurate, candid reports would be made if a privilege protected critical analysis from disclosure.

relying on *O'Connor*'s "factors," the *Resnick* court's finding that voluntary undertaking of self-critical analysis does not implicate the policies supporting the privilege entirely misses the original basis for the self-critical-analysis privilege—to encourage companies to evaluate thoroughly and candidly their antidiscrimination policies to effectuate greater compliance with antidiscrimination legislation.

The *Resnick* court specifically stated that no enforcement scheme like that under Executive Order 11,246 is "implicated" to support its *O'Connor*-based argument that no confidentiality is needed to create an effective incentive structure for voluntary self-evaluations, because they are not a part of the enforcement scheme set forth in Title VII and its regulations.[370] Yet, the U.S. Supreme Court has condoned voluntary action to effectuate Title VII's purpose and has stated that Congress' intent in promulgating Title VII was to eradicate all vestiges of discrimination, which includes voluntary actions by employers to meet this fundamental goal.[371]

The *Resnick* court's focus on the "implications" of the enforcement scheme set forth in Executive Order 11,246 as a justification for the voluntary-mandatory distinction makes no theoretical sense. The fact that voluntary evaluations are not required by law means greater incentives are needed to encourage employers to undertake self-assessment; indeed, *O'Connor* essentially argued that because government reports are required, the need for confidentiality to encourage their production is lessened.[372] The *Resnick* court's reasoning subverts Title VII's overall purposes to one specific regulation promulgated to help effectuate those purposes and gives employers a disincentive to work voluntarily towards EEO in any way that produces documents. *Resnick* exemplifies the way in which sensible privilege law analysis has been lost through reliance on the "standards" set forth in *Webb* and *O'Connor*, resulting in irreconcilable policy arguments that in some instances clash with the

---

Note also that the court did not accept the attorney-client privilege as a basis for nondisclosure since the attorney's involvement was "tangential." *See id.*

370. *See id.*

371. *See, e.g.*, United Steelworkers of Am. v. Weber, 443 U.S. 193, 204 (1979) (upholding voluntary AA plan as consistent with Title VII's purpose and stating that in enacting Title VII, Congress's intent was to spur "employers and unions to self-examine and self-evaluate their employment practices and to endeavor to eliminate, so far as possible, the last vestiges of an unfortunate and ignominious page in this country's history") (quoting Albemarle Paper Co. v. Moody, 422 U.S. 405, 417–18 (1975)).

372. Chrysler Corp. v. O'Connor, 86 F.R.D. 211, 217 (D. Mass. 1980).

basic concern that prompted Executive Order 11,246—to encourage self-evaluation as a means to eradicate discriminatory practices.

I. Focus on Confidentiality

Two years after *Resnick*, a Maryland federal district court in *Witten v. A.H. Smith & Co.*[373] reiterated the "necessary" standards set forth in *Webb* and *O'Connor*, but then declined to recognize the privilege by modifying the pertinent inquiries and virtually eviscerating the privilege.[374] The issue was whether to compel production of the defendant's AA plan materials in a race discrimination suit brought pursuant to § 1981 and Title VII.[375] The defendants argued that the reports were irrelevant or constituted privileged information based on the self-critical-analysis privilege.[376] After acknowledging the split of authority and the *O'Connor* standards for applying the privilege, the *Witten* court held that FRE 501 gives federal courts the flexibility to develop rules of privilege on a case-by-case basis.[377]

The *Witten* court then set forth Wigmore's "four fundamental conditions necessary to establish a privilege against disclosure of communications" and analyzed the case under these factors.[378] First, the court said that there could be no expectation of confidentiality.[379] The federal procurement regulations state that the information in the reports will be used to effectuate the purposes of Title VII, so consequently little merit exists in the argument that such reports should remain confidential and precluded from use in Title VII suits.[380] Moreover, the split in authority regarding the privilege militates against an expectation of confidentiality; the fact that many courts have ordered production despite a claim of a self-critical-analysis privilege makes an expectation insupportable.[381]

---

373. 100 F.R.D. 446 (D. Md. 1984).

374. *See id.* at 450–52.

375. *See id.* at 445.

376. *See id.* at 449–50.

377. *See id.* at 451 (citing 120 Cong. Rec. 40,891 (1974) (statement of Rep. Hungate)).

378. *Id.* at 452; see *infra* text accompanying note 432 for Wigmore's four factors.

379. *See Witten*, 100 F.R.D. at 452.

380. *See id.*

381. *See id.*

Second, the *Witten* court deemed protection of the reports "not necessary to assure the continued status of the relationship presently existing between the federal procurement agencies and the parties contracting with the government."[382] The court stated that "it has not been *demonstrated* that disclosure would significantly discourage the amount of voluntary compliance" and referred to the *O'Connor* reasoning that deterrents to candid self-evaluation exist that are unrelated to the confidentiality issue.[383] Other deterrents to candor and high goal setting present in every case include compliance reviews (which encourage low AA plan goal setting because failure to meet the goals can cost the company the contract or its status as a prime contractor) and the use of plans in litigation (a company's compliance with its AA plan is "regularly considered by courts as evidence of a defendant corporation's attention to a problem or as an affirmative defense to a particular claim of discrimination").[384] Thus, the *Witten* court concluded that not assuring confidentiality provides no more incentive to lessen candor and goal setting than is already present by these factors.[385]

The third factor—that the relationship must be one that the community thinks worthy of sedulously fostering—was "not a matter of dispute," and the fourth factor was a judgment call: a "cost-benefit" analysis of harm to the "confidential" relationship versus benefit to society if disclosure is required.[386] The *Witten* court found that AA plans could contain useful information of intent and motivation and decided that considering the difficulty of proving intent in discrimination cases, the plaintiff's need significantly outweighed any possible decrease in quality.[387]

The *Witten* court placed undue weight on Wigmore's factors and failed to conduct a full privilege analysis as called for by FRE 501. Although the court recognized its discretion and flexibility under FRE 501, it applied Wigmore's factors in lieu of the original privilege analysis based on reason and experience envisioned by FRE 501. Wigmore's

382. *Id.* at 453.

383. *Id.* at 452 (emphasis added). The court appears to want empirical proof of the privilege's *effect* on communications but does not mention that there is no proof that the privilege does *not* encourage more candid communication, relying instead on theories about "other deterrents" set forth by *O'Connor* and courts that relied on *O'Connor*.

384. *Id.* at 453.

385. *See id.* at 452–54.

386. *Id.* at 453–54.

387. *See id.* at 454.

factors were not developed with the corporate context in mind and should be refashioned as balancing social benefits of confidentiality against its social costs.[388] By using Wigmore's factors instead of a full utilitarian analysis or even *Webb*'s and *O'Connor*'s factors, the *Witten* court eviscerated the privilege because the first Wigmore factor can never be met with regard to government reports, particularly considering the split of authority in the privilege's application. Furthermore, the U.S. Supreme Court has held that under FRE 501, courts are not bound by any prior rules or standards for applying privileges.[389] That is, a court's duty under FRE 501 is to conduct original analysis, not simply to apply old privilege standards.[390] Just as an entrenched expectation of confidentiality is not controlling, neither should a lack of an expectation be determinative or given any more weight than is appropriate in light of reason and experience. It is as if the *Witten* court had decided to begin the demise of the self-critical-analysis privilege and "found" firm grounds against the privilege's application in Wigmore's first factor, therefore giving it great weight.

Witten's other stated basis for no expectation of confidentiality—that the regulation put companies on notice that reports are not confidential—is also not persuasive. The court assumed that the regulation's language, stating that the reports "'shall be used only in connection with the administration of the Order [Executive Order 11,246], the administration of the Civil Rights Act of 1964, and in furtherance of the Order and that Act,'" equates to a legislative mandate that the reports be turned over to private litigants.[391] Instead, this language could mean that the documents are limited to administrative use because no specific mention is made of non-administrative use and because "only" qualifies "used," indicating a limitation on the documents' use. But even assuming the regulatory language supports a litigant's right to the documents,[392] the U.S. Supreme

388. In Wigmore's day, corporations were not nearly as commonplace as today; hence, Wigmore's analysis focused solely on personal privileges and cannot simply be applied mechanically to corporations. *See* Bush, *supra* note 277, at 639–41 & n.282.

389. *See* Trammel v. United States, 445 U.S. 40, 43–48 (1980) (rejecting previous formulation of spousal testimonial privilege and holding that only witness-spouse has privilege to refuse to testify against other spouse).

390. Of course, the fact that Wigmore was against the spousal testimonial privilege may be another reason the *Trammel* Court did not consider Wigmore's four factors. *See id.* at 45.

391. *Witten*, 100 F.R.D. at 452 (citing 41 C.F.R. § 1-12.811 (1983)).

392. For example, affirmative action studies must be forwarded to the appropriate government agencies who may disclose them to other government agencies; the government policy is to produce them (unless excepted from discovery), and they are subject to FOIA requests. *See* Flanagan, *supra*

Court's general support of the "right to every man's evidence" is at least as important a rule to uphold.[393] Yet, even this fundamental right to evidence is subverted to the transcendent societal benefits conferred by privileges every time a privilege is recognized.

The *Witten* court, like prior courts, assumed that Title VII is an efficacious private antidiscrimination enforcement mechanism and thus that allowing plaintiffs access to self-critical analysis would increase a plaintiff's chances of proving the motivation and intent necessary to win a disparate treatment discrimination claim. The court acknowledged the difficulty of proving intent in Title VII cases, and indeed based its decision in part on trying to help Title VII plaintiffs acquire the ammunition necessary to meet their burden of proof. But, as argued in Part II of this Article, Title VII litigation is not an effective way of redressing most instances of discrimination. By basing its "cost-benefit" analysis on the faulty assumption that additional evidence of intent furthers antidiscrimination efforts, the *Witten* court ignores the nature and prevalence of unconscious bias and the contribution truly candid self-assessment could make to an antidiscrimination effort.

J. The Demise of the Self-Critical-Analysis Privilege

Federal courts' momentum in destroying the self-critical-analysis privilege had become strong. Courts consistently relied on factors and analyses of previous cases, giving particular weight to *O'Connor* and *Witten*. Courts continued, however, to find additional reasons why the privilege was unnecessary, ineffective, or both.

In *Hardy v. New York News, Inc.*,[394] a New York federal district court relied on the *O'Connor* factors in rejecting the self-critical-analysis privilege in a discrimination and retaliation action brought pursuant to 42 U.S.C. § 1981 and Title VII.[395] The documents were created voluntarily, so the first *O'Connor* factor was not met.[396] The court clarified that "in the area of employment discrimination, virtually every court has limited

---

note 275, at 565–66 & nn.78–82; *see also* 5 U.S.C. §§ 552, 706(2)(a) (1994); 41 C.F.R § 60-40.1 (1998).

393. *See, e.g.*, United States v. Bryan, 339 U.S. 323, 331 (1950) (stating that "'[f]or more than three centuries it has been recognized as a fundamental maxim that the public . . . has a right to every man's evidence'") (quoting 8 J. Wigmore, *Evidence* § 2192 (3d ed. 1940)).

394. 114 F.R.D. 633, 640–42 (S.D.N.Y. 1987).

395. *See id.*

396. *See id.* at 641.

the privilege to information or reports that are mandated by statute or regulation."[397] Therefore, the *O'Connor* court's perception of unfairness in a rule that required companies to engage in self-critical analysis and then surrender that ammunition to the plaintiffs did not apply.[398]

The *Hardy* court found that *O'Connor*'s "balancing" factor militated against the privilege's application.[399] The court referred to the analysis set forth in *Witten* and *O'Connor* that deterrents to candid self-evaluation exist apart from threat of disclosure in discovery and, based thereon, questioned the underlying assumption of the privilege that disclosure would significantly discourage self-critical activity.[400] On balance, the *Hardy* court found that the plaintiffs' need to prove intent outweighed the interest in promoting candid self-analysis and voluntary EEO compliance, especially given the insignificant deterrent effect that the threat of disclosure would have on the document preparation at issue in that case.[401] The court stated that companies have an obligation to comply with the law and as a matter of "sound business management" will take steps to prevent litigation by making evaluations.[402] Thus, any disincentive to conduct candid evaluations based on potential disclosure is contrary to the "basic principles of risk management," including a company's need to resolve minorities' grievances without litigation.[403]

The *Hardy* court did not consider that some businesses will exercise "sound business judgment" and "risk management" by bulletproofing themselves against potential discrimination lawsuits by setting low goals, making no potentially harmful subjective findings of fact, and hiring just enough minorities to meet minimal goals. The *Hardy* court's theory of sound business judgment rests on the premise that only effective antidiscrimination efforts lower the risk of litigation. But the ubiquitous use of "bean counters" by automobile and other product manufacturers belies the notion that companies seek first to comply with the law if there are other ways to minimize "risk"—which generally means "costs" to companies. Human justice, even human life, is but one factor in deciding when and how to comply with the law. To assume companies' interest in

397. *Id.*

398. *See id.*

399. *See id.* at 643.

400. *See id.* at 641.

401. *See id.* at 641–42.

402. *Id.* at 642.

403. *Id.*

risk management necessarily converges with the interests at stake in antidiscrimination legislation is dangerous and naive. By refusing to recognize the privilege, the *Hardy* court not only made truly critical and candid analysis unlikely, but the court actually encouraged companies to cover their tracks in meeting their reporting requirements.[404]

The *Hardy* court also considered the fact that the privilege has never been applied to prevent government agencies from obtaining the documents.[405] Because both the Equal Employment Opportunity Commission (EEOC) and the private plaintiffs sought the documents in this case and because the EEOC indicated it would bring its own motion to compel if the private plaintiffs' motion to compel were denied, the denial of the plaintiffs' motion would neither protect the documents from disclosure nor eliminate the alleged chilling effect.[406] Theoretically, however, if a self-critical-analysis privilege were recognized seriously, it would protect the privileged documents against discovery by anyone, such as is the case with attorney-client privileged documents.

In this particular context, however, it is arguable that the U.S. Supreme Court has spoken. In *University of Pennsylvania v. EEOC*,[407] the Court held that Congress struck a balance in relation to the purported privilege against discovery of peer-review materials in Title VII cases by not addressing the privilege when amending Title VII to include educational institutions.[408] Since the self-critical-analysis privilege could have been addressed in the amendments to Title VII, but was not, the same reasoning regarding peer reviews could apply to the self-critical-analysis privilege. Indeed, at least one magistrate has held that under the principles of *University of Pennsylvania v. EEOC* and Congress' failure to provide for a self-critical-analysis privilege, recognition of the privilege is inappropriate.[409] This is probably the strongest argument

404. Indeed, one company argued that if the self-critical-analysis privilege is not recognized, it would be "forced" to conduct its reports in a manner "inconsistent with the basic goals of affirmative action." Capellupo v. FMC Corp., 46 Fair Empl. Prac. Cas. (BNA) 1193, 1197 (D. Minn. May 3, 1988).

405. *Hardy*, 114 F.R.D. at 643 (quoting FTC v. TRW, Inc., 628 F.2d 207, 210 (D.D.C. 1980)).

406. *See id.* The U.S. Supreme Court has held that private litigants are entitled to documents in the EEOC's possession. *See* EEOC v. Associated Dry Goods Corp., 449 U.S. 590, 604 (1981).

407. 493 U.S. 182 (1990).

408. *See id.* at 194–95 (holding that Congress had already struck appropriate balance of interests in Title VII actions by providing that EEOC can obtain "relevant" evidence).

409. *See* Aramburu v. Boeing Co., 885 F. Supp. 1434, 1440 (D. Kan. 1995).

against recognizing the privilege in cases relating to Title VII compliance.[410]

In *Capellupo v. FMC Corp.*,[411] which was decided one year after *Hardy*, a Minnesota federal district court rejected the privilege after FMC boldly claimed that potential disclosure of self-evaluative materials could cause it to be "forced to exercise the substantial discretion left to it by the federal regulations *in a manner inconsistent with the basic goals of affirmative action.*"[412] The court called FMC's position "an affront" and emphasized the company's obligation to comply with the "law of the land."[413] Ironically, FMC's position essentially restates the *Dickerson* court's reasoning and is consistent with the theory underlying the traditional justification. The *Capellupo* court's response implicitly rejects the notion that privileges produce candor and, as argued above, rather naively exhibits an expectation that companies consider their legal duties over their bottom lines. The court's position on this fundamental basis for the self-critical-analysis privilege—the creation of better EEO policy through candid self-evaluation—underscores the complete rejection, based upon the reasoning in cases such as *O'Connor* and *Witten*, of the core premise of the self-critical-analysis privilege by this time.

The *Capellupo* court's opinion is a harbinger of the way federal courts in the 1990s relied largely on prior tests and analyses in determining that no self-critical-analysis privilege exists in employment discrimination cases. For example, although *Witten*'s analysis is questionable and places undue emphasis on Wigmore's factors, later courts relied heavily on the lack of expectation of confidentiality to find the privilege inappropriate. After discussing *Witten* and Wigmore's four "necessary" conditions, the *Capellupo* court held that the defendant had "no expectation of privacy" in AA plans and related documents because the regulations clearly state that such documents must be made available to the Office of Federal Contract Compliance Programs and that the information can be disclosed to others, including private parties.[414] The *Capellupo* court also noted that

410. It would not apply, however, to self-critical evaluations unrelated to Title VII compliance, such as purely voluntary evaluations, including unconscious-bias testing.

411. 46 Fair Empl. Prac. Cas. (BNA) 1193 (D. Minn. May 3, 1988).

412. *Id.* at 1197 (emphasis added). The documents at issue evaluated FMC's progress under its AA plan and discussed the company's AA goals in hiring women and minorities. *See id.* at 1194.

413. *Id.* at 1197 (citing Hardy v. New York News, Inc., 114 F.R.D. 633 (S.D.N.Y 1987)).

414. *See id.* (citing 41 C.F.R. §§ 60-40.1, 60-2.12(a)). Note that, like *O'Connor*'s factors, Wigmore's factors are referred to as "necessary" in spite of the fact that no single factor is necessary under the broad language of FRE 501.

the U.S. Supreme Court has held that individual plaintiffs are entitled to EEOC investigation information and appeared to base its denial of the self-critical-analysis privilege primarily on Wigmore's first factor, which concerns the fact that the communication did not originate in confidence.[415]

K. Rejection of the Self-Critical-Analysis Privilege

In 1990, federal courts began completely rejecting the self-critical-analysis privilege in employment discrimination cases.[416] In *Martin v.*

---

415. *See id.* (citing EEOC v. Associated Dry Goods Co., 449 U.S 590, 604 (1981)). The *Capellupo* court also noted the similarity of defendant's argument to FRE 407, which renders evidence of measures taken subsequent to an alleged harmful event inadmissible to prove negligence or culpable conduct in connection with the event. *See id.* at 1197–98. The court distinguished FRE 407 because it is a rule of evidence not applicable to pretrial discovery. "The better view is to permit discovery, not only because Rule 407 is essentially a rule of public policy rather than of relevancy, but also because subsequent remedial measures might be admissible to prove a consequential, material fact in issue . . . ." *Id.* (citing 2 Weinstein, *Evidence* ¶ 407[7], at 407-37 to 407-38). In the instant case, the court said that proving the defendant's intent to discriminate in a Title VII suit is precisely the kind of "consequential material fact in issue" that Weinstein refers to. *Id.* Whether the documents indicate a failure to perform under an AA plan is relevant to the issue of discriminatory intent under Eighth Circuit law. *See id.* (citing Craik v. Minnesota State Univ. Bd., 731 F.2d 465, 472 (8th Cir. 1984)); *see also* Aramburu v. Boeing Co., 885 F. Supp. 1434 (D. Kan. 1995); Siskonen v. Stanadyne, Inc., 124 F.R.D. 610, 612 (W.D. Mich. 1989) (rejecting self-critical-analysis privilege under Michigan law). The *Capellupo* court did not, however, notice that the U.S. Supreme Court did not address the expectation-of-confidentiality factor or even mention any Wigmore factor in making its decision in *Trammel* just one year before ruling on the availability of EEOC documents to private litigants. Rather, the *Capellupo* court relied on *Witten* and, like *Witten*, places undue emphasis on Wigmore's first factor.

416. *See* Spencer Sav. Bank v. Excell Mortgage Corp., 960 F. Supp. 835, 838, 844 (D.N.J. 1997) (setting forth "proper test" for recognizing common law privilege and holding "a self-critical analysis privilege does not exist at federal common law"); *Aramburu*, 885 F. Supp. at 1441 ("[I]t is inappropriate to recognize the privilege of self-critical analysis in Title VII cases."); Tharp v. Sivyer Steel Corp., 149 F.R.D. 177, 182 (S.D. Iowa 1993) (holding "the 'self-critical analysis' privilege should not be recognized in the field of employment discrimination litigation"); Etienne v. Mitre Corp., 146 F.R.D. 145, 147–49 (E.D. Va. 1993) (calling privilege "largely undefined" in age discrimination suit and finding it inapplicable as contrary to public interests at stake in employment discrimination cases); Williams v. Vulcan-Hart Corp., 136 F.R.D. 457, 459–60 (W.D. Ky. 1991) (holding no self-critical-analysis privilege under Kentucky law for violation of state antidiscrimination statute); Martin v. Potomac Electric Power Co., 58 Fair Empl. Prac. Cas. (BNA) 355 (D.D.C. May 25, 1990). *But see* Whittingham v. Amherst College, 164 F.R.D. 124, 130 (D. Mass. 1995) (recognizing privilege, but stating that document at issue "hardly amounts to the kind of self-critical analysis deserving of the court's protection"); Abel v. Merril Lynch & Co., 91 Civ. 6261, 1993 WL 33348, at *2 (S.D.N.Y. Feb. 4, 1993) (regarding demographic statistical data in disparate treatment and disparate impact ADEA claim and stating that "given that this Court has doubts as to the viability of this privilege, . . . and that the reports sought by Plaintiff contain only 'non-evaluative facts, statistics, or other data,' the Court declines to apply the so-called privilege of self-critical analysis in this context"); Vanek v. Nutrasweet, 59 Empl. Prac. Dec. (CCH) ¶ 41,600, at

Potomac Electric Power Co.,[417] a District of Columbia federal district court held that no privilege for self-critical analysis exists protecting documents from discovery in private employment discrimination cases.[418] The *Martin* court condoned the *Bredice* decision, stating that when a situation concerns public health and safety, it is critical not to block efforts to identify and "correct dangerous conditions."[419] However, the *Martin* court found that private employment discrimination suits do not involve similarly strong policy considerations to justify shielding self-critical documents from discovery.[420] After noting the general standards for recognizing a privilege under FRE 501 and relying on the analyses of *O'Connor* and *Witten*, the *Martin* court went through the reasons why the self-critical-analysis privilege would not affect internal evaluations or discourage candid preparation of government reports.[421]

Fundamentally, the *Martin* court did not believe that disclosure would hinder progress toward equal employment opportunity. This is because either (1) companies engage in aggressive self-critical analysis, resulting in better EEO policies and a lack of fear of litigation or disclosure, or (2) if the degree of candor in the reports is unrelated to success in creating equal employment opportunities, then society should not care about encouraging candor:

71,455 (N.D. Ill. 1992) (recognizing split of authority and discussing privilege as "not yet fully established in the law" but ordering discovery); Steinle v. Boeing Co., 62 Fair Empl. Prac. Cas. (BNA) 272, 279 (D. Kan. Feb. 3, 1992) (acknowledging split of authority regarding self-critical-analysis privilege). Some courts, however, continued to recognize the privilege to a limited degree in the employment context after 1990. *See, e.g.*, Troupin v. Metropolitan Life Ins. Co., 169 F.R.D. 546, 549 (S.D.N.Y. 1996) (finding privilege applied as to narrative, evaluative, or analytical portions of self-evaluative documents pursuant to self-critical-analysis privilege despite "questionable force of the reasoning behind the privilege"); Brem v. Decarlo, 162 F.R.D. 94, 101–02 (D. Md. 1995) (protecting physician's opinion about competence of former residents based on Maryland's peer-review statute and self-critical-analysis privilege in Title VII suit, but distinguishing documents at hand from documents unprotected in employment context because documents at hand were peer-review type documents); Cobb v. Rockefeller Univ., 58 Fair Empl. Prac. Cas. (BNA) 184, 186 (S.D.N.Y. Oct. 21, 1991) (holding review process portion of affirmative action plan privileged). The author of this Article could find no cases outside of the southern district of New York after 1990 where the self-critical-analysis privilege was held to prevent discovery of EEO assessment in an employment discrimination suit.

417. 58 Fair Empl. Prac. Cas. (BNA) 355 (D.D.C. May 25, 1990).

418. *See id.* at 359.

419. *Id.* at 356.

420. *See id.* (citing FRE 407).

421. *See id.* at 357–59.

> Thus, either disclosure will not give companies a disincentive to be candid, or disclosure will give a disincentive to be candid but a decrease in candor will not slow progress toward the goal of equal employment opportunity. In either case, disclosure will not undermine the goal of equal employment opportunity.[422]

This decision reflects the most common attack on the traditional justification—that changes in privilege law have no impact on behavior and ultimately employers' EEO attitudes and efforts will not be affected by privilege law.

Finally, the *Martin* court found that private litigation is itself an important means of promoting EEO. Self-evaluative documents improve private plaintiffs' ability to prove discrimination, furthering EEO goals by allowing discovery of the documents.[423] For these reasons, the court concluded that protecting EEO reports from discovery by private plaintiffs would not promote sufficiently important interests to outweigh plaintiffs' need for probative evidence.[424] This case and others denouncing the privilege place trust in private litigation as a means of promoting equal employment opportunity and ultimately find private litigants' need for probative evidence a more compelling consideration.[425]

The history of the self-critical-analysis privilege in employment discrimination cases demonstrates the inherent tension in privilege law analysis between the need for confidentiality and the need for evidence. Because the self-critical-analysis privilege is traditionally grounded in utilitarian justifications, courts that refused to uphold the privilege did so

---

422. *Id.* at 359.

423. *See id.*

424. *See id.* The court went on to note a number of "additional considerations" that further convinced it that its decision was correct, including (1) because employers may use positive reports to prove their good intentions, fairness dictates that plaintiffs should have the same opportunity with respect to negative reports; (2) there is no expectation of confidentiality; (3) the self-critical-analysis privilege is not grounded in any constitutional provision and has no statutory or historical basis, and although privileges do not require such a basis, the lack thereof carries some weight; and (4) no court of appeal has recognized the privilege and prior decisions in the Circuit, and other courts have argued against the privilege in the employment context. *See id.* at 359–60.

425. *See id.* at 360–61; *see also* Gray v. Board of Higher Educ., 692 F.2d 901, 904 (2d Cir. 1982) (holding where proof of intent is required, tension between plaintiff's need for evidence and employer's confidentiality tips in favor of plaintiff) (citing Herbert v. Lando, 441 U.S. 153, 170–75 (1979)); Etienne v. Mitre Corp., 146 F.R.D. 145, 147 (E.D. Va. 1993) ("Over the years, this balancing of public and private interests has become the essential consideration when a court decides whether the [self-critical-analysis] privilege should prevent disclosure of relevant information.").

primarily by attacking the basic premise of the traditional rationale that the privilege would affect people's behavior positively. The courts' analyses have been unjustifiably reliant on prior holdings, elevating fact-specific holdings and unsound analysis to privilege "tests" and policy statements that are unrelated to true privilege law analysis and inconsistent with the courts' obligation to develop rules of privilege in accordance with sound public policy. Factual situations in early cases, such as the fact that most documents protected by the privilege prior to *Webb* were government mandated, became entrenched in self-critical-analysis case law. Policy analysis was thus lost to the prior cases' factual similarities. One of the strongest arguments for not recognizing the privilege in later cases—that no expectation of confidentiality exists given that the employer had notice that the documents must be produced to government agencies—does not apply to voluntarily created documents. Yet, by the time the courts' opinions focused on this confidentiality expectation argument, the requirement that the documents be government mandated was already entrenched in the case law. There was no careful judicial analysis of whether purely voluntary EEO self-critical-analysis documents should be privileged. If there had been, courts would have probably protected purely voluntary critical self-assessments, likely encompassing voluntary unconscious-bias testing.

Nonetheless, fundamental justifications for recognizing a self-critical-analysis privilege emerge from these cases and build an argument for protecting the results of unconscious-bias testing. First, when the desired communications are vulnerable to "chilling," recognition of a privilege is more likely. Second, when a great benefit to society is believed to result from protecting the communications, or conversely, when great harm could result from not protecting the communications, a privilege is more probable. Third, when little or no evidence is lost on account of the privilege, it is more likely to be recognized.

VI. COURTS SHOULD RECOGNIZE A PRIVILEGE FOR UNCONSCIOUS-BIAS TESTING

The immediate truth exposed by unconscious-bias testing is subordinate to the greater, but less immediate, truth that will come as a result of increased understanding about what motivates individuals in intergroup relations. Recognizing a self-critical-analysis kind of privilege for unconscious-bias testing protects the more transcendent, individual and societal truth-finding process necessary to achieve interracial and intergroup peace from the more temporal objectives of individual

plaintiffs in discrimination lawsuits. Considering that unconscious-bias test results do not prove intent under Title VII,[426] any loss of evidence from recognizing a privilege is minimal and outweighed by the benefits. Does traditional privilege law theory support a self-critical-analysis privilege for employer testing of cognitive bias and subjective employer analysis of its employees' unconscious mental processes? What values must we balance in analyzing this issue?

Before embarking on an analysis of why a privilege for unconscious-bias testing is warranted by the policies expressed by courts in relation to the formerly recognized privilege for self-critical analysis, it is important to get a basic understanding of privilege law theory. Selected U.S. Supreme Court cases on privilege law clarify justifications for a new privilege and guide privilege law analysis under FRE 501. After discussing privilege law theory and case law, this Part argues that recognition of a qualified privilege for unconscious-bias testing is appropriate. Finally, this Part addresses objections to unconscious-bias testing.

A. The Theory of Privileges

Privileges are generally viewed as impediments to truth. They subordinate the truth-finding process to higher societal interests and thus "impede the realization of a central objective of the legal system in order to advance other, often less immediate, goals."[427] Two primary justifications exist in defense of privileges.[428] The first and most

426. Disparate impact theory does not require proof of intent, just an employer's facially neutral practice that impacts a protected group more harshly than other groups. Unconscious bias would not help plaintiffs in disparate impact cases because attitude, motivation, and intent are not at issue. Under the ADEA, unconscious-bias testing for age bias could constitute proof, since "stereotyping" of older people is illegal under the ADEA. However, ADEA cases have generally required tangible statements demonstrating stereotyping, so purely unconscious bias demonstrated by scientific test results would probably not be admissible; they do not prove anything except a potential to stereotype. *See generally* Krieger, *supra* note 6. Also, the studies have not shown a clear and definite connection between unconscious bias and discriminatory actions, making the unconscious-bias test results probably more prejudicial than probative. But even if the testing was probative evidence, this author maintains that the benefits to be gained from keeping test results out of litigation outweigh the loss of evidence.

427. *See Developments in the Law—Privileged Communications*, *supra* note 327, at 1454.

428. Two other theories hold that neither of these justifications is coherent because they both attempt to rationalize an incoherent body of law; instead, privileges can be explained only in terms of political power. "Power Theory" argues that privilege law is "special treatment won by the power of those privileged." *Id.* at 1493–98. "Image Theory" holds that privilege law can be understood as a

influential is the traditional, utilitarian justification enunciated by Wigmore.[429] This theory holds that privileges are justified only if the benefit gained by protecting the information outweighs the impediment in the search for truth.[430] The goal is to encourage communication in a relationship, but the theory is based on the unproven assumption that the existence of the privilege encourages people to speak more candidly because they know that what they say is protected from disclosure. Thus, the contention is that the relationship of the parties to the communication will suffer if the communications are not privileged.[431] Wigmore set forth four conditions to establish a privilege: (1) the communications must originate in confidence; (2) the confidentiality must be essential to the full maintenance of the relation between the parties; (3) the relation must be one that the community strongly wants to foster; and (4) the injury to the relation caused by the disclosure of the communications must be greater than the benefit gained by the correct disposal of litigation.[432]

Wigmore was a strong believer in the duty to testify and narrowly construed the fourth factor to constrict application of the privilege to worthy communications[433] and thus to systemic harms. The traditional justification is not concerned about injury to specific litigants, but rather balances the encouragement of communications in the relevant class against the cost of its obstruction to truth finding. Most courts have followed Wigmore's approach, requiring systemic harm in the balancing test established by the fourth factor.[434] By considering only extrinsic social policy unrelated to the interests of the individual litigants in a particular dispute, the justification "elevates the interests advanced by privileges to the same plane as the societal interest in ascertaining the

---

means of preserving the image and legitimacy of the legal system, avoiding embarrassment to the legal system by masking its inability to compel obedience and minimizing the possibility that facts discovered after trial would undermine the judgments made. *See id.* at 1498–1500.

429. *See id.* at 1472; *see also* Ronald J. Allen et al., *Evidence* 990 (2d ed. 1977); 8 John Wigmore, *Evidence* § 2285 (J. McNaughton ed. 1961).

430. *See* Wigmore, *supra* note 429, § 2285.

431. *See* Allen et al., *supra* note 429, at 990.

432. *See* Wigmore, *supra* note 429, § 2285.

433. *See Developments in the Law—Privileged Communications*, *supra* note 327, at 1472–74.

434. *See id.* at 1473; *see also* Jaffee v. Redmond, 518 U.S. 1, 10 (1996) (holding privileges must serve public ends).

truth."[435] Examples of privileges justified under this theory include the attorney-client privilege and medical and counseling privileges.[436]

The most common attack on this justification is that the privilege does not affect people's behavior: either people do not know the law, so the unknown privilege does not encourage communication, or if they do know the law, people's need to communicate would overcome their fear of disclosure.[437] To the extent data exists, it is mixed. Some studies have found that most people believe that they communicate more openly when they are assured of confidentiality, and this may be especially true of legally sophisticated corporate actors.[438] Other studies demonstrate that whether a psychotherapist-patient privilege is recognized has little impact on patients.[439] Of course, the effect of the privilege remains unproven and is, in fact, probably incapable of proof. Social psychology studies indicate that people are often unable to say what really motivated them. Thus, empirical studies relying on self-reporting about whether the existence of a privilege affected the privilege-holders' behavior are inherently indeterminate.

The rejoinder to this attack is that the costs imposed by a privilege are no less speculative. If the privilege encouraged a communication that would not have otherwise occurred and then shielded that communication from discovery, the result is a net "wash" for evidentiary purposes. The exceptions to privileges and concepts such as waivers

---

435. *Developments in the Law—Privileged Communications*, *supra* note 327, at 1474.

436. *See id.* at 1501, 1530. Note that other theories support these privileges as well.

437. *See id.* at 1474–80. For example, a person's need to communicate freely with a doctor to gain the utmost in medical advice for her condition probably outweighs any embarrassment concerns over disclosure of the information.

438. *See* Daniel W. Shurman & Myron S. Weiner, *The Privilege Study: An Empirical Examination of the Psychotherapist-Patient Privilege*, 60 N.C. L. Rev. 893 (1982); Note, *Functional Overlap Between the Lawyer and Other Professionals: Its Implications for the Privileged Communications Doctrine*, 71 Yale L.J. 1226 (1962); *see also* Bush, *supra* note 277, at 637–38 (citing studies showing that 93% of corporate counsel and 79% of executives indicated concern over attorney-client privilege at least once in preceding five years; 89% of corporate counsel and 75% of executives believe attorney-client privilege fosters candor; and 70% of executives indicated that if "control group" test were used to apply attorney-client privilege, they would restrict or prohibit attorney access to lower employees) (citing Vincent C. Alexander, *The Corporate Attorney-Client Privilege: A Study of the Participants*, 63 St. John's L. Rev. 191, 260–61 (1989)); Jessica G. Weiner, *"And The Wisdom to Know the Difference": Confidentiality vs. Privilege in the Self-Help Setting*, 144 U. Pa. L. Rev. 243, 260–64 & nn.120–22 (1995).

439. *See* Edward Imwinkelried, *The Rivalry Between Truth and Privilege: The Weakness of the Supreme Court's Instrumental Reasoning in* Jaffee v. Redmond, 518 U.S. 1 (1996), 49 Hastings L.J. 969, 976 (1998).

diminish the costs of the privilege and can provide access to the most important evidence.[440] Furthermore, communications that are determined not to be privileged cannot be forced out of unwilling witnesses, so compelled disclosure may often result in false testimony. Finally, the loss of a single piece of evidence will rarely make or break a case.

The emergent justification for privileges is the privacy rationale.[441] Here the focus is not on the privilege's beneficial impact to society, but on the protection that the privilege affords to individual privacy. That is to say, the interest in securing privacy and confidentiality to the individuals involved in the communication is itself considered justification for whatever obstruction to truth seeking results.[442] Some say that the marital communications privilege is based on the privacy rationale so as to recognize and protect intimate aspects of the marital relationship.[443] The privacy rationale raises these three issues: (1) whether people need to keep certain communications confidential; (2) whether this need is socially cognizable; and (3) whether the privacy interest outweighs the need for evidence.[444]

Privacy is said to serve several societal objectives, such as promoting self-evaluation, permitting emotional release, and allowing for personal autonomy.[445] Respecting privacy against compelled disclosure is also an end in itself, as it avoids exposing embarrassing intimacies to the public and protects against forced breaches of confidential relationships. Whether an alleged privacy interest is legally cognizable turns on what courts deem worthy of protection. The balance between privacy and society's need to find the truth depends on the facts and normative considerations involved. Categorical balancing makes the privacy rationale as indeterminate as the traditional justification; neither justification can be resolved empirically.

---

440. *See Developments in the Law—Privileged Communications*, *supra* note 327, at 1477–79.

441. *See* Imwinkelried, *supra* note 439, at 988; *see also* Thomas G. Krattenmaker, *Testimonial Privileges in Federal Courts: An Alternative to the Federal Proposed Rules of Evidence*, 62 Geo. L.J. 61 (1973); David W. Louisell, *Confidentiality, Conformity and Confusion: Privileges in Federal Court Today*, 31 Tul. L. Rev. 101 (1956).

442. *See Developments in the Law—Privileged Communications*, *supra* note 327, at 1480.

443. *See* Allen et al., *supra* note 429, at 990; *see also* Charles L. Black, Jr., *The Marital and Physician Privileges—A Reprint of a Letter to a Congressman*, 1975 Duke L.J. 45, 47.

444. *See Developments in the Law—Privileged Communications*, *supra* note 327, at 1481.

445. *See, e.g.*, Imwinkelried, *supra* note 439, at 985–88.

A utilitarian balancing of all interests may be more useful than the "mutually exclusive" traditional and privacy alternatives.[446] The case for a new privilege is strengthened by considering the privacy rationale and the traditional justification in tandem. In the context of unconscious-bias testing, the traditional justification can be supplemented by the need to protect people's psyches from invasive exploitation.

The burden is on those advocating privileges to show that the benefit to society or to the individual is greater than the costs of allowing the privilege.[447] But the rationale used to justify the privilege will affect the way the privilege is applied. The traditional justification requires predictable application; otherwise, the benefit of candid communication will be chilled by fear that a court will find the communication unprotected in a particular case.[448] On the other hand, because the privacy rationale does not focus on behavioral consequences of the privilege, certainty in application is not a major issue.[449] Thus, privileges justified by the traditional justification should be absolute and certain, while privacy concerns can be addressed on a case-by-case basis. Because the main goal of the privilege for unconscious-bias testing is to improve the content of communications, as opposed to fostering a relationship, a traditional, utilitarian approach should be used so that application is consistent and employers can rely upon the privilege to protect their self-critical employment assessments. Thus, while a privacy rationale may strengthen an argument for recognizing a privilege based upon the tested individual's privacy,[450] the primary justification must be utilitarian.

Neither a strict rule nor a case-by-case approach is without potential drawbacks. A predictable rule invites abuse because it allows the "bad man" to "walk the line." Also, problems of underinclusiveness and overinclusiveness will likely arise. Some commentators counter that

---

446. The two primary justifications are generally regarded as mutually exclusive alternatives even though they can be compatible and share similar methodologies. *See Developments in the Law—Privileged Communications*, *supra* note 327, at 1483–86.

447. *See id.* at 1480; *see also* Swidler & Berlin v. United States, 524 U.S. 399, 409–10 (1998).

448. *See, e.g.*, *Swidler & Berlin*, 524 U.S. at 409.

449. *See Developments in the Law—Privileged Communications*, *supra* note 327, at 1486–87.

450. A privilege against discovery of an individual's test results could be grounded in the privacy rationale, but because here the goal is to encourage *employers* to test employees, an individual's privacy interests are not directly on point. Nonetheless, recognizing an individual's privacy is important for a full analysis of the privilege. For a discussion about unconscious-bias testing and the right to privacy, see *supra* Part VI.D.

certainty encourages private activity to follow a desired pattern.[451] Applying uncertain standards on a case-by-case basis risks arbitrary results, but discretion can beneficially reduce underinclusive and overinclusive effects of the privilege. The best privilege may be one that appears certain in form, so that communication is not chilled, but is uncertain in application so judges can fine-tune the privilege to reduce inappropriate outcomes.[452]

B. Standards for Recognizing and Analyzing Privileges

Privilege law embodies federal common law privileges and a diverse collection of state law privileges mostly of statutory origin. As part of the Federal Rules of Evidence adopted by Congress in 1975, FRE 501 governs privilege law in federal courts.[453] Congress declined to adopt the Proposed Rules of Evidence,[454] which set forth nine discrete privileges,[455] instead opting for the more flexible language of the current FRE 501.

FRE 501 grants federal courts discretion to modify common law privileges and to create new ones. As the U.S. Supreme Court has explained: "In rejecting the proposed rules and enacting Rule 501, Congress manifested an affirmative intention not to freeze the law of

---

451. *See, e.g.*, Upjohn Co. v. United States, 449 U.S. 383, 393 (1981); Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 89 Harv. L. Rev. 1685, 1688–89 (1976).

452. *See Developments in the Law—Privileged Communications*, *supra* note 327, at 1489.

453. FRE 501 provides:

> Except as otherwise required by the Constitution of the United States or provided by Act of Congress or in rules prescribed by the Supreme Court pursuant to statutory authority, the privilege of a witness, person, government, State, or political subdivision thereof shall be governed by the principles of the common law as they may be interpreted by the courts of the United States in the light of reason and experience. However, in civil actions and proceedings, with respect to an element of a claim or defense as to which State law supplies the rule of decision, the privilege of a witness, person, government, State or political subdivision thereof shall be determined in accordance with State law.

454. The Proposed Rules were drafted by the Advisory Committee and proposed by the U.S. Supreme Court in 1971. After considerable controversy and criticism, Congress adopted the new Federal Rules of Evidence in 1975. *See* Allen et al., *supra* note 429, at 992–95.

455. The privileges proposed were required reports, attorney-client confidential communications, psychotherapist-patient confidential communications, prevention of spousal testimony, clergy-communicant confidential communications, political votes, trade secrets, state secrets and other official information, and the identity of an informer. *See* Proposed FRE 502–510 and Advisory Committee Notes, 51 F.R.D. 260, 260–80 (1971). For a thorough discussion of FRE 501's history, see Edward J. Imwinkelried, *An Hegelian Approach to Privileges Under Federal Rule of Evidence 501: The Restrictive Thesis, The Expansive Antithesis, and the Contextual Synthesis*, 73 Neb. L. Rev. 511 (1994).

privilege. Its purpose rather was to 'provide the courts with the flexibility to develop rules of privilege on a case-by-case basis' and to leave the door open to change."[456] Thus, Congress recognized that the evidence rules needed the flexibility to keep pace with society's changing values and relational dynamics. Indeed, FRE 501 confers to federal courts a legislative-type function to change privilege law in accordance with changing societal norms and is thus an effective way to keep privilege law up-to-date without the need for congressional action each time privilege law needs modification.

The extent to which new privileges are recognized under state law varies.[457] State privileges are usually created by state legislatures that consider the same utilitarian and privacy policy rationales that animate federal court privilege review under FRE 501. These shared policy concerns encourage crossover of privilege analysis. States often borrow federal privilege analysis to determine the propriety of recognizing a privilege.[458] Federal courts in federal question cases often look to state law for guidance in the area of privilege, and commentators have condoned such practice in the absence of strong countervailing federal policy.[459] Therefore, all of the arguments supporting federal court recognition of the privilege also support state recognition of the privilege.

Trammel v. United States[460] exemplifies the Court's flexible "all things considered" approach to effectuating FRE 501. In *Trammel*, the Court modified the previous rule barring the testimony of one spouse against the other unless both consented and held that, under modern policy considerations, only the witness-spouse holds a privilege to refuse

456. Trammel v. United States, 445 U.S. 40, 47 (1980) (quoting statement by Rep. Hungate).

457. *See McCormick on Evidence* § 76.2, at 109 (John William Strong ed., 4th ed. 1992).

458. *See, e.g.*, Siskonen v. Stanadyne, Inc., 124 F.R.D. 610, 612 (W.D. Mich. 1989) (analyzing federal self-critical-analysis privilege and concluding that no state privilege would be recognized); *In re* Subpoena Duces Tecum to the Wayne County Prosecutor, 477 N.W.2d 412 (Mich. Ct. App. 1991) (analyzing federal and state privilege law to determine applicability of work product doctrine to work of nonparty's counsel) (Griffin, J., dissenting); State v. Cecos Int'l, Inc., 583 N.E.2d 1118 (Ohio Ct. App. 1990) (noting test for determining whether adoption of privilege is appropriate relied on factors set forth in federal cases); Commonwealth v. Hancharik, 633 A.2d 1074 (Pa. 1993) (relying on federal privilege analysis of marital communications privilege and privilege against adverse spousal testimony).

459. *See Developments in the Law—Privileged Communications*, *supra* note 327, at 1470 & n.127.

460. 445 U.S. 40 (1980).

to testify against the other spouse.[461] The Court clarified the authority to continue the development of privileges, but maintained that privileges contravene the principle of the "public['s] . . . right to every man's evidence"[462] and should be accepted only to the extent that excluding relevant evidence advances a public good that overcomes the interest in ascertaining truth.[463] The Court found that in light of "reason and experience," the privilege against adverse spousal testimony did not promote interests sufficient to outweigh the need for probative evidence in the administration of criminal justice.[464] The privilege against adverse spousal testimony stood in marked contrast to the attorney-client, priest-penitent, and physician-patient privileges, all of which are limited to confidential communications.[465] Furthermore, the privilege originated in the now-obsolete concept of a woman as chattel with no legal identity separate from her husband's. Even the contemporary justification for the privilege—to protect marital harmony—is unpersuasive; if one spouse is willing to testify against the other, "there is probably little in the way of marital harmony for the privilege to preserve."[466]

The Court's methodology in *Trammel* was to recognize historical tenets and scholarly commentary where relevant and to employ FRE 501 to make privilege law decisions that comport with contemporary societal norms and the Court's experience. While the Court acknowledged Dean Wigmore's harsh criticism of the spousal privilege and appears even to adopt Wigmore's fourth "factor" under the traditional justification,[467] it remains unbound by prior privilege analysis or "factors" per se; instead, the Court relied on a more current, holistic analysis in light of its own reason and experience.

In *Jaffee v. Redmond*,[468] a recent case recognizing a psychotherapist-patient privilege,[469] the U.S. Supreme Court showed a pronounced reliance on privilege law analysis rooted in the traditional utilitarian

461. *See id.* at 53.

462. *Id.* at 50 (quoting United States v. Bryan, 339 U.S. 323, 331 (1950)).

463. *See id.*

464. *Id.* at 51–53.

465. *See id.* at 51.

466. *Id.* at 52.

467. *See id.* at 45.

468. 518 U.S. 1 (1996).

469. *See id.* at 15. The Court held that the privilege extends to social workers as well because all of the same policy considerations applied with equal force, and as a practical matter, less-expensive social workers are utilized more often by the poor. *See id.* at 15–16.

justification. The Court held that FRE 501 requires that privileges serve public ends,[470] a position consistent with the traditional utilitarian justification and reiterated that an exception to the general rule prioritizing truth finding can be justified only by a transcendent public good served by the privilege.[471]

The *Jaffee* Court found that because psychotherapy cannot be effective without confidentiality, important private interests in mental well-being are served by the privilege.[472] The Court also determined the mental health of our citizenry to be a public good of transcendent importance.[473] Furthermore, the likely evidentiary benefit from denying the privilege is modest because much of the sought-after communications would not occur but for the privilege.[474]

The *Jaffee* Court feared the Seventh Circuit's "balancing" approach would eviscerate the effectiveness of the privilege:

> [I]f the purpose of the privilege is to be served, the participants in the confidential conversation must be able to predict with some degree of certainty whether particular discussions will be protected. An uncertain privilege, or one which purports to be certain but results in widely varying applications by the courts, is little better than no privilege at all.[475]

Here, again, the Court focuses on traditional justification arguments.

The U.S. Supreme Court has continued to reject balancing tests, most recently in the attorney-client privilege context in *Swidler & Berlin v.*

---

470. *See id.* at 11 (citing Upjohn Co. v. United States, 449 U.S. 383, 389 (1981)).

471. *See id.* at 9; *see also, e.g.*, Benton v. Superior Court, 897 P.2d 1352 (Ariz. Ct. App. 1994) (considering societal utilitarian policy concerns by holding that public interest in protecting victims of crime outweighed privacy interest reflected in physician-patient privilege).

472. *See Jaffee*, 518 U.S. at 10–11.

473. *See id.* at 11.

474. *See id.* at 12. The Court also stated that the fact that all 50 states and the District of Columbia have enacted some form of psychotherapist-patient privilege made it appropriate for federal courts to recognize the privilege, while not recognizing the privilege in federal courts could frustrate the purposes of the state legislation. *See id.* The fact that the psychotherapist-patient privilege was one of the nine privileges proposed by the Judicial Conference Advisory Committee in 1972 reinforces the states' uniform judgment that the privilege is appropriate. *See id.* at 14. *But see* Imwinkelried, *supra* note 439, at 980–82 (arguing that *Jaffee* Court misconstrued empirical data on whether the privilege encourages communications that otherwise would not have occurred, and that contrary to Court's opinion, recognition of privilege is much more costly in terms of lost evidence than Court assumes).

475. *Jaffee*, 518 U.S. at 18 (citing Upjohn Co. v. United States, 449 U.S. 383, 393 (1981)).

United States.[476] The Court justified extending the attorney-client privilege to protect communications after death, not through balancing, but by finding that full candor in the attorney-client relationship could not be otherwise assured.[477] The balancing approach is, of course, the usual method of applying privileges under the privacy rationale: the individual's specific privacy interest is balanced against the more general societal need for evidence.[478] By repeatedly rejecting a balancing approach in favor of more clear demarcations of the privilege's parameters, the Court implicitly rejected the privacy rationale in favor of the traditional justification and its philosophical underpinnings.

Although FRE 501's language is liberal, the U.S. Supreme Court requires substantial justification to recognize a privilege, particularly where Congress has declined an opportunity to provide protection. In *University of Pennsylvania. v. EEOC*,[479] the U.S. Supreme Court was asked to recognize a common law privilege protecting academic peer-review materials from subpoena.[480] The Court affirmed the Third Circuit's holding that a university does not enjoy a special privilege that would require a judicial finding of particularized necessity beyond mere relevance before pertinent peer review could be disclosed to the EEOC.[481] The Court reiterated the general rule that privileges are created only when they promote interests that outweigh the need for probative evidence.[482] The Court also stated that FRE 501's authority should not be exercised expansively,[483] especially where, as here, Congress did not provide for the privilege even after it had considered the relevant competing concerns when it extended Title VII to educational institutions.[484] Congress had already struck the balance of interests: unless specified otherwise in the statute, the EEOC may obtain

476. 524 U.S. 399, 409 (1998).

477. *See id.* at 408.

478. *See supra* notes 441–52 and accompanying text.

479. 493 U.S. 182 (1990).

480. *See id.* at 188.

481. *See id.* at 189–92.

482. *See id.* at 189.

483. *See id.* at 189–92.

484. *See id.* at 192–93. Moreover, the fact that Congress afforded a "modicum of protection" for confidential records such as peer-review materials in 42 U.S.C. § 2000e-8 (1994), which made it unlawful for EEOC employees to make public any information obtained pursuant to its discovery powers, weakened the petitioner's contentions that Title VII's subpoena enforcement provisions do not give the commission an unqualified right to acquire all "relevant" evidence and that the Court should create a privilege for safeguarding confidential records. *Id.*

"relevant" evidence.[485] The Court feared that recognizing a privilege for universities may inspire a "wave of similar privilege claims" by other employers who further speech and learning in society, such as writers and musicians.[486] Finally, the fact that the petitioner's privilege claim lacked a constitutional, statutory, or historical basis distinguished it from the precedents the petitioner cited in support of its privilege claim.[487]

The Court recently discussed the burden under Fed. Evid. R. 501 of showing that "reason and experience" require a departure from the prevailing common law rule. In *Swidler & Berlin*, Deputy White House Counsel Vincent Foster consulted an attorney about an investigation by the Office of Independent Counsel.[488] The attorney took handwritten notes of the meeting with Foster, and nine days later, Foster committed suicide.[489]

The grand jury subpoenaed the attorney's notes, but the law firm of Swidler & Berlin invoked the attorney-client privilege and work product doctrine. The federal district court refused to order production of the notes.[490] The Court of Appeals for the District of Columbia reversed, concluding that the attorney-client privilege was not absolute, as some posthumous exceptions to the privilege have been recognized, such as the testamentary exception.[491] The appellate court felt that the risk of posthumous revelation in the criminal context would have little or no chilling effect on client communications, but the costs of protecting the communications after death were high.[492]

The U.S. Supreme Court reversed,[493] finding that most cases discussing the attorney-client privilege have presumed or held outright that the privilege survives death and that the testamentary "exception" is grounded in furtherance of the client's interests, as opposed to the interest in criminal justice.[494] Therefore, under FRE 501, the Independent Counsel bears the burden of showing that "reason and experience"

485. *See id.* at 194.

486. *See id.*

487. *See id.* at 194–95.

488. 524 U.S. 399 (1998).

489. *See id.* at 401–02.

490. *See id.* at 402.

491. *See id.*

492. *See id.*

493. *See id.* at 411. Justice O'Connor dissented, joined by Justices Scalia and Thomas.

494. *See id.* at 405.

require overruling the common law rule that the attorney-client privilege survives death.[495] The Independent Counsel argued that the testamentary exception reflects a policy judgment that settling estates outweighs any posthumous interest in confidentiality and that the interest in determining whether a crime has been committed should trump client confidentiality.[496] The Court disagreed, stating that this interpretation did not square with the lower courts' acceptance of testamentary disclosure.[497]

On policy grounds, the *Swidler & Berlin* Court recognized that posthumous application of the privilege encourages full and frank disclosure and that clients' fear that disclosure after death could hurt family members or create civil liability may diminish candor.[498] The Court stated that any loss of evidence caused by the privilege is justified in part because without the privilege, the client may decide not to disclose the evidence in the first place.[499] The Court disagreed that limiting the proposed exception to criminal cases would have a minimal impact on client candor.[500] The Court noted that clients do not always know whether or not their confidential disclosures may become relevant to a civil or criminal matter and determined that this uncertainty could inhibit disclosure.[501] This fear of uncertainty is why the Court has consistently rejected a "balancing test" for the privilege.[502] Furthermore, the Court rejected the argument that recognizing one more exception would not result in harm because such recognition would erode the privilege without reference to common law principles or "reason and experience."[503]

---

495. *See id.* at 405–06.

496. *See id.* at 406.

497. *See id.*

498. *See id.* at 407.

499. *See id.* at 408.

500. *See id.* at 408–09.

501. *See id.* at 409.

502. *Id.* (citing Upjohn Co. v. United States, 449 U.S. 383 (1981)).

503. *See id.* at 410. In part because of the unavailability of empirical evidence on the effect of posthumous termination of the privilege, the petitioner failed to make a showing sufficient under FRE 501 to overturn the common law. *See id.* at 411. The dissent argued that the cost of recognizing an "absolute" posthumous privilege is "inordinately high" and advocated a balancing test. *Id.* at 413–14. One example the dissent gave when the privilege could be outweighed by fairness concerns was when a criminal defendant seeks disclosure of the deceased client's confession of the crime for which he is being tried. *See id.* at 416. Again, the Court features traditional justification principles such as encouraging candor through certain application of the privilege and recognizes that very little evidence may be lost through the privilege, because without it, the communication may not

The U.S. Supreme Court's jurisprudence indicates that recognizing or modifying a privilege should not be done without the proponent proving that the privilege is justified first by "reason and experience" in light of any historical bases and contemporary societal norms, and second by a public good transcending the general rule that litigants have "a right to every man's evidence."[504] The Court is more persuaded by traditional justifications and disinclined to accept the privacy rationale's "individual" justification per se, at least when it is inconsistent with the nature and facts or circumstances of the relationship. Thus, in the context of the spousal privilege—a privilege generally considered grounded in privacy concerns—the Court was willing to subvert the individuals' privacy interests to the societal interests of truth finding in criminal cases, at least where the marital relationship is in "disrepair" and the privilege has no parameters to prevent abuse.[505] So, while a privilege proponent should make policy arguments based on the privacy rationale where appropriate, it is probably more effective to justify the privilege under the traditional, utilitarian rationale and to focus on the public benefits the privilege would serve.[506] Also, where Congress has declined

---

have been made. *See* Great Am. Surplus Lines Ins. Co. v. St. Paul Fire & Marine Ins. Co., 540 So. 2d 1357, 1358 (Ala. 1989) (noting purpose of attorney-client privilege is to encourage candor); State v. Pavin, 494 A.2d 834 (N.J. Super. Ct. App. Div. 1985) (noting same and holding no public policy interest in granting privileged status to purported lie).

504. *See Swidler & Berlin*, 524 U.S. 399; Jaffee v. Redmond, 518 U.S. 1 (1996); Trammel v. United States, 445 U.S. 40 (1980).

505. *See Trammel*, 445 U.S. at 52.

506. Although corporations do not enjoy a right to privacy, privacy arguments can still persuasively buttress other privilege arguments. *See* Bush, *supra* note 277, at 599. Although this Article argues for a privilege for *employers*, the employees/test-takers' right to privacy in protecting their psychological processes also supports an unconscious-bias testing privilege. The U.S. Supreme Court has recognized that individuals' privacy interests may include controlling the dissemination of private information about themselves, and the privacy rationale supports controlling dissemination of information exposed by unconscious-bias testing. *See, e.g.*, Whalen v. Roe, 429 U.S. 589 (1977); *Developments in the Law—Privileged Communications*, *supra* note 327, 1543–48 & n.110. For a discussion on testing as an invasion of employees' privacy interests, see *infra* Part VI.D.1. Note also that psychology testing done as part of the individual's therapy would be covered by the psychotherapist-patient privilege, at least when done with the expectation of confidentiality for personal testing. *See, e.g.*, *In re* Doe, 649 P.2d 510 (N.M. Ct. App. 1982) (finding that for purposes of privilege, "communication" is defined broadly and includes information gained by personal examination and observation, verbal communications, exhibition of body parts, and inferences and conclusions drawn therefrom). However, psychology tests performed at the behest of the employer for EEO policy purposes probably would not be covered by the psychotherapist-patient privilege because there is no expectation of confidentiality. *See* State *ex rel.* Leas, 303 N.W.2d 414, 420 (Iowa 1981) (holding that psychologist's testimony regarding psychological testing was not privileged at hearing on parental rights because communications were not related to diagnosis or treatment and

an opportunity to enact the privilege, some deference will be paid to Congress' decision, particularly where no historical, constitutional, or statutory privilege basis exists and the privilege sought may precipitate a "wave" of similar privilege claims.[507]

C. *Privilege Law Principles and Court Analysis Support a Privilege for Unconscious-Bias Testing*

Courts should recognize a privilege for unconscious-bias test results to encourage companies interested in taking proactive measures to eradicate discrimination. The promotion of unconscious-bias testing not only has the immediate benefit of enabling employers to make prudent personnel decisions and initiate individual self-analysis and improved behavior, but it also has the more long-term benefit of contributing to the data on discrimination and providing insight on the direction of future research and changes in employment discrimination law. If unconscious-bias testing proves effective in predicting and preventing unlawful discrimination, such testing may become at the least a practical necessity for risk-management purposes. However, if discoverable, test results could expose employers to Title VII liability under Professor Oppenheimer's model of liability, as they may demonstrate constructive knowledge of the propensity for discrimination.[508]

Principles emerge from the cases discussing privilege law and the self-critical-analysis privilege that support the propriety of a privilege for unconscious-bias testing along the lines of the self-critical-analysis privilege. First, courts are more likely to recognize a privilege when they believe the desired communication will be chilled without a privilege. The U.S. Supreme Court emphasized the privilege holders' need to rely on a privilege to encourage candor in *Jaffee* and *Swidler & Berlin*.[509] Although early courts addressing a self-critical-analysis privilege, such as the *Banks* and *Dickerson* courts, manifested a belief that a privilege would impact the content, candor, and quality of government-mandated reports, later courts were not impressed by this argument.[510] After

---

were not expected to be confidential); *see also* Hager v. Bellingham Sch. Dist., 74 Wash. App. 49, 871 P.2d 1106 (1994).

507. *See* University of Pa. v. EEOC, 493 U.S. 182, 194–95 (1990).

508. *See supra* notes 111–23 and accompanying text.

509. *See supra* notes 468–507 and accompanying text.

510. *See supra* notes 416–26 and accompanying text.

O'Connor, courts began doubting that government reports would be chilled; because they were required, companies would prepare them regardless of a privilege, and in any case, other "deterrents" to candid self-criticism exist. Coupled with these arguments, some courts discussing the self-critical-analysis privilege, such as the *Martin* court, exhibited a basic disbelief that companies' EEO practices would change on account of changes in the law.[511]

Voluntary unconscious-bias testing presents a more compelling case for privilege than evaluative portions of government-mandated reports. First, little question exists that without an evidentiary privilege, the testing will not be done. Test results will likely indicate a high level of unconscious bias among most test takers.[512] The test results may serve as fodder for accusations of bigotry and intentional discrimination even though the tests do not measure the kind of intent required in a Title VII disparate treatment claim and despite the fact that a high level of unconscious bias does not necessarily result in discriminatory conduct.[513] Although an employer could rightly object that the test results are irrelevant to prove discrimination or that the test results are more prejudicial than probative,[514] the fear that this potentially damaging evidence could come out in trial makes unconscious-bias testing risky for employers without a privilege. The U.S. Supreme Court has repeatedly criticized a case-by-case privilege application method because uncertainty regarding a privilege's application could create fear on behalf of the privilege holder, chilling the disclosure sought to be fostered and making the privilege "little better than no privilege at all."[515]

Second, a privilege is more likely to be recognized where the proponent justifies the privilege on the utilitarian grounds that the privilege will greatly benefit society or that failing to recognize the privilege could greatly harm society. The U.S. Supreme Court in *Jaffee* emphasized the need for a privilege to advance a public good.[516] As the *Martin* court put it, "better [EEO] reports and evaluations are not an end

511. *See supra* notes 416–26 and accompanying text.

512. *See supra* note 254.

513. There is still some question about the relationship between unconscious bias and discriminatory action. *See supra* notes 48–57 and accompanying text.

514. *See* FRE 402, 403.

515. Jaffee v. Redmond, 518 U.S. 1, 18 (1996) (citing Upjohn v. United States, 449 U.S. 383, 393 (1981)); *see also* Swidler & Berlin v. United States, 524 U.S. 399, 409 (1998) (citing Upjohn v. United States, 449 U.S. 383, 393 (1981)).

516. *See Jaffee*, 518 U.S. at 9.

in themselves, but the means to an end."[517] So even assuming recognition of a privilege encourages unconscious-bias testing, will such testing benefit society?

Unconscious-bias testing's potential benefit to society in terms of better race relations is unprecedented and may prove to be unparalleled.[518] We are taught as children to be fair. Yet without an understanding of what part each of us plays in perpetuating a racially divided society—with the attendant injustice, danger, and unhappiness—we cannot hope to achieve fairness or peace. Unconscious-bias testing may be a first step in breaking down the ubiquitous ignorance that has permeated and perpetuated unhealthy interracial interaction for decades.

Furthermore, one cannot overstate the danger in not taking whatever steps are possible to create a more harmonious community. Not only are hate crimes and riots endemic in our recent history, but they are probably only the tip of an iceberg: we cannot continue to oppress growing segments of our society without increased violence and social unrest. Oppressed out-group members are dying unnecessarily as a direct result of socioeconomic hardship and discrimination in health care.[519] Many minorities cannot even secure employment, with health coverage or not, due to unconscious and unintended assumptions about their worth as employees.[520] Discriminatory employment practices deprive minorities of income, employment benefits, money for education, and hope for a better life. The *Bredice* court was concerned with creating better health care among elite hospitals by recognizing a privilege for doctors to criticize each other's practices without fear of disclosure.[521] *Bredice*'s concern about better health care among prestigious hospitals pales by comparison to the danger our society imposes on minorities, and ultimately society as a whole, by allowing discrimination to continue. If a means for uncovering and dispelling discriminatory attitudes or at least mitigating their effects exists, it is hard to imagine a more compelling reason to recognize a privilege to encourage use of that means. There is arguably

517. Martin v. Potomac Elec. Power Co., 58 Fair Empl. Prac. Cas. (BNA) 355, 357 (D.D.C. May 25, 1990).

518. Some empirical proof of the societal benefit unconscious-bias testing can achieve is currently available, and the data is growing. *See supra* Part III.A.

519. *See supra* notes 156–58 and accompanying text.

520. *See supra* notes 156–58 and accompanying text.

521. *See* Bredice v. Doctors Hosp., Inc., 50 F.R.D. 249, 250 (D.D.C. 1970), *aff'd*, 479 F.2d 920 (D.C. Cir. 1973).

no greater benefit to society than creating more fair opportunities and more justice for all of its members.

Third, when little or no evidence is lost on account of the privilege, the privilege is more likely to be recognized. As the self-critical-analysis privilege evolved, courts expressed increasing concern over Title VII plaintiffs' need for probative evidence, particularly considering that Title VII requires proof of intent.[522] Thus, the loss of potentially the best proof of intent was found to significantly outweigh any loss in the quality of mandatory government reports.[523]

Little or no evidence would be lost by recognizing a privilege for unconscious-bias testing. First, the test results probably will not be produced without a privilege protecting them from discovery. If the evidence is produced and then protected only as a result of the privilege, the result is a net evidentiary "wash"—a traditional, utilitarian argument in favor of privileges for communications that otherwise would not take place.[524] Second, proof of unconscious bias is not relevant evidence under Title VII. Unconscious bias does not prove intent to discriminate in disparate treatment cases. In disparate impact cases, unconscious bias does not prove facial neutrality or disparate impact. While unconscious-bias testing arguably provides evidence of the potential for stereotyping, which may be probative in mixed-motive or ADEA cases, the test results are probably legally insufficient proof because unconscious bias does not necessarily cause the subject to apply the stereotype in any particular context. Third, the test results are probably much more prejudicial than probative and therefore would be inadmissible regardless of a privilege. In any event, the evidence lost would be outweighed by the expected societal benefit warranting protection of the test results.[525]

---

522. *See, e.g.*, Hardy v. New York News, Inc., 114 F.R.D. 633, 641–42 (S.D.N.Y. 1987).

523. *See* Witten v. A.H. Smith & Co., 100 F.R.D. 446, 454 (D. Md. 1984).

524. *See* paragraph accompanying note 440.

525. "Long term accessibility to vital information must not be sacrificed on the altar of immediate discovery needs." Note, *supra* note 276, at 1088. Also, to the extent unconscious-bias testing is conducted in response to discrimination complaints, it may be excluded from trial use under FRE 407 as a "subsequent remedial measure." Note, *supra* note 277, at 1341–42. However, FRE 407 does not prevent discovery of the documents, and the need for confidentiality of this type of information makes it more amenable to privilege analysis, that is, a discovery bar is necessary to protect the interests at stake. *See id.* at 1346–47, 1352–55. Also, evidence excluded under FRE 407 can be admissible to prove control or feasibility, making less than a privilege for unconscious-bias testing dangerous for employers and contrary to the public policy of encouraging testing. *See id.* at 1342 n.18.

Fourth, to the extent privacy arguments support recognizing a privilege in addition to utilitarian arguments, the privilege is more likely to be accepted.[526] Unconscious bias test takers have a privacy interest in their test results. The testing promotes self-evaluation and self-actualization in addition to personal, emotional, and spiritual growth—privacy interests specifically recognized by cases and scholars addressing the privacy rationale.[527] Exposing test results could humiliate an individual who holds an egalitarian self-image—the same individual likely to benefit the most personally from unconscious-bias testing. Probing people's unconscious minds makes them particularly vulnerable, and forced disclosure of our minds' contents beyond the limited scope necessary to achieve the benefits of the testing is an unreasonable intrusion.

Because the policy justifications for utilitarian-based privileges require predictability in the privilege's application, it is important to clarify standards for the privilege's application and to focus on how to assure privilege holders that their communications will not be disclosed and how to curb employer abuse of that privilege. As a practical matter, most companies will employ consultants to test their employees for unconscious bias. Although some testing is offered via the Internet,[528] it would be difficult to monitor employees' testing and to keep track of the test results without some organized structure. Considering that at least one company will start offering unconscious-bias testing to corporate clients,[529] employers will likely turn to psychology experts rather than attempt to handle specialized psychological testing in-house.

Principles governing the admissibility of scientific evidence may inspire a "test" for the applicability of the unconscious-bias privilege that could assure employers that the results of their tests are valid, meet the privilege's objectives, and are not subject to discovery. One way to assure the inviolability of the privilege may be to condition its application on the employer utilizing a method generally accepted as scientifically accurate by the scientific community similar to the former *Frye v. United States*[530] test for admissibility of expert witness testimony

526. For a discussion on invasion of privacy, see *infra* Part VI.D.1.

527. *See supra* note 445 and accompanying text.

528. *See, e.g., supra* note 246.

529. Social psychologists Anthony G. Greenwald and Mahzarin R. Banaji are setting up a company to offer this service. Telephone Interview with Anthony G. Greenwald, *supra* note 11.

530. 293 F. 1013 (D.C. Cir. 1923).

relating to scientific evidence.[531] Alternatively, as a condition precedent to recognizing the privilege, the employer could be required to make a showing under FRE 702 as interpreted by *Daubert v. Merrell Dow Pharmaceuticals*[532] before conducting unconscious-bias testing in order to have a trial judge make a preliminary assessment of the validity of the test methodology. Because companies run by social psychology experts would presumably employ accepted unconscious-bias tests—such as Greenwald and Banaji's IAT—such testing would satisfy either of these conditions. In time, certain methodologies would become accepted for legal purposes, and only unconscious-bias testing performed by experts using an accepted methodology would be protected.

The privilege must have parameters to prevent abuse. Because the premise for an unconscious-bias-testing privilege is to encourage exploration and analysis of EEO practices and is based in part on the notion that employment discrimination litigation is ineffective, test results should not be discoverable unless they are abused by the employer in such a way that demonstrates an employer's intent to misuse the test results. The hope is that a privilege will encourage employers to use unconscious-bias test results to create more fair supervisory relationships, performance reviews, promotional decisions, and overall EEO policy. Also, fundamentally, the expectation is that as people begin to recognize the dissociation between their conscious beliefs and unconscious bias, this will spur processes to reverse automatic processing, making employment discrimination less likely. The results of these tests should be kept out of litigation to the extent possible to create employer trust in the privilege's application.

---

531. *See id.* at 1014. Under the *Frye* rule, scientific testimony was admissible only if the tests at issue had gained "general acceptance" within the scientific community. Thus, if a party sought to introduce an expert, it was up to the court to pass judgment on the general acceptance of the expert's methodology, based on testimony or judicial notice, prior to allowing the expert to testify before the jury. The U.S. Supreme Court held in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), that the *Frye* test was superseded by the Federal Rules of Evidence. Under *Daubert*, general acceptance is not a necessary precondition to the admissibility of scientific evidence; rather, the trial judge acts as a "gatekeeper," making his or her own preliminary assessment of the scientific validity of the methodology and reasoning underlying the proffered scientific testimony. *See id.* at 592–93. One "pertinent consideration" the *Daubert* Court discussed is whether the scientific theory or technique has been subjected to peer review and publication, although "[t]he inquiry envisioned by Rule 702 is . . . a flexible one." *Id.* at 594. In the case of unconscious-bias testing, judges would probably initially need to hear from experts regarding the validity of the testing, although in time certain methodologies would be recognized as valid.

532. 509 U.S. 579 (1993).

Employers should be barred from introducing favorable unconscious-bias testing as well. One of the reasons courts disallowed a self-critical-analysis privilege is because employers had been using EEO self-assessments to shield themselves from liability, so it was unfair to prevent plaintiffs from using them to prove liability. Not allowing either party to use unconscious-bias test results will dispel this fairness argument. At the same time, not allowing employers to use the results for their own benefit in litigation will screen out employers who exploit unconscious-bias testing for motives ulterior to creating a more egalitarian work environment and will prevent attempts to misrepresent test results.

Review of employer's response to the test results could also curb abuse. For example, availability of the privilege may be conditioned on the employer actually removing personnel decisionmaking authority from supervisors who demonstrate unconscious bias against their subordinates. This condition must be easy to satisfy or it could vitiate the privilege's promise of nondisclosure. The employer would not need to prove that fewer complaints of discrimination occurred after the personnel changes; instead, the employer must demonstrate only that it took steps consistent with the test results to minimize the potential impact of its supervisors' unconscious bias. Thus, if a supervisor showed a strong unconscious bias against blacks, but no bias against Hispanics or Asians, she could continue to review the work of the Hispanics and Asians, but her black subordinates should be assigned to a different supervisor. The employer would have to keep records of its test results and subsequent personnel changes in order to present to a court for *in camera* review in case the privilege is challenged. The expert psychologists conducting the tests should also keep records of the testing results, in the event the employer fails to produce test results upon challenge.

It is impossible to overemphasize that personnel changes consistent with the test results, rather than an increase or decrease in discrimination complaints, is the key to this condition. The goal of the privilege is to encourage employers to dismantle discrimination by attacking it from where it begins, to expose the depth of bias to employers so they can take action against it, and to inform and motivate individuals to "break the habit" of discrimination. While unconscious-bias testing promises to reduce discrimination and complaints of discrimination, few employers would dare conduct the tests if doing so risked inviting liability. Therefore, whatever conditions are required for the privilege's

application, they must be clear, easy to understand, and within the employer's control.

By the same token, however, employers who conduct unconscious-bias testing for impermissible motives may expose themselves to liability if they do not take antidiscrimination actions consistent with the test results. An extreme and unlikely example illustrates this point. Suppose a company owned by racists tests for unconscious bias in order to identify employees who appear unbiased for public relations purposes, but who still have underlying racist political beliefs. Then, the company promotes these most-biased individuals or gives them supervisory authority over group members against whom they demonstrated bias. This practice would not satisfy the second condition that test results and subsequent personnel changes be consistent with antidiscrimination efforts, so the test results would not be privileged and could be used to show intent to discriminate by the employers.[533]

The result is a privilege that appears very reliable and yet can be fine-tuned by judges to deny nondisclosure to employers who use unconscious-bias testing for impermissible purposes. With overcrowded dockets and mandatory alternative dispute resolution, some may question whether we have the resources to support the judicial oversight of an unconscious-bias testing privilege and its conditions. But with Title VII's effectiveness in serious question, AA plans on the decline, and hate crime and other societal ills on the rise, perhaps the better question is not "Can we afford to do this," but rather, "Can we afford not to?" If the societal benefits are even close to what the research anticipates, the immediate costs of encouraging unconscious-bias testing will be substantially outweighed, and as a result courts should recognize a privilege for unconscious-bias testing by employers.

D. Response to Expected Objections to Unconscious-Bias Testing

Objections to unconscious-bias testing may include the assertions that it invades individual privacy or deprives freedom of belief protected by the First and Fourteenth Amendments. The federal constitutional rights to privacy and freedom of thought are implicated only where the employer is a state actor[534] or where federal or state law requires the

533. Again, unconscious bias does not prove intent to discriminate. However, in the example given, the employer's actions create an inference of intent to discriminate per se.

534. *See* The Civil Rights Cases, 109 U.S. 3 (1883).

testing. Therefore, employees would have no federal constitutional grounds to challenge voluntary testing conducted by private employers. However, some state constitutions extend protections for privacy and freedom of thought to private, as well as government, actors.[535] Furthermore, private employees may bring state tort actions to vindicate the same privacy concerns that the federal Due Process Clause protects. More fundamentally, the policy concerns expressed in constitutional analysis of privacy and free speech are equally important in the private sector and must be addressed.

1. The Right of Privacy

The Fourteenth Amendment protects citizens from government deprivations of "liberty" without due process.[536] Privacy is one aspect of liberty secured by the Due Process Clause.[537] The government cannot infringe upon fundamental rights such as privacy unless strict scrutiny is met—that is, the government must have a compelling interest in the activity or prohibition that infringes on the right and the government's method must be narrowly tailored to meet the compelling interest.[538]

In 1965, the U.S. Supreme Court recognized in *Griswold v. Connecticut*[539] the constitutional right to privacy rather than grounding privacy in substantive due process.[540] Justice Douglas found the fundamental right to privacy implicit in the provisions of the Bill of

535. *See, e.g.*, Cal. Const. art. 1, § 1; Soroka v. Dayton Hudson Corp., 1 Cal. Rptr. 2d 77 (Cal. Ct. App. 1991).

536. U.S. Const. amend. XIV.

537. *See* Griswold v. Connecticut, 381 U.S. 479, 499–502 (1965) (Harlan, J., concurring); *see generally* Moore v. City of E. Cleveland, 431 U.S. 494 (1977); Roe v. Wade, 410 U.S. 113 (1973); Whalen v. Roe, 429 U.S. 589 (1971).

538. *See generally Moore*, 431 U.S. 494; *Roe*, 410 U.S. 113; *Whalen*, 429 U.S. 589.

539. 381 U.S. 479 (1965).

540. *See id.* at 482 (concluding that Connecticut law prohibiting use and distribution of contraceptives violated right to privacy by prohibiting married people from using contraceptives). There were several other opinions in *Griswold*. Justice Goldberg, joined by Chief Justice Warren and Justice Brennan, concurred, but emphasized that the Ninth Amendment was authority for the Court to protect nontextual rights such as privacy. *See id.* at 487. Justice Harlan's concurrence prefigured the modern view that the right to privacy protected by the liberty aspect of the Due Process Clause of the Fourteenth Amendment. *See id.* at 499–502. Justice White also concurred, and stated that the law would be unconstitutional even under a rational basis standard. *See id.* at 505. Justices Black and Stewart dissented, finding no right to privacy protected by the Constitution. *See id.* at 508.

Rights,[541] declaring that the "specific guarantees in the Bill of Rights have penumbras, formed by emanations from those guarantees that help give them life and substance [and that various] guarantees create zones of privacy."[542] Later cases did not follow Douglas's penumbra approach, and in any event, the application of the Bill of Rights to the states through the Due Process Clause of the Fourteenth Amendment makes even the penumbral approach to privacy essentially a due process analysis.[543] The right of privacy has also been grounded in the Equal Protection Clause, such as in *Eisenstadt v. Baird*,[544] in which the Court deemed unconstitutional a law that prohibited distributing contraceptives to unmarried persons but allowed distribution through physicians to married persons.[545] Justice Brennan wrote for the Court that "[i]f the right to privacy means anything, it is the right of the *individual*, married or single, to be free from unwarranted governmental intrusion into matters so fundamentally affecting a person as the decision whether to bear or beget a child."[546] Of course, the analysis of fundamental rights is essentially the same under the Due Process Clause and Equal Protection Clause, the main difference being how the constitutional arguments are phrased.[547] The Ninth Amendment also provides textual support for the Court's protection of nontextual rights, such as privacy.[548] However, the Court has almost never used the Ninth Amendment as support for the right to privacy per se.[549]

Fundamental privacy rights other than reproductive rights have also been recognized under the Constitution pursuant to the due process and Equal Protection Clauses and include the right to marry, medical

---

541. *See id.* at 484–85. Justice Douglas lived through the *Lochner* era and probably avoided a substantive due process analysis based on concerns over putting too stringent a limit on states' ability to exercise their police power to control societal ills and inhumane practices. *See* Lochner v. New York, 198 U.S. 45 (1905) (invalidating New York law making it criminal for baker to allow his employees to work more than 60 hours per week as unreasonable and unnecessary interference with individual's right of personal liberty). *Lochner*, of course, has been explicitly rejected by the Court on several occasions. *See, e.g.*, *Whalen*, 429 U.S. at 596 & n.18.

542. *Griswold*, 381 U.S. at 484–85 (citations omitted).

543. Erwin Chemerinsky, *Constitutional Law Principles and Policies* § 10.3.2, at 658–61 (1997).

544. 405 U.S. 438 (1972).

545. *See id.* at 453.

546. *Id.*

547. *See* Chemerinsky, *supra* note 543, § 10.1.1, at 639.

548. U.S. Const. amend. IX.

549. *See* Griswold v. Connecticut, 381 U.S. 479, 486–98 (1965) (Goldberg, J., concurring).

treatment rights, and the right to travel.[550] After *Griswold*, the Supreme Court has grounded privacy rights in the liberty aspect of the Due Process Clause of the Fourteenth Amendment rather than in the Equal Protection Clause or the "penumbra" of the Bill of Rights.[551] Privacy has been said to encompass two categories: an individual's interest in avoiding disclosure of personal matters and the interest in making certain kinds of important decisions.[552]

Employees involuntarily subjected to unconscious-bias testing could claim the testing invades their right to privacy. They could argue that they have a fundamental right to keep secret their innermost unconscious beliefs about other people—a privacy interest arguably more sensitive even than one's contraceptive privacy, for at least contraceptive choices are conscious and voluntary unlike one's cognitive functioning. Intuitively, the right to control access to unconscious thoughts seems more fundamental than conscious choices such as traveling, medical treatment, or even marriage.

The U.S. Supreme Court has indicated that a right to control personal, private information may exist. In 1977, in *Whalen v. Roe*,[553] the Court considered whether the right to privacy includes the right to control access to one's medical treatment history.[554] New York State maintained a centralized computer file that identified patients with prescriptions for certain addictive drugs and the prescribing doctors.[555] The computer files were a response to concerns that these addictive drugs were being diverted into illegal channels and abused. The patients claimed both types of privacy interests in their medical treatment and expressed concern that misuse of the computerized data would stigmatize them as drug addicts.[556]

The *Whalen* Court rejected the argument that these public records infringed upon the patients' right of privacy.[557] The Court found the state

550. *See* Chemerinsky, *supra* note 543, ch. 10; *see also* Carey v. Population Servs. Int'l, 431 U.S. 678 (1977); Whalen v. Roe, 429 U.S. 589 (1971).

551. *See, e.g.*, *Whalen*, 429 U.S. at 603–04; *see also* Moore v. City of East Cleveland, 431 U.S. 494 (1977); Roe v. Wade, 410 U.S. 113 (1973).

552. *See Whalen*, 429 U.S. at 598–600.

553. 429 U.S. 589 (1977).

554. *See id.* at 591.

555. *See id.*

556. *See id.* at 595.

557. *See id.* at 598. The doctors also claimed that the law interfered with their freedom to practice medicine, which the Court summarily rejected. *See id.* at 604.

to have an important interest in monitoring abuse and illegal distribution of prescription drugs.[558] Medical information is routinely given to hospital personnel, insurance companies, and public health agencies, and disclosure to the state is not an impermissible invasion of privacy.[559] The Court discussed the security system set up to prevent inappropriate access to the data, the destruction of the data after five years, and the fact that employees who failed to maintain proper security over the information were subject to civil and criminal penalties.[560] Thus, the Court found the state database to be a reasonable and constitutional exercise of the state's police power.[561] However, the Court also acknowledged the concept of a constitutional privacy right to control information that is "personal in character and potentially embarrassing or harmful if disclosed."[562] As a result, a government employer that conditions employment on an unconscious bias test could arguably violate test-takers' constitutional right to privacy.[563]

At least one federal district court has upheld a government employer's psychological testing of employees against privacy challenges. In *McKenna v. Fargo*,[564] the court considered the constitutionality of standardized psychology tests that included questions on religious, political, and familial relationships as a condition of employment as a firefighter.[565] The *McKenna* court relied on the dicta of *Whalen v. Roe* to conclude that the right of privacy extends to employer-mandated psychological testing because of the disclosure of highly personal information:

> In this case, the degree and character of the disclosure is far greater and more intrusive [than in *Whalen v. Roe*]. The evaluation looks deeply into an applicant's personality, much as a clinical

558. *See id.* at 508.

559. *See id.* at 602.

560. *See id.* at 593–95.

561. *See id.* at 598–600.

562. *Id.* at 605.

563. In accordance with the "unconstitutional conditions" doctrine, termination by a government employer for failing to submit to testing could raise issues regarding deprivation of a property right without due process, if the testing is a privacy violation, and continued employment is conditioned on submitting to it. *See* Chemerinsky, *supra* note 543, § 7.3.1–7.3.2; *see also* Bishop v. Wood, 426 U.S. 341 (1976); Arnett v. Kennedy, 416 U.S. 134 (1974).

564. 451 F. Supp. 1355 (D.N.J. 1978), *aff'd*, 601 F.2d 575 (3d Cir. 1979).

565. *See id.* at 1357. The tests included the Minnesota Multiphasic Personality Inventory (MMPI) and a number of other standardized tests. *See id.* at 1359–61.

> psychologist would if requested to do so by an applicant . . . [This testing] involves a loss of the power individuals treasure to reveal or conceal their personality or their emotions as they see fit, from intimacy to solitude. Involuntary disclosure of such a unique kind . . . distinguishes this case from *Whalen* . . . . That privacy interest is sufficiently burdened . . . to call constitutional protection into play.[566]

However, in applying strict scrutiny, the court found that the psychological testing was not an impermissible violation of the right to privacy and upheld the government's use of the tests.[567] The court found that the tests were narrowly tailored to meet the state's interest in an effective fire department because they identified "applicants whose emotional make-up makes them high risk candidates for the job of firefighting."[568] The court stated that the revelations of the psychological evaluation may save the lives of the unqualified applicants or others.[569]

In the same vein, to the extent unconscious-bias testing invades an employee's right to privacy, the testing is justified by the employer's compelling interest in preventing employment discrimination.[570] Social psychology research indicates that to redress employment discrimination effectively, one must redress unconscious bias. Unconscious-bias testing is narrowly tailored to achieve the compelling purpose of eradicating invidious discrimination. Unconscious-bias testing tests only for bias, not for general personality traits, such as religious beliefs, sexual preferences, or other unrelated aspects of one's personhood, and thus unconscious-bias testing does not unnecessarily invade the test-takers' privacy. Employers have a legal obligation to prevent discrimination similar to the fire department's obligation to provide reliable, emotionally stable firefighters in fire emergencies: failure of either has a serious impact on the well-being of innocent citizens.

---

566. *Id.* at 1380–81.

567. *See id.* at 1382.

568. *Id.* at 1381.

569. *See id.* The court also found that there were not less intrusive alternatives available but directed the city to limit the length of time the test records would be kept. *See id.* at 1382.

570. *See, e.g.*, United States v. Paradise, 480 U.S. 149, 167 (1990) (holding government "unquestionably has a compelling interest in remedying past and present [race] discrimination by a state actor"); Board of Dir. of Rotary Int'l v. Rotary Club, 481 U.S. 537, 549 (1987) (finding any infringement on Rotary members' right of expressive association on account of California's Unruh Civil Rights Act is justified by state's compelling interest in eliminating discrimination against women and minorities).

Other professionally developed psychology tests, such as the Minnesota Multiphasic Personality Inventory (MMPI) and the California Psychological Inventory (CPI) are often used in the employment context without regard to employees' or applicants' privacy rights.[571] Personality tests, like unconscious-bias testing, expose traits the subject may be unaware of and may even adamantly deny.[572]

In California, at least one case has indicated that standardized psychological testing may be an unconstitutional invasion of privacy under the California Constitution. Under the California Constitution, the right to privacy is liberally construed and applies to private actors as well as state actors.[573] In California, both state and private employees might assert that unconscious-bias testing is an unconstitutional intrusion into private thoughts that is not justified by an employer's desire to identify employees likely to discriminate. A review of a California Court of Appeal's analysis of personality testing challenged on privacy grounds, however, militates against the success of a privacy challenge to unconscious-bias testing.

In *Soroka v. Dayton Hudson Corp.*,[574] a California Court of Appeal reversed a lower court's denial of job applicants' motion for a preliminary injunction prohibiting the use of MMPI and CPI testing pending the outcome of the lawsuit.[575] The court found the injunction appropriate because the plaintiffs had shown a reasonable probability of

571. *See, e.g.*, Cort v. Bristol-Myers Co, 431 N.E.2d 908 (Mass. 1982) (determining employer not liable for terminating employees for their refusal to answer questions on employer's questionnaire because most questions were relevant to employer's job qualifications and presented no invasion of privacy, affirming trial court's directed verdict for employer on privacy claim); *see also* Bennett v. County of Suffolk, 30 F. Supp. 2d 353 (E.D.N.Y. 1998) (denying summary judgment to challenge to MMPI and CPI on Title VII and religious freedom grounds); Shirsat v. Mutual Pharm. Co., 169 F.R.D. 68 (E.D. Pa. 1996) (holding employer was entitled to conduct MMPI on plaintiff-former employee in action in which plaintiff's mental state was at issue); Colbert v. H-K Corp., 4 Fair Empl. Prac. Cas. (BNA) 529 (N.D. Ga. July 2, 1971) (upholding use of intelligence and personality tests, as they were professionally developed, related to job performance, recommended by psychologist, and not intended to discriminate); Scott v. State, No. 96A-06-001-RRC, 1996 WL 769222 (Del. Super. Ct. Dec. 6, 1996) (ordering MMPI by employer's expert witness).

572. *See* Cuddy v. Wal-Mart Super Ctr., Inc., 993 F. Supp 962 (W.D. Va. 1998) (granting defendant's summary judgment on ADEA claim in situation where although plaintiff denied such traits as being "bitter" or believing the "world owed him something," those traits were confirmed by standardized personality inventory (Orion) test).

573. *See* Cal. Const. art. 1, § 1; Soroka v. Dayton Hudson Corp., 1 Cal. Rptr. 2d 77, 83 (Cal. Ct. App. 1991).

574. 1 Cal. Rptr. 2d 77, 83 (Cal. Ct. App. 1991).

575. *See id.* at 79.

prevailing on the merits of their claim that the tests infringed on their constitutional right to privacy.[576]

To pass muster under the California Constitution, any violation of the right to privacy of job applicants must be justified by a compelling interest and the employer must demonstrate a clear, direct nexus between the nature of the employment duties and the content of the test that constitutes a privacy infringement—essentially a strict scrutiny standard.[577] The test questions at issue related to the applicants' religious beliefs and sexual orientation; the security officer positions involved approaching, questioning, and possibly detaining shoplifting suspects at Target stores. The employer's claim that the test bore a relationship to emotional fitness and that it had seen an overall improvement in the security officers' performance since implementing the tests did not constitute a compelling interest or satisfy the nexus requirement.[578] Although a final determination was never reached,[579] the opinion clarifies that a compelling interest and clear nexus between psychological tests and job responsibilities are necessary for an employer to infringe lawfully upon employees' right to privacy in California.

Even under California's more liberal privacy standards, unconscious-bias testing is constitutionally permissible. First, employers have a compelling interest—indeed, a legal obligation—to take steps to avoid workplace discrimination, particularly among employees who make employment decisions for other employees. The weight of empirical psychological studies shows that unconscious-bias testing advances that interest by pinpointing where discrimination is most likely to occur and allowing the employer to take steps to prevent discrimination. The need for unconscious-bias testing, and all of its potential benefits to society, outweighs any privacy interests involved.[580]

---

576. *See id.* at 86.

577. *See id.* at 85.

578. *See id.* at 86.

579. The parties stipulated to dismiss review. *See* Soroka v. Dayton Hudson Corp., 862 P.2d 148 (Cal. Ct. App. 1993).

580. Cases upholding the use of employer drug testing based on the employer's interest in protecting the safety of its employees and the public outweighing the right to privacy can be used as additional support for unconscious-bias testing. *See, e.g.*, Plane v. United States, 796 F. Supp. 1070 (W.D. Mich. 1992). While drug testing is distinguishable in that it reveals employees' conscious, deliberate use of drugs, the privacy concerns are similar, as are the employer's interests in protecting its employees from unnecessary and preventable harm.

Recognition of a privilege for unconscious-bias testing actually enhances the privacy interests of individuals tested. Under current law, employers are free to subject employees to psychological testing, and the test results are not privileged per se.

Finally, individual privacy concerns can be lessened substantially by redacting identifying information from test results. Under a qualified privilege, a court deciding the privilege's applicability can compare test results and subsequent personnel changes by referring to positions and job responsibilities on the one hand and test results on the other, rather than by referring to names or other identifying information. To the extent a plaintiff seeks to prove discrimination through a negligence theory of liability or by showing the employer intentionally kept highly biased persons in decisionmaking positions, the plaintiff can demonstrate either without specifically naming the particular employees tested. Thus, neither an *in camera* review to determine the privilege's applicability nor evidence presented to a factfinder need identify the particular individuals tested or their test results.

Privacy analysis turns on how important it is to engage all available methods to eradicate discrimination. It simply seems wrong to protect people's privacy interests in harboring negative stereotypes about others—stereotypes they are unaware of and often would like to change—to the detriment of innocent people routinely oppressed on account of these stereotypes. Justice requires that the balance tip in favor of allowing unconscious-bias testing, considering that the limited infringement on privacy interests is necessary to combat discrimination.

2. *Freedom of Speech and Beliefs*

Unconscious-bias testing does not violate the First Amendment. Employees may try to resist unconscious-bias testing based on First Amendment concerns of freedom of thought and conscience, arguing that unconscious-bias testing interferes with their right to think and express stereotypical thoughts. Employees could argue that unconscious-bias testing is politically correct "thought control" and the thin edge of the wedge of governmental oppression of individual autonomy and ideology. A short review of the First Amendment's purpose and analysis, however, reveals that unconscious-bias testing actually enhances the underlying social benefits the First Amendment seeks to encourage and protect.

Although there is no single, universally accepted purpose of the First Amendment, there are a number of traditional arguments for why freedom of speech and beliefs should be regarded as fundamental rights.

These include self-governance (democracy works better if candidates are encouraged to speak freely and voters are better informed);[581] discovery of truth (truth emerges from open debate in the "marketplace of ideas");[582] advancement of autonomy (human spirit "demands self-expression");[583] and promotion of tolerance (free speech helps segments of society to understand one another's ideas).[584]

Fundamentally, the First Amendment seeks to protect people from censorship or punishment for expressing their views—particularly criticism of the government. England's use of prior restraints on publication and law of sedition motivated the Framers to write the First Amendment. The First Amendment's *raison d'etre* is to protect the expression and dissemination off all viewpoints and data, to prevent oppression, and to allow people to speak freely, gather information, and learn about themselves and the world. Justice Brandeis eloquently summed up the First Amendment's ideals in *Whitney v. California*:[585]

> Those who won our independence believed that the final end of the State was to make men free to develop their faculties, and that in its government the deliberative forces should prevail over the arbitrary . . . . They believed that freedom to think as you will and to speak as you think are means indispensable to the discovery and spread of political truth; that without free speech and assembly discussion would be futile; that with them, discussion affords ordinarily adequate protection against the dissemination of noxious doctrine; that the greatest menace to freedom is an inert people; . . . it is hazardous to discourage thought, hope, and imagination.[586]

Unconscious-bias testing does not violate the First Amendment because the tests do not suppress speech or beliefs, but rather they identify subconscious stereotypes to prevent acts of discrimination. Obviously, a state actor could not punish an employee for unconscious bias any more than it could for explicit racist speech or beliefs.[587] The

581. *See* Chemerinsky, *supra* note 543, § 11.1.2.

582. *See id.*

583. *See* Procunier v. Martinez, 416 U.S. 396, 427 (1974) (Marshall, J., concurring).

584. *See* Chemerinsky, *supra* note 543, § 11.1.2 & n.39.

585. 274 U.S. 357 (1927).

586. *Id.* at 375 (Brandeis, J., concurring).

587. *See, e.g.*, National Socialist Party of Am. v. Village of Skokie, 432 U.S. 43 (1977); Chemerinsky, *supra* note 543, § 11.3.3.4.

testing does not dictate what beliefs employees may hold or express.[588] The purpose of unconscious-bias testing is to protect others from the harm that may result from those beliefs. The individuals tested are free to maintain their bias, but they are not free to manifest that bias in illegal acts of discrimination.

The First Amendment right of free speech and beliefs assumes that the individual knows the content of the speech or belief protected. The overall objective is to allow expression of information or opinions individuals wish to advance or advocate. With unconscious bias, the test-taker is necessarily uninformed about her bias, and therefore, there is no "speech" or "belief" to express, disseminate, or protect. Unconscious-bias testing simply does not implicate the First Amendment because the test-takers are not even aware of their unconscious bias until after the testing and are free to maintain whatever beliefs they choose thereafter.[589] Thus, this argument is consistent with federal court analysis of free speech and personality testing.

The court in *McKenna v. Fargo* held that testing for emotional fitness does not implicate freedom of beliefs.[590] The employer did not see the raw testing data, but instead relied on summaries made to predict whether the subject could withstand the psychological pressures involved

588. The U.S. Constitution does not preclude private employers from making employment decisions based on employees' speech or beliefs. However, Title VII and some state statutes protect against adverse employment decisions by private employers based on certain personal beliefs, such as religious beliefs. *See, e.g.*, Cal. Gov. Code § 12940(a).

589. Note also that, to the extent social psychologists are correct in believing that our unconscious bias results largely from cultural and media messages that we do not "choose," but are inundated with from an early age, the entire "marketplace of ideas" concept underlying the First Amendment is called into question: the "free trade of ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market" is hindered when ideas creating unconscious bias are planted in our minds without our acceptance or consent. Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting). That is, according to some social psychologists, we are not choosing our biases in an open market, but rather, have no choice in the matter, making First Amendment analysis inapposite. At least one court has opined that subliminal messages are not protected by the First Amendment because (1) they do not advance any theories supporting free speech (such as Holmes's marketplace-of-ideas concept, or individual self-fulfillment and self-realization); (2) an individual has a First Amendment right to be free from unwanted speech; and (3) the listener's right to privacy outweighs the speaker's right of free speech when subliminal messages are used. *See* Vance v. Judas Priest, Nos. 86-5844, 86-3939, 1990 WL 130920 (D. Nev. Aug. 24, 1990). Unconscious bias is similar to subliminal messages in that it affects us beyond our level of awareness, so we do not exercise conscious choice about it.

590. *See* McKenna v. Fargo, 451 F. Supp. 1355, 1377–78 (D.N.J. 1978), *aff'd*, 601 F.2d 575 (3d Cir. 1979).

in firefighting.[591] The employer thus did not attempt to deny employment based on protected beliefs, but rather attempted to identify personality traits incompatible with firefighters' duties. The court noted, however, that the potential for abusing psychological testing is clear, and it would be unconstitutional to reject a candidate based on his beliefs under the guise of rejecting the emotionally unfit.[592]

The employer is obligated to use the unconscious-bias test results to match employees with job responsibilities that reduce the potential for discrimination. The testing is not conducted to control the employees' beliefs, but instead to adjust personnel to decrease the likelihood of unconscious bias manifesting in unlawful and harmful acts toward innocent out-group members. The First Amendment protects racist views, not racist acts in employment; the purpose of unconscious-bias testing is to prevent "bias-inspired conduct"[593] and not to dampen beliefs or speech.

Employees may argue that compulsory unconscious-bias testing violates their right not to speak. The U.S. Supreme Court has recognized the right not to speak.[594] In *West Virginia State Board of Education v. Barnette*,[595] the Court declared unconstitutional a state law requiring children to salute the flag because it compelled an affirmation of a belief.[596] More recently, in *McIntyre v. Ohio Elections Commission*,[597] the Court held that the First Amendment protects a person's right to

---

591. *See id.* at 1377.

592. *See id.* at 1377–78.

593. *See* Chemerinsky, *supra* note 543, § 11.3.3.4. The U.S. Supreme Court has held that enhanced penalties for hate crimes are constitutional, stating that the law

> singles out for enhancement *bias-inspired conduct* because this conduct is thought to inflict greater individual and societal harm .... [B]ias-motivated crimes are more likely to provoke retaliatory crimes, inflict distinct emotional harms on their victims, and incite community unrest. The State's desire to redress these perceived harms provides an adequate explanation for its penalty-enhancement provision over and above mere disagreement with offenders' beliefs or biases.

Wisconsin v. Mitchell, 508 U.S. 476, 487–88 (1993) (emphasis added and citations omitted).

594. *See* Wooley v. Maynard, 430 U.S. 705, 714 (1977) ("[T]he right of freedom of thought protected by the First Amendment ... includes both the right to speak freely and the right to refrain from speaking at all .... The right to speak and the right to refrain from speaking are complementary components of the broader concept of 'individual freedom of mind.'").

595. 319 U.S. 624 (1943).

596. *See id.* at 633, 642; *see also Wooley*, 430 U.S. 705 (holding it unconstitutional to punish person for blocking out portion of his license plate which contained New Hampshire state motto "Live Free or Die").

597. 514 U.S. 334 (1995).

speak anonymously.[598] In invalidating an Ohio law that required contact information on voter publications, the Court determined that the decision to remain anonymous, as with other publication content decisions, is protected by the First Amendment's freedom of speech provision.[599] The Court emphasized that authors may choose anonymity to preserve their privacy or because of fear that their expressions will cause retaliation or ostracism.[600] But whatever the motivation for choosing anonymity, any interest in disclosure as a condition for entry into the marketplace is clearly outweighed by the interest in ensuring the entry of anonymous works into the public domain.[601] This case supports the right not to speak because it protects the speaker's choice about the content of speech and allows the speaker to withhold speech at the speaker's discretion.

However, unconscious-bias testing can be distinguished from "right not to speak" cases because those cases assume that the would-be speaker knows the content of the speech and has decided not to advance it, thereby expressing an opinion or choice by refusing to speak. Unconscious-bias test-takers are not conscious of their cognitive content and are therefore unable to form an opinion about that content or to exercise the right to disagree with it or withhold it for content-based purposes by not speaking. Nothing in unconscious-bias testing or application of test results in the employment context prevents the test-taker from maintaining the bias or refusing to talk about it. The testing is done for the purpose of preventing discriminatory acts, not regulating speech.

Unconscious-bias testing not only does not impinge on an individual's right to believe and speak as she pleases, but actually enhances the values underlying the First Amendment by exposing information important to an individual's right to choose her beliefs in accordance with the true facts. The testing advances the discovery of truth in the "marketplace of ideas" by uncovering bias and by contributing critical information about the reality of racism and discrimination. It enhances autonomy because people cannot freely

598. *See id.* at 357.

599. *See id.* at 341–42. Note that the Court applied "exacting scrutiny," which requires that a statute be "narrowly tailored to serve an overriding state interest," to analyze the constitutionality of the Ohio law's burden on the "core political speech" involved in the case. *Id.* at 347. Previously, the Court applied "strict scrutiny" to test content-based restrictions on political speech. *See, e.g.*, Meyer v. Grant, 486 U.S. 414 (1988); Talley v. California, 362 U.S. 60 (1960).

600. *See McIntyre*, 514 U.S. at 341–42.

601. *See id.* at 342.

express themselves when they are ignorant about who they are and why they do what they do. Understanding oneself is a prerequisite for true autonomy and self-actualization. Perhaps most importantly, unconscious-bias testing promotes tolerance of others. White Americans who feel affirmative action has "gone too far" or that minorities take advantage of their status may experience an epiphany when confronted with statistics regarding unconscious bias.

CONCLUSION

Racial and gender statistics demonstrate that America is a land entrenched with injustice and stark socioeconomic stratification. Equal employment opportunity is critical to creating socioeconomic justice for all Americans.

Employment discrimination statutes were enacted to provide a remedy for victims of discrimination, as well as to deter discrimination. But discrimination laws do not, and cannot, redress what social psychologists believe is the most common form of discrimination—stereotyping resulting from implicit associations made beyond people's conscious control. New psychology research offers an alternative to litigation as a means to deter and redress discriminatory practices. Unconscious-bias testing provides hope for an objective, clear, and efficacious means of identifying, and working to solve, the problem of discrimination in employment.

The potential benefits of unconscious-bias testing are far-reaching and many. Merely recognizing that we harbor biases against others gives us the power of knowledge to take proactive steps to minimize the likely impact of such bias. Perhaps more importantly, discovering our own biases is a necessary first step toward reversing implicit associations and also provides data necessary to break the chain of "behavioral confirmation." As more people understand the serious social implications and magnitude of injustice resulting from unconscious bias, in part from confrontation with the raw statistics that unconscious-bias testing can achieve, our society may begin the process of racial enlightenment we need to work toward racial harmony and justice.

Use of unconscious-bias testing in employment settings is a critical step in addressing economic injustice. However, employers are unlikely to conduct the testing at this time because evidence law does not encourage it, but rather discourages it by failing to provide protection against liability based on test results. Thus, a qualified evidentiary privilege should be recognized so that unconscious-bias testing can be

used as an effective tool to make changes in employment practices to further equal employment opportunity.

The law should respond to recent breakthroughs in the psychological understanding of discrimination by encouraging unconscious-bias testing. The potentially metamorphic social benefits unconscious-bias testing can achieve underscore the need for a privilege to protect test results to encourage their use. Hopefully, in time, we will break away from unfair, unconscious stereotypes and create an America where prosperity and peace is equally attainable for all.