

7-1-2004

Implicit Racial Attitudes of Death Penalty Lawyers

Theodore Eisenberg

Cornell Law School, ted-eisenberg@lawschool.cornell.edu

Sheri Lynn Johnson

Cornell Law School, slj8@cornell.edu

Follow this and additional works at: <http://scholarship.law.cornell.edu/facpub>

 Part of the [Applied Statistics Commons](#), [Criminal Law Commons](#), and the [Legal Profession Commons](#)

Recommended Citation

Eisenberg, Theodore and Johnson, Sheri Lynn, "Implicit Racial Attitudes of Death Penalty Lawyers" (2004). *Cornell Law Faculty Publications*. Paper 353.

<http://scholarship.law.cornell.edu/facpub/353>

This Article is brought to you for free and open access by the Faculty Scholarship at Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in Cornell Law Faculty Publications by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact jmp8@cornell.edu.

IMPLICIT RACIAL ATTITUDES OF DEATH PENALTY LAWYERS

*Theodore Eisenberg**
*Sheri Lynn Johnson***

INTRODUCTION

Defense attorneys commonly suspect that the defendant's race plays a role in prosecutors' decisions to seek the death penalty, especially when the victim of the crime was white.¹ When the defendant is convicted of the crime and sentenced to death, it is equally common for such attorneys to question the racial attitudes of the jury. These suspicions are not merely partisan conjectures; ample historical,² statistical,³ and anecdotal⁴ evidence supports the inference that race matters in capital cases. Even the General Accounting Office of the United States concludes as much.⁵ Despite *McCleskey v. Kemp*,⁶ in which the United States Supreme Court concluded that strong, well-controlled statistical correlations with race do not demonstrate causa-

* Henry Allen Mark Professor of Law, Cornell Law School. The authors would like to thank Mary Mulhearn for her research assistance.

** Professor of Law, Cornell Law School.

1. A race-of-victim effect is widely supported by empirical studies. The existence of a broad race-of-defendant effect has been much more difficult to detect. David Baldus et al., *Racial Discrimination and the Death Penalty in the Post-Furman Era: An Empirical and Legal Overview, with Recent Findings from Philadelphia*, 83 CORNELL L. REV. 1638, 1658-59 n.61, 1662, app. B (1998) (collecting studies and reporting that race-of-defendant effect has not been found in over fifty empirical studies). For recent evidence of such an effect, see John H. Blume et al., *Explaining Death Row's Population and Racial Composition*, 1 J. EMPIRICAL LEGAL STUD. 165, 167, 200 (2004).

2. See generally *Furman v. Georgia*, 408 U.S. 238 (1972).

3. See e.g., DAVID BALDUS ET AL., EQUAL JUSTICE AND THE DEATH PENALTY: A LEGAL AND EMPIRICAL ANALYSIS (1990); Baldus et al., *supra* note 1; SAMUEL R. GROSS & ROBERT MAURO, DEATH AND DISCRIMINATION: RACIAL DISPARITIES IN CAPITAL SENTENCING (1989); John H. Blume et al., *Post-McCleskey Racial Discrimination Claims in Capital Cases*, 83 CORNELL L. REV. 1771, 1790 (1998); RICHARD C. DIETER, THE DEATH PENALTY IN BLACK AND WHITE: WHO LIVES, WHO DIES, WHO DECIDES 17 fig. 7 (1998) (summarizing studies).

4. For a particularly egregious example, see *Andrews v. Shulsen*, 485 U.S. 919, 920 (1988) (Marshall, J., dissenting) (A juror handed the bailiff a napkin with a drawing of a man on a gallows above the inscription "Hang the Niggers.").

5. U.S. GEN. ACCOUNTING OFFICE, DEATH PENALTY SENTENCING: RESEARCH INDICATES PATTERN OF RACIAL DISPARITIES 1-6 (1990).

6. 481 U.S. 279 (1987).

tion, half of all Americans believe that race does influence the administration of the death penalty.⁷

In investigating the influence of racial bias, commentators (ourselves included) have focused on prosecutors and jurors,⁸ generally neglecting the question of whether bias affects the representation defense counsel provides his or her client. In part, this may be due to the greater difficulty in uncovering evidence of bias in the ranks of defense counsel; most of the evidence of bias of other death penalty actors has been uncovered through the efforts of defense counsel, a class one might suspect would be less likely to scrutinize itself. It may also be faith in the adversary system that disinclines observers to suspect bias on the part of defense counsel: One would hope that those who represent capital defendants (or at least African-American capital defendants) would themselves be free of racialized thinking as they establish trust with their clients, direct both fact and mitigation investigations, select experts, choose witnesses to call, and decide what arguments to make. A final reason for this inattention to defense lawyers may be the apparent *sincerity* of the ideological commitment of most such lawyers to racial equality; not only do they proclaim that commitment,⁹ but the poor pay and low status of their work suggests some amount of sacrifice for those ideological commitments. But ideological commitment need not translate into racially unbiased evaluations, as a large accumulation of literature discussing social and cognitive psychology demonstrates.¹⁰ Indeed, the defense lawyer's commitment to the formal norm of equality is probably shared with most prosecutors and jurors. So the question should be asked, better late than not at all: What do we know about the capital defense lawyer's racial attitudes?

Nothing. Virtually nothing is known about the racial attitudes of lawyers in general, let alone defense lawyers or capital defense law-

7. See Tom Moranthau & Peter Annin, *Should McVeigh Die?*, NEWSWEEK, June 16, 1997, at 20 (citing a *Newsweek* poll conducted by Princeton Survey Research Associates polling seventy-five adults, eighteen and older, June 5-6, 1997 in which 49% of respondents agreed). Non-white respondents overwhelmingly agreed. *Id.*

8. See, e.g., BALDUS ET AL., *supra* note 3; Baldus et al., *supra* note 1; GROSS & MAURO, *supra* note 3; Blume et al., *supra* note 3; Jeffrey J. Pokorak, *Probing the Capital Prosecutor's Perspective: Race of the Discretionary Actors*, 83 CORNELL L. REV. 1811 (1998).

9. See, e.g., Bryan A. Stevenson & Ruth E. Friedman, *Deliberate Indifference: Judicial Tolerance of Racial Bias in Criminal Justice*, 51 WASH. & LEE L. REV. 509 (1994); Stephen B. Bright, *Race, Poverty, the Death Penalty and the Responsibility of the Legal Profession*, 1 SEATTLE J. SOC. JUST. 73 (2002).

10. See Sheri Lynn Johnson, *Racial Imagery in Criminal Cases*, 67 TUL. L. REV. 1739, 1760-66 (1993) (reviewing the literature).

yers specifically. The demographic characteristics,¹¹ compensation patterns,¹² career paths,¹³ and occasionally the daily activities¹⁴ of lawyers are studied, but researchers to date have expressed little interest in their attitudes, with the exception of attitudes concerning job satisfaction.¹⁵

In contrast, quite a lot is known about the racial attitudes of the general population. The prevalence of hostile, overt racism has been declining at least since the 1960s.¹⁶ Some researchers have observed that polls may overstate this trend, given the growing social unacceptability of racial hostility. Indeed, when experiments have tried to weed out social desirability effects, they do find greater levels of conscious stereotyping and hostility, though still clearly lesser levels than in the past.¹⁷ For the most part, however, old-fashioned, “Bull Connor-style” racism has not been replaced with colorblindness but with subtler manifestations of racial bias. Some social psychologists have labeled this newer racism “aversive racis[m],” documenting the prevalence of subjects who subscribe to a formal norm of equality, but desire to keep their distance from other racial groups, and often covertly disparage those groups.¹⁸ Cognitive psychologists have focused more on stereotypes, observing how thinking and judgment may be altered by stereotypes that the subject would not endorse, and often consciously rejects.

11. See, e.g., Deborah R. Hensler, *Studying Gender Bias in the Courts: Stories and Statistics*, 45 STAN. L. REV. 2187 (1993); *Report of the Working Commission to the Second Circuit Task Force on Gender, Racial and Ethnic Fairness in the Courts*, 1997 ANN. SURV. AM. L. 117.

12. See, e.g., Catherine Green Burnett et al., *In Pursuit of Independent, Qualified and Effective Counsel: The Past and Future of Indigent Criminal Defense in Texas*, 42 S. TEX. L. REV. 595 (2001); Robert R. Rigg, *The Constitution, Compensation, and Competence: A Case Study*, 27 AM. J. CRIM. L. 1 (1999).

13. See, e.g., Robert Granfield & Thomas Koenig, “*It’s Hard To Be a Human Being and Lawyer*”: *Young Attorneys and the Confrontation with Ethical Ambiguity in Legal Practice*, 105 W. VA. L. REV. 495 (2003); Fiona M. Kay & John Hagan, *Building Trust: Social Capital, Distributive Justice and Loyalty to the Firm*, 28 LAW & SOC. INQUIRY 483 (2003).

14. See, e.g., Austin Sarat & William L.F. Felstiner, *Law and Strategy in the Divorce Lawyer’s Office*, 20 LAW & SOC’Y REV. 93 (1986).

15. See, e.g., Granfield & Koenig, *supra* note 13; Kenneth G. Dau-Schmidt & Kaushik Mukhopadhyaya, *The Fruits of Our Labors: An Empirical Study of the Distribution of Income and Job Satisfaction Across the Legal Profession*, 49 J. LEGAL EDUC. 342 (1999).

16. See, e.g., HOWARD SCHUMAN ET AL., *RACIAL ATTITUDES IN AMERICA: TREND AND INTERPRETATIONS* (1997); T. Alexander Aleinikoff, *The Constitution in Context: The Continuing Significance of Race*, 63 COLO. L. REV. 325 (1992).

17. See, e.g., Harold Sigall & Richard Page, *Current Stereotypes: A Little Fading, a Little Faking*, 18 J. PERSONALITY & SOC. PSYCHOL. 247, 252 (1971).

18. See, e.g., Sheri Lynn Johnson, *Unconscious Racism and the Criminal Law*, 73 CORNELL L. REV. 1016, 1027-28 (1988) (reviewing the literature).

Both of these (related) conceptions of modern racism raise the troubling *possibility* that defense counsel, who are charged with undivided loyalty to their clients, and presumed to serve as a shield against racial bias on the part of other criminal justice system actors, may in fact experience both compromised loyalty and judgment when they serve African-American or Latino clients. On the other hand, perhaps capital defense attorneys, either by self-selection or by training, are different than the rest of the population in this regard. This Article describes preliminary data suggesting that they are not. That they exhibit similar automatic racial attitudes does not, of course, prove that their performance is impaired by those attitudes; we leave the implications of our findings for the discussion section.

II. GATHERING THE DATA

A. *The Instrument*

The Implicit Association Test (IAT) was developed to measure the relative strength with which groups (or individuals) are associated with positive and negative evaluations. It has been used to measure attitudes about a variety of issues, including race, gender, age, and political candidates, and is accepted as a valid research tool. All variations of the IAT use some form of response latency to assess those attitudes that are “automatic,” as opposed to attitudes that are subject to intention or control,¹⁹ and operate on the principal that it should be easier to make the same behavioral response to concepts that are associated than to concepts that are not associated. Computerized versions of the IAT achieve these pairings by assigning a keyboard key to be pressed in response to items from categories such as *old* or *bad*, and another key to be pressed in response to items from the opposite categories, such as *young* or *good*. Then the pairings are switched, with the subject being asked to press one key in response to either *old* or *good*, and another key in response to items from either the *young* or *bad* category. The differential speed required to complete these two opposite pairings is measured, yielding information both about the direction of the implicit attitude and the strength of the association; thus, if it takes longer for a subject to complete the pairings of *old* and *good* than the pairings of *young* and *good*, then the researcher infers that young is automatically associated with good for that subject, and old with bad.

19. This is not to say that the subject cannot thwart the accurate measure of these automatic attitudes, should he or she choose to do so, but only that subjects who faithfully follow the directions will reflect their unconscious associations, not their conscious preferences.






More than 500,000 IATs that focus on race have been taken online using this format.²⁰ In the race IAT, subjects are first asked to pair “good” words with pictures of white faces and “bad” words with pictures of black faces, and then to reverse the pairings.²¹ Here, if the subject can more quickly complete the task when white and good (and black and bad) are paired than when black and good (and white and bad) are paired, it means that the subject automatically pairs white with good—and black with bad.






A paper and pencil version of the race IAT is also available and was used in collecting these data because of time and computer accessibility constraints. In the paper and pencil version, subjects are faced with a column of words and faces, which he or she is asked to categorize “as quickly as possible without making too many mistakes” in twenty seconds. First the subjects complete two practice tests; the first test pairs flowers with good and insects with bad, and the second pairs insects with good and flowers with bad.²² This practice is designed to make the subject familiar with the pairing and check-off process, and accustomed to the idea of switching which items are paired. Then the subjects are asked to familiarize themselves with four new categories. One category is “good,” which is composed of the words flower, pretty, and love; a second category is “bad,” composed of the words ugly, vomit, and hate; the third category is “white,” which is composed of five white faces; and the fourth category is “black,” which is composed of five black faces. A short column of these words and pictures appears, and the subjects are instructed to go down the column checking the items that are “white or good” on the left of the item and items that are “black or bad” on the right of the item. After permitting questions, the subjects are told that when they turn the next page, they will be asked to check white faces or good on the left, and black faces or bad on the right, completing as many as possible in the allotted time, as the (shorter) sample below indicates:

20. Craig Lambert, *Stealthy Attitudes*, at www.harvard-magazine.com/on-line/070270.html. (last visited Jan. 20, 2004).






21. The order of the tasks may slightly alter the strength, but not the direction of the effect. Brian A. Nosek et al., *Understanding and Using the Implicit Association: Test II. Methodological Issues*, at <http://www.briannosek.com/papers/ngb.IATmethodII.2003.pdf> (last visited Jan. 20, 2004).


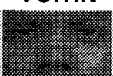



22. In the practice test, subjects attempt to go through as many items as possible in the time limit, checking items that are either flowers or good on the left side of the page and either insects or bad on the right side of the page; then on the next page, the subject is asked to check off items that are either insects or good on the left and items that are either flowers or bad on the right.

White Good		Black Bad
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>

White Good		Black Bad
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>

After completing this task, the subjects are asked to turn the page, and the new pairing of black with good and white with bad is explained. Subjects then complete the same task with the new pairing, as the sample below indicates:

Black Good		White Bad
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>

Black Good		White Bad
<input type="radio"/>	joy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>

The number of items correctly completed on each test is then counted; it is not the number of items a particular subject can complete that is of significance, but the difference in the number of items

he or she completes when white is paired with good and black with bad, as contrasted with the number completed when black is paired with good and white with bad.

B. The Subjects

Two of the three data sets are the product of presentations made by one of the authors at training sessions for capital defense lawyers.²³ The first was obtained at an annual gathering of lawyers who represent death row inmates in federal habeas corpus proceedings. Most of these attorneys are experienced capital litigators, but some novices were invited. The training session was held in Nashville, Tennessee, but the attorneys came from all over the country. These subjects we will call the “habeas lawyers.” The second data set was collected at a training session for Georgia trial lawyers involved in representing defendants charged with capital crimes; these subjects we will call the “trial lawyers.” We asked all of these subjects for their race. For the trial lawyers, we also collected the age and gender of each subject. The third data set is composed of most of the students in a first-year constitutional law class at Cornell Law School; we will call them the “law students.” For the law students we collected race, age, and gender data.

III. RESULTS

As shown in Table 1, on average, the subjects in all three groups completed more items when white was paired with good and black with bad than when black was paired with good and white with bad, and in all three groups these differences are statistically significant. For example, the table’s first row shows that, for all subjects combined, an average of 16.4 correct responses were given when white was paired with good and black with bad compared to 13.5 correct responses when black was paired with good and white with bad. A test of the statistical significance of the difference in means, computed using a t-test, yields a p-value of less than 0.0001 for our 321 subjects, as reported in column (4). The difference in means is significant at that level for each group as well as for all groups combined.

23. The primary purpose in administering all of these tests was educational, and after the tests were administered, the subjects were asked to compare the number of items they were able to complete in the time period, and then the interpretation of the test suggested by the authors was presented. In all three cases, substantial discussion followed.

TABLE 1. NUMBER OF ITEMS COMPLETED WHEN "GOOD" IS PAIRED WITH WHITE FACES V. WHEN "GOOD" IS PAIRED WITH BLACK FACES, BY SUBJECT GROUP

	Average Number of Correct Responses When "Good" is Paired with:		Difference = (1) - (2) (3)	Significance of Difference		N (6)
	(1) White Faces	(2) Black Faces		(4) Means	(5) Medians	
All subjects	16.4	13.5	2.8	<.0001	<.0001	321
Habeas lawyers	14.4	12.1	2.3	<.0001	<.0001	146
Trial lawyers	16.7	13.6	3.0	<.0001	<.0001	92
Law students	19.5	15.9	3.6	<.0001	<.0001	83

To eliminate the possibility that the significance of the difference is the product of a few extreme individuals and to assure that the results are not sensitive to the distributive assumptions associated with the t-test, we also tested the statistical significance of the differences in median correct scores. Column (5) shows that the median differences are also significant beyond the 0.0001 level for all groups and for the groups aggregated into a single sample. Thus, the probability of observing by chance differences as large as, or larger than, those observed is vanishingly small.

As shown in the Table 5 regression models reported below, differences between the three groups—habeas lawyers, trial lawyers, and law students—are not statistically significant. Within each group, however, the differences between black and white subjects are highly significant. Table 2 expands on Table 1 by reporting the results broken down by race within each subject group. Table 2 shows that, for all groups combined, as well as for each subject group, the average white subject completed more items when white was paired with good than when black was paired with good, while the average black²⁴ subject completed more items when black was paired with good. For example, for all white subjects combined, an average of 16.6 correct responses were given when white was paired with good and black with bad compared to 13.2 correct responses when black was paired with good and white with bad. In contrast, for all black subjects combined, an average of 14.4 correct responses were given when white was paired with good and black with bad compared to 15.8 correct responses when black was paired with good and white with bad.

24. We do not refer to our subjects as "African American" because, although virtually all of the black subjects in the two lawyer groups were African American, in the constitutional law class, Caribbean and African students are a significant part of the total.

Column (3) of Table 2, which is the difference in correct responses between columns (1) and (2), shows that this racial difference persists for all three groups. White subjects in all these groups provided more correct responses when white was paired with good. Black subjects in all three groups provided more correct responses when black was paired with good. For the most part, we limited our analysis to white and black subjects because the subjects in the two lawyer groups were almost exclusively black and white. In the constitutional law class, however, there were fourteen Asian²⁵ subjects, enough to make analysis of their results potentially meaningful. We found that the Asian subjects were not significantly different than the white subjects, but were significantly different than the black subjects.

TABLE 2. NUMBER OF CORRECT RESPONSES WHEN "GOOD" IS PAIRED WITH WHITE FACES V. WHEN "GOOD" IS PAIRED WITH BLACK FACES, BY RACE OF SUBJECTS

	Average Number of Correct Responses When "Good" is Paired with:		Difference = (1) - (2)	Significance of Difference		N
	(1)	(2)		(4)	(5)	
	White Faces	Black Faces	(3)	Means	Medians	(6)
All subjects	16.4	13.5	2.8	<.0001	<.0001	321
Whites	16.6	13.2	3.4	<.0001	<.0001	281
Blacks	14.4	15.8	-1.4	.097	.039	40
Habeas lawyers	14.4	12.1	2.3	<.0001	<.0001	146
Whites	14.5	12.0	2.4	<.0001	<.0001	140
Blacks	12.7	14.0	-1.3	.669	.399	6
Trial lawyers	16.7	13.6	3.0	<.0001	<.0001	92
Whites	17.4	13.1	4.3	<.0001	<.0001	70
Blacks	14.4	15.5	-1.1	.112	.086	22
Law Students	19.5	15.9	3.6	<.0001	<.0001	83
Whites	20.2	15.6	4.5	<.0001	<.0001	71
Blacks	15.4	17.3	-1.8	.384	.432	12

Columns (4) and (5) show that, for all groups of white subjects, the differences in the numbers of correct responses between white being paired with good and black being paired with good are highly statistically significant. In part because of the smaller number of black subjects in each group, and in part because the mean difference in their white/good and black/good scores is smaller, differences in black subjects' performance on the white/good pairing as compared to the

25. The constitutional law class comprised both United States citizens and nationals of other countries, and each racial group included foreign students. To determine whether the Asian (as opposed to Asian-American) subjects might be driving these results, we removed them from the analysis, but the results were unchanged.

black/good pairing are not statistically significant for the three sub-groups considered separately. When the three groups are combined, however, the differences in the median numbers of correct responses for black subjects between white being paired with good and black being paired with good are statistically significant at the .039 level, and the means differ at the .097 level.

Thus, both white and black subjects score higher when their own racial group is paired with good. But Table 2 indicates that the difference is not of the same magnitude. Table 3 explores this difference by reorganizing some of the data in Table 2. Table 3 again shows that both black and white subjects on average scored higher when their own racial group was paired with good than when it was paired with bad. In all three groups, white subjects' scores were more affected by which race was paired with good than were the scores of black subjects. Column (3) reports these differences in differences for each subject group.

TABLE 3. DIFFERENCE BETWEEN NUMBER OF ITEMS COMPLETED WHEN "GOOD" IS PAIRED WITH OWN-RACE FACES V. WHEN "GOOD" IS PAIRED WITH OTHER-RACE FACES, BY SUBJECT GROUP

	Mean Difference Between Number of Correct Responses When "Good" is Paired with Subject's Own Race v. When "Good" is Paired with Other Race		Difference = (1) - (2) (3)	Significance of Difference		N (6)
	(1) Whites	(2) Blacks		(4) Means	(5) Medians	
All subjects	3.4	1.4	2.1	.008	.018	321
Habeas lawyers	2.4	1.3	1.1	.548	.839	146
Trial lawyers	4.3	1.1	3.2	.0003	.001	92
Law Students	4.5	1.8	2.7	.130	.123	83

Columns (4) and (5) show that, for the groups combined (as well as for the trial lawyers taken individually), the differences in differences are statistically significant. For the two other groups, the direction of the effect is the same but the effect is not statistically significant at traditional levels, a result likely attributable to the small number of black subjects. The races' different reaction also emerges in another measure. Only a quarter of white subjects scored the same or better on the black/good pairing as on the white/good pairing, while almost half of the black subjects scored the same or better on the white/good pairing as on the black/good pairing.

We also examined the effect of gender on the performance, holding race constant. Table 4 below reports the results. (Its sample is

smaller than the samples in the other tables because we did not record the gender of the subjects for the first group we sampled, the habeas lawyers.) Table 4 can be used to explore two separable questions. First, do the results in Tables 1 and 2 hold when the sample is disaggregated by gender? That is, do males and females separately confirm the pattern of more correct responses when white faces are paired with “good” than when black faces are paired with “good?” Second, Table 4 explores whether the pattern of responses varies not only by race, as suggested in Tables 2 and 3, but also by gender.

Table 4 shows that the core result of significantly different reactions to white/good pairings and black/good pairings holds for black males, white females, and white males. However, for white females and males the significant advantage is in white/good pairings, while for black males, the significant advantage is in black/good pairings. For black females, the tendency is toward easier black/good pairings, but the difference, as reported in columns (3) through (5), is not statistically significant.

Column (6) explores whether the difference in differences varies statistically significantly across the race/gender groups. White women (as well as black women and black men) are significantly different from white men. For both blacks and whites, the tendency is for men to reflect a greater own-race advantage than women, but for black subjects, the difference is smaller and not significant.

TABLE 4. NUMBER OF ITEMS CORRECTLY COMPLETED WHEN “GOOD” IS PAIRED WITH WHITE FACES V. WHEN “GOOD” IS PAIRED WITH BLACK FACES, BY RACE AND GENDER

	Average Number of Correct Responses When “Good” is Paired with:		Difference = (1) – (2) (3)	Significance of Difference (1)–(2) From White Males’ Difference		N (7)	
	(1) White Faces	(2) Black Faces		(4) Means	(5) Medians		(6)
Black females	15.0	16.1	–1.1	.365	.397	.002	22
Black males	14.3	16.1	–1.8	.020	.025	.010	12
White females	18.1	15.1	2.9	.0001	.0002	.041	55
White males	18.6	13.8	4.8	<.0001	<.0001		100

Note. Column (6), which shows the level significance of difference from white males, reports a test of the difference in the median number of good characteristics correctly associated with one’s own race and the other race. A t-test of the difference in means is also statistically significant.

The analysis so far explores subject group differences and race/gender differences in isolation from one another. To assess all factors simultaneously, we employ regression analysis and report the results

in Table 5 and include age in the models for those subjects for whom age is known. Models reported in columns (1) and (2) explore the differences described in Tables (1) and (2)—the difference between subjects' scores when whites are paired with good characteristics and when blacks are paired with good characteristics. We report both ordinary least squares models and negative binomial models. Negative binomial models are appropriate because the dependent variables in our models are count data—the difference in the number of correct responses.²⁶ Columns (3) and (4) report models that explore the difference in differences between the white/good and black/good scores. It thus assesses whether the explanatory variables in the model help explain the degree of within-race favoritism.

TABLE 5. REGRESSION MODELS OF DIFFERENCES IN SCORES
BETWEEN PAIRINGS

	(1) Dependent variable = Difference between subject's white/good and black/good scores		(2) Dependent variable = Difference between subjects own-race/good and other-race/good scores	
	OLS	Negative binomial	OLS	Negative binomial
Black female	-6.465** (4.88)	-0.344** (4.30)	-4.272** (3.37)	-0.285** (3.06)
Black male	-7.001** (8.59)	-0.381** (7.99)	-3.149** (3.50)	-0.206** (3.39)
White female	-2.035* (2.51)	-0.098* (2.50)	-2.099* (2.59)	-0.133* (2.54)
White male = reference category				
Trial lawyers	0.802 (0.91)	0.041 (0.95)	0.645 (0.73)	0.045 (0.78)
Habeas lawyers	-2.437 (1.02)	-0.128 (0.98)	-1.947 (0.81)	-0.136 (0.79)
Law students = reference category				
Age	-0.620 (1.56)	-0.031 (1.59)	-0.835* (2.12)	-0.054* (2.13)
Age missing	-1.613 (0.58)	-0.075 (0.51)	-2.794 (0.98)	-0.178 (0.90)
Gender missing	0.166 (0.12)	0.014 (0.19)	0.344 (0.25)	0.035 (0.35)
Constant	6.278** (6.43)	3.154** (69.42)	6.912** (7.12)	2.954** (51.01)
Observations	321	321	321	321
R-squared or pseudo R-squared	0.16	.030	0.10	.017

Robust t statistics in parentheses

* significant at 5%; ** significant at 1%

26. ALAN AGRESTI, CATEGORICAL DATA ANALYSIS 559 (2d ed. 2002). To assure positive values of the dependent variable in the negative binomial analyses, a positive integer was added to the values of the dependent variable for all observations. *Id.*

Table 5 confirms the results in the earlier tables. Models (1) and (2) show that black females, black males, and white females all tend to have less extreme differences between the white/good and black/good conditions than do white males. The results are all significant at or beyond the 0.05 level. The results do not depend on the functional form, least squares, or negative binomial of the models used. The story is constant across models (1) and (2). Similarly, models (3) and (4) show that the difference between the number of good characteristics correctly associated with one's own race and the other race is associated with a race/gender effect. White males again differ statistically significantly from the other three race/gender combinations. They tend to record greater differences between own-race/good and other-race/good pairings than do the other race/gender combinations. The result is statistically significant and persists in both least squares and negative binomial models.

Again, differences between the three subject groups—habeas lawyers, trial lawyers, and law students—are not significant. The only surprise is that age, which viewed in isolation was not a significant predictor of performance, is slightly but significantly correlated, and in a *negative* direction, with size of mean white/good and black/good differences

IV. DISCUSSION

Because speed at these tasks reflects relative ease in associating two categories of items, the creators of the IAT describe a subject who is more adept at pairing white with good than black with good as having an “automatic preference” for white. In adopting this terminology, we note that “preference” as used here does not imply a conscious choice, but merely an automatic association; when we say that subjects have an “automatic preference for white,” we mean nothing more than that they automatically associate white with good and black with bad.²⁷ Part III of this Article reported both the direction of the automatic associations and their relative strength for each professional group (habeas lawyers, trial lawyers and law students) and for demographic subgroups of each group (dividing the groups by race, gender,

27. This preference is not the result of greater familiarity with white faces. Researchers have eliminated the familiarity hypothesis by using familiar black faces (or names) and unfamiliar white faces (or names), and find that the automatic preference of whites subjects for white survives unchanged. See generally Nilanjan Dasgupta et al., *Automatic Preference for White Americans: Eliminating the Familiarity Explanation*, 36 J. EXPERIMENTAL SOC. PSYCHOL. 316 (2000), available at <http://faculty.washington.edu/agg/pdf/DasguptaEtAl.JESP2000.pdf> (last visited Apr. 13, 2004).

and age). Because we used a paper and pencil version of the IAT, we are unable to compare the *strength* of observed preferences with those found in the larger IAT databases,²⁸ but we can compare the *direction* of those preferences with those observed in the larger, more diverse (and overwhelmingly lay) subject pool, as well as relative differences between the demographic subgroups in each data set. After making these comparisons, we will turn to the question of what behavioral implications, if any, these observed automatic preferences entail.

The direction and demographic distribution of the automatic preferences we observed are strikingly consistent with those observed in the larger trials not targeted at occupational subgroups.²⁹ Data from the websites, like our data, reflect significant automatic preference for whites among white subjects; this is true when good and bad are in turn paired with black and white faces, and it is also true when they are paired with first names that are associated with black or white persons.³⁰ Black web subjects, on the other hand, show an automatic preference for black faces or names, but that preference, like the preference of our defense lawyer and law student subjects, is weaker than is white subjects' automatic preference for white faces or names.³¹ Our responses and the web responses both mimic earlier laboratory results with college student subjects.³² Moreover, the web researchers found, as we find, that Asian respondents show a pro-white bias level

28. We were, however, able to compare the strength of automatic preferences in our law student and defense lawyer groups with the strength of the automatic preferences of *undergraduates* who attended the Race to Execution Symposium, and found no significant differences. We do not report the symposium data because it reflects such a hodge-podge of respondents that we were uncertain as to the meaning of any results, but the undergraduate component of the symposium was large enough to analyze, and gives us some reason to believe that our sample does not differ in strength of the automatic-race preference from lay populations, just as the web data gives us confidence that the direction of our sample's preferences is indistinguishable from that of the general population.

29. The IAT has been administered in laboratory settings to college students, but most of the reported data has been harvested from demonstration web sites. As the authors who report web-site findings acknowledge, there are selection effects that render these web subjects different from a random sample of the population; most obviously, these subjects must have access to a computer, awareness of the test, and interest in taking the test. (Recruitment occurred through four known channels: media coverage, links from other sites, search engines, and word of mouth, with media coverage being the most significant source.) See Brian A. Nosek et al., *Harvesting Implicit Group Attitudes and Beliefs from a Demonstration Web Site*, 6 GROUP DYNAMICS: THEORY RES. PRAC. 101, 102-103 (2002), available at <http://projectimplicit.net/nosek//papers/harvesting.GroupDynamics.pdf> (last visited Apr. 13, 2004). Nonetheless, the data represents a much more diverse subject pool than those ordinarily surveyed in psychology experiments, *id.* at 102, and one likely to be largely composed of nonlawyers.

30. *Id.* at 105-06.

31. *Id.* at 105.

32. *Id.*

comparable to that of white respondents.³³ Finally, while our regression reflected that white women showed slightly, but significantly less, negativity toward black faces than did white men, web researchers reported that women subjects (not broken down by race, but predominantly white) showed slightly less automatic preference for whites.³⁴

The reader will recall that our results on age were of marginal significance in the regression, with older age predicting *slightly less* of an automatic preference for white. This is the only place our results differ even slightly from those of the web research; there, older participants did not differ at all from younger ones on the analogous face test. Several explanations for these disparate findings are possible. While it is possible that our older capital defense attorneys, unlike their age peers in the general population, actually do harbor weaker automatic preferences for white than do otherwise similar subjects, web research on the performance of various measuring algorithms suggests that our results may be an artifact of our scoring method, which does not account for the fact that older subjects perform all of the tasks more slowly.³⁵

Thus, our capital defense attorneys, both trial and post-conviction (trial lawyers and habeas lawyers), look like our law students in their implicit attitudes about race and, as far as we can tell, pretty much like the rest of the population. White men have the strongest automatic preference for white, followed by white women. The responses of Asian subjects look like those of white subjects. In contrast, black subjects have an automatic preference for black, but it is significantly smaller than the preference white and Asian subjects have for white. If one imagines that automatic reactions are the combined product of culture and individual experience, these patterns are not surprising.

Looking at other empirical evidence concerning lawyers, the observed similarities between defense lawyers, law students, and the lay population are also not surprising. Popular impression to the con-

33. Nosek, *supra* note 29, at 110. The web data also has a substantial number of Hispanic respondents, who also show a pattern similar to that of white respondents. *Id.*

34. *Id.* at 109.

35. In our scoring, a *proportionally* equal impairment would result in a smaller absolute difference score for an older subject; for example, if an older subject completed fifteen items in the white/good test, but only twelve in the black/good test, while a younger subject completed twenty items on the white/good pairing, but only sixteen on the black/good pairing, the older subject, by our measurement, would have a weaker preference for white. In contrast, web researchers have developed a more complicated algorithm for measuring automatic preference, one that includes calibration by each respondent's latency variability. Anthony G. Greenwald et al., *Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm*, 85 J. PERSONALITY & SOC. COGNITION 197, 214 (2003). That algorithm is not transferable to the paper and pencil test, so we could not benefit from its development.

trary, comparisons between judges and juries generally find little difference in the damages they award.³⁶ And, most relevantly, judges appear to be just as susceptible as are jurors to three cognitive illusions that hinder accurate decision making: anchoring, hindsight bias, and egocentric bias.³⁷ Finally, judges' political orientation has been shown to have some influence in politically charged cases.³⁸ As is slowly being recognized across the field, "lawyers are like other people and suffer from the same human failings as those not admitted to the bar."³⁹ Though we have presented evidence that white capital defense lawyers, like the rest of the population, have automatic reactions that make associating white with good easier than associating white with bad, we have by no means proved that they treat black clients (or witnesses, jurors, attorneys, or judges for that matter) differently than they treat their white counterparts. Indeed, even the evidence on the relationship between *explicit* prejudice and discrimination is complex. Subjects who acknowledge negative attitudes toward vulnerable groups do not always discriminate against them; either lack of opportunity or social disapproval may inhibit the expression of those attitudes.⁴⁰ Likewise, subjects who disavow negative attitudes toward a vulnerable group may nonetheless discriminate against that group for a variety of reasons, including social pressure to do so and covert or unconscious stereotypes about the target group.⁴¹ Because measurement of implicit racial attitudes is quite new, even less is known about how those attitudes affect behavior.

It is, however, possible to note what external factors are known to make discrimination more likely and consider their presence or ab-

36. See, e.g., Theodore Eisenberg et al., *Juries, Judges, and Punitive Damages: An Empirical Study*, 87 CORNELL L. REV. 743 (2002) (finding no substantial difference in the rates judges and juries award punitive damages and no substantial difference in the relation between the size of punitive and compensatory damages for judges and juries); Neil Vidmar & Jeffrey J. Rice, *Assessments of Noneconomic Damage Awards in Medical Negligence: A Comparison of Jurors with Legal Professionals*, 78 IOWA L. REV. 883, 896 (1993); Roselle L. Wissler et al., *Decisionmaking About General Damages: A Comparison of Jurors, Judges and Lawyers*, 98 MICH. L. REV. 751, 812 (1999).

37. Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 816 (2001). Judges also fall prey to framing effects and the representative heuristic, though they are less susceptible than others.

38. Richard Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717 (1997).

39. Fred C. Zacharias, *The Humanization of Lawyers*, 2002 PROF. LAW. 9, 10 (arguing that this is one of the significant transformations in professional responsibility scholarship in recent years).

40. GORDON W. ALLPORT, *THE NATURE OF PREJUDICE* 56-57 (1954).

41. The sharp differences between explicit attitudes and implicit attitudes found by web researchers attests to the frequency of unconscious (or covert) negative attitudes. See, e.g., Nosek et al., *supra* note 29, at 111.

sence in the context of capital representation of a minority-race defendant. As already mentioned, social approval (or even the absence of social disapproval) enhances the likelihood of discrimination. The factor of social approval must vary widely in capital representation, depending in part on whether the relevant audience is the local prosecutor (with whom the next case must be plea-bargained), the local judge (who may be the source of the next appointment as counsel), the local defense bar (whose attitudes differ by area), the jurors (who in some small localities are potential clients), or the national capital defense bar. Stereotypes are more likely to alter judgement when a task is complex,⁴² or a decision difficult,⁴³ or when the context activates stereotypes.⁴⁴ Certainly the task of representing a capital defendant or habeas petitioner is complex and involves many difficult decisions. Moreover, the capital litigation context—in which the defendant is accused of heinous crimes—seems especially likely to activate stereotypes of violence and criminality.

There is some evidence that awareness of automatic reactions can trigger attempts to counteract them, and that such attempts are sometimes successful.⁴⁵ Here we have no clear indication of whether capital defense attorneys struggle against their own automatic reactions. On the one hand, we would predict from their ideological commitments that they would be inclined to do so, *if they were aware that they had such reactions*; on the other hand, anecdotal experience in administering these tests leads us to believe that many capital defense attorneys were surprised at their own automatic preferences and,

42. ALLPORT, *supra* note 40, at 56-57.

43. Thus, for example, mock jury studies of race and guilt attribution find race-of-defendant effects in the “marginal evidence” cases; when proof of guilt is either weak or strong, these effects do not appear. See Sheri Lynn Johnson, *Black Innocence and the White Jury*, 83 MICH. L. REV. 1611, 1626-34 (1985) (reviewing the literature). The same pattern is reflected in the Baldus data on capital cases in Georgia; race-of-victim effects are found in the cases that are in the middle range of aggravating factors. See *McCleskey v. Kemp*, 481 U.S. 279, 364 (1987) (Stevens, J., dissenting).

44. See, e.g., HOWARD EHRLICH, *THE SOCIAL PSYCHOLOGY OF PREJUDICE* 40 (1973). See also Patricia Linville & Edward E. Jones, *Polarized Appraisals of Outgroup Members*, 38 J. PERSONALITY & SOC. PSYCHOL. 689 (1980). IAT research reveals that automatic race evaluations are significantly stronger when the black exemplars were disliked than when they were liked. Jason P. Mitchell et al., *Contextual Variations in Implicit Evaluation*, 132 J. EXPERIMENTAL PSYCHOL. 455, 460 (2003).

45. Annie Murphy Paul, *Where Bias Begins: The Truth About Stereotypes*, PSYCHOL. TODAY, May-June 1998, at 52 (quoting researcher Margo Monteith). However, attempts to suppress stereotypes may also sometimes backfire, causing them to return in a stronger form. *Id.*

therefore, would not have previously realized that they should struggle against those preferences.⁴⁶

IV. CONCLUSION: WHO IS POLICING THE BIAS POLICE?

Our initial foray into the racial attitudes of capital defense lawyers permits us a modest conclusion: Like the rest of the population, race influences their automatic reactions. This proves neither that race does, nor that it does not, often influence the quality of representation afforded the black clients of these attorneys. For judges reviewing the effectiveness of the assistance of counsel provided to capital defendants, our data suggest that they should not *assume* that race has not influenced the actions of defense counsel, and that they should not *assume* that counsel will be sensitive to the racial bias of other criminal justice system actors. For the capital defense lawyers themselves, it suggests that introspection about racial stereotypes and reactions, as well as vigilance concerning those effects on others, is necessary. For the public, it may be yet another reason to doubt the evenhandedness of capital punishment.

46. It is clear that the process of "de-automatization" works only for people disturbed by the discrepancy between their conscious and unconscious beliefs. *Id.* Creation of awareness of the discrepancy, in both the lawyers and the law students, was the primary purpose of administering the IAT to these groups.