

# Module 3: The Analysis of $2 \times 2 \times 2$ (Multiway) Contingency Tables: Explaining Simpson's Paradox and Demonstrating Racial Bias in the Imposition of the Death Penalty

It is the mark of a truly intelligent person to be moved by statistics.

– George Bernard Shaw

**Abstract:** This module discusses the two major topics of Simpson's paradox and the Supreme Court decision in *McCleskey v. Kemp* (1987). Simpson's paradox is ubiquitous in the misinterpretation of data; it is said to be present whenever a relationship that appears to exist at an aggregated level disappears or reverses when disaggregated and viewed within levels. A common mechanism for displaying data that manifests such a reversal phenomenon is through a multiway contingency table, often of the  $2 \times 2 \times 2$  variety. For example, much of the evidence discussed in *McCleskey v. Kemp* was cross-categorized by three dichotomous variables: race of the victim (black or white), race of the defendant (black or white), and whether the death penalty was imposed (yes or no). Despite incontrovertible evidence that the race of the victim plays a significant role in whether the death penalty is imposed, the holding in *McCleskey v. Kemp* was as follows: Despite statistical evidence of a profound racial disparity in application of the death penalty, such evidence is insufficient to invalidate defendant's death sentence.

## Contents

<b>1 A Few Introductory Examples of Simpson’s Paradox</b>	<b>2</b>
<b>2 Statistical Sleuthing and the Imposition of the Death Penalty: McCleskey v. Kemp (1987)</b>	<b>15</b>
2.1 United States Supreme Court, McCleskey v. Kemp (1987): Syllabus . . . . .	18
<b>3 Appendix: United States Supreme Court, McCleskey v. Kemp (1987): Majority Opinion and Dissent</b>	<b>26</b>

## 1 A Few Introductory Examples of Simpson’s Paradox

An enjoyable diversion on Saturday mornings is the NPR radio show, *Car Talk*, with Click and Clack, The Tappet Brothers (aka Ray and Tom Magliozzi). A regular feature of the show, besides giving advice on cars, is The Puzzler; a recent example on September 22, 2012 gives a nice introductory example of one main topic of this chapter, Simpson’s paradox. It is called, Take Ray Out to the Ball Game, and is stated as follows on the Car Talk website:

Take Ray Out to the Ball Game:

RAY: As you might guess, I’m a baseball fan. And now that the season is in its waning days, I thought I’d use this baseball Puzzler I’ve been saving.

There are two rookie players, Bluto and Popeye, who started the

season on opening day and made a wager as to which one would have the best batting average at the end of the season.

Well, the last day of the season arrives, and not much is going to change—especially considering that neither one of them is in the starting lineup.

Bluto says, “Hey, Popeye, what did you bat for the first half of the year?”

Popeye answers, “I batted .250.”

And Bluto responds, “Well, I got you there. I batted .300. How about after the All-Star break?”

Proudly, Popeye pipes up, “I batted .375.”

Bluto says, “Pretty good, but I batted .400. Fork over the 20 bucks that we bet.”

The bat boy, Dougie, saunters over and says, “Don’t pay the 20 bucks, Popeye. I think you won.”

TOM: Why is someone who batted .375 not playing in the last game of the season? That’s what I want to know!

RAY: Good point. But the question is this: How could Popeye have won?

---

RAY: Here’s the answer. Let’s assume that they both had 600 at-bats.

TOM: Yeah.

RAY: If Bluto batted .300 for the first half of the season and he had 500 at-bats during that first half of the season.

TOM: Oooh. Yeah.

RAY: He got 150 hits. One hundred fifty over 500 is a .300 average, right?

TOM: Mmm-hmm. So he would have gotten 150.

RAY: Yeah. OK? If Popeye batted .250 and had 100 at-bats, he would have had 25 for 100. The second half of the season, Bluto bats .400. How does he do that? Well, we know he had 500 at-bats in the first half.

TOM: So he's only been up 100 times in the second half of the season.

RAY: And he got 40 hits.

Popeye bats .375.

TOM: But he's up 500 times.

RAY: And he gets 187 and a half hits. One of them was a check-swing single over the infield. They only count that as half a hit. So now, let's ... let's figure it all out.

Bluto batted 600 times. How many total hits did he get?

TOM: 190.

RAY: Right. How about Popeye? How many hits did he get?

TOM: 212 and a half.

RAY: And when you figure that out, Bluto batted .316 for the season. Even though he batted .300 and .400 in each half.

TOM: Yeah.

RAY: And Popeye bats .353 and wins the batting title.

TOM: No kidding!

RAY: Pretty good, huh?

Putting the data about Bluto and Popeye in the form of a  $2 \times 2$  table that gives batting averages both before and after the All-Star break as well as for the full year should help see what is happening:

	Before Break	After Break	Full Year
Bluto	$\frac{150}{500} = .300$	$\frac{40}{100} = .400$	$\frac{190}{600} = .317$
Popeye	$\frac{25}{100} = .250$	$\frac{187.5}{500} = .375$	$\frac{212.5}{600} = .354$

Thus, the batting averages of Popeye before and after the break (.250 and .375) can be less than for Bluto (.300 and .400), even though for the full year, Popeye’s average of .354 is better than Bluto’s .317. This type of counterintuitive situation is referred to as a “reversal paradox” or more usually by the term, “Simpson’s paradox.”

The unusual phenomenon presented by the example above occurs frequently in the analysis of multiway contingency tables. Basically, various relations that appear to be present when data are conditioned on the levels of one variable, either disappear or change “direction” when aggregation occurs over the levels of the conditioning variable. A well-known real-life example is the Berkeley sex bias case applicable to graduate school (Bickel, Hammel, & O’Connell, 1975). The table below shows the aggregate admission figures for the fall of 1973:

	Number of applicants	Percent admitted
Men	8442	44
Women	4321	35

Given these data, there appears to be a *prima facie* case for bias because a lower percentage of women than men is admitted.

Although a bias seems to be present against women at the aggregate level, the situation becomes less clear when the data are

broken down by major. Because no department is significantly biased against women, and in fact, most have a small bias against men, we have another instance of Simpson’s paradox. Apparently, women tend to apply to competitive departments with lower rates of admission among qualified applicants (for example, English); men tend to apply to departments with generally higher rates of admission (for example, Engineering).<sup>1</sup>

A different example showing a similar point can be given using data on the differential imposition of a death sentence depending on the race of the defendant and the victim. These data are from twenty Florida counties during 1976-1977 (Radelet, 1981):

Defendant	Death Penalty	
	Yes	No
White	19 (12%)	141
Black	17 (10%)	149

Because 12% of white defendants receive the Death penalty and only 10% of blacks, at this aggregate level there appears to be no bias against blacks. But when the data are disaggregated, the situation appears to change:

---

<sup>1</sup>A question arises as to whether an argument for bias “falls apart” because of Simpson’s paradox. Interesting, in many cases the authors have seen like this, there is a variable that if interpreted in a slightly different way would make a case for bias even at the disaggregated level. Here, why do the differential admission quotas interact with sex? In other words, is it inherently discriminatory to women if the majors to which they apply most heavily are also those with the most limiting admission quotas?

Victim	Defendant	Death Penalty	
		Yes	No
White	White	19 (13%)	132
White	Black	11 (17%)	52
Black	White	0 (0%)	9
Black	Black	6 (6%)	97

When aggregated over victim race, there is a higher percentage of white defendants (12%) receiving the death penalty than black defendants (10%), so apparently, there is a slight race bias against whites. But when looking within the race of the victim, black defendants have the higher percentages of receiving the death sentence compared to white defendants (17% to 13% for white victims; 6% to 0% for black victims). The conclusion is disconcerting: the value of a victim is worth more if white than if black, and because more whites kill whites, there appears to be a slight bias against whites at the aggregate level. But for both types of victims, blacks are more likely to receive the death penalty.<sup>2</sup>

A common way to explain what occurs in Simpson’s paradox is to use contingency tables. For convenience, we restrict discussion

---

<sup>2</sup>Simpson’s paradox is a very common occurrence, and even through it can be “explained away” by the influence of differential marginal frequencies, the question remains as to why the differential marginal frequencies are present in the first place. Generally, a case can be made that gives an argument for bias or discrimination in an alternative framework, for example, differential admission quotas or differing values on a life. A more recent study similar to Radelet (1981) is from the *New York Times*, April 20, 2001, reported in a short article by Fox Butterfield, “Victims’ Race Affects Decisions on Killers’ Sentence, Study Finds.”

to the simple  $2 \times 2 \times 2$  case, and use the “death penalty” data as an illustration. There are two general approaches based on conditional probabilities. One that is presented below involves weighted averages; the second that we do not discuss relies on the language of events being conditionally positively correlated, but unconditionally negatively correlated (or the reverse).

To set up the numerical example, define three events:  $A$ ,  $B$ , and  $C$ :

$A$ : the death penalty is imposed;

$B$ : the defendant is black;

$C$ : the victim is white.

For reference later, we give a collection of conditional probabilities based on frequencies in the  $2 \times 2 \times 2$  contingency table:

$$P(A|B) = .10; P(A|\bar{B}) = .12; P(A|B \cap C) = .17;$$

$$P(A|\bar{B} \cap C) = .13; P(A|\bar{B} \cap \bar{C}) = .00;$$

$$P(C|B) = .38; P(\bar{C}|B) = .62; P(C|\bar{B}) = .94;$$

$$P(\bar{C}|\bar{B}) = .38; P(C) = .66; P(\bar{C}) = .34.$$

The explanation for Simpson’s paradox based on a weighted average begins by formally stating the paradox through conditional probabilities: It is possible to have

$$P(A|B) < P(A|\bar{B}) ,$$

but

$$P(A|B \cap C) \geq P(A|\bar{B} \cap C) ;$$

$$P(A|B \cap \bar{C}) \geq P(A|\bar{B} \cap \bar{C}) .$$

So, conditioning on the  $C$  and  $\bar{C}$  events, the relation reverses.

In labeling this reversal as anomalous, people reason that the conditional probability,  $P(A|B)$ , should be an average of

$$P(A|B \cap C) \text{ and } P(A|B \cap \bar{C}) ,$$

and similarly, that  $P(A|\bar{B})$  should be an average of

$$P(A|\bar{B} \cap C) \text{ and } P(A|\bar{B} \cap \bar{C}) .$$

Although this is true, it is not a simple average but one that is weighted:

$$P(A|B) = P(C|B)P(A|B \cap C) + P(\bar{C}|B)P(A|B \cap \bar{C}) ;$$

$$P(A|\bar{B}) = P(C|\bar{B})P(A|\bar{B} \cap C) + P(\bar{C}|\bar{B})P(A|\bar{B} \cap \bar{C}) .$$

If  $B$  and  $C$  are independent events,  $P(C|B) = P(C|\bar{B}) = P(C)$  and  $P(\bar{C}|B) = P(\bar{C}|\bar{B}) = P(\bar{C})$ . Also, under such independence,  $P(C)$  and  $P(\bar{C})$  ( $= 1 - P(C)$ ) would be the weights for constructing the average, and no reversal would occur. If  $B$  and  $C$  are not independent, however, a reversal can happen, as it does for our “death penalty” example:

$$.10 = P(A|B) = (.38)(.17) + (.62)(.06);$$

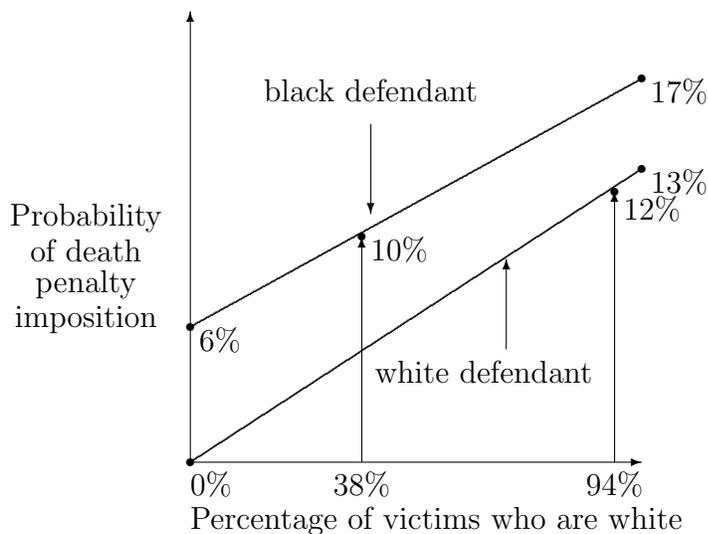
$$.12 = P(A|\bar{B}) = (.94)(.13) + (.06)(.00).$$

So, instead of the weights of .66 ( $= P(C)$ ) and .34 ( $= P(\bar{C})$ ), we use .38 ( $= P(C|B)$ ) and .62 ( $= P(\bar{C}|B)$ ); and .94 ( $= P(C|\bar{B})$ ) and .06 ( $= P(\bar{C}|\bar{B})$ ).

Figure 1 provides a convenient graphical representation for the reversal paradox in our “death penalty” illustration. This representation generalizes to any  $2 \times 2 \times 2$  contingency table. The  $x$ -axis is labeled as percentage of victims who are white; the  $y$ -axis has a label indicating the probability of death penalty imposition. This probability generally increases along with the percentage of victims that are white. Two separate lines are given in the graph reflecting this increase, one for black defendants and one for white defendants. Note that the line for the black defendant lies wholly above that for the white defendant, implying that irrespective of the percentage of victims that may be white, the imposition of the death penalty has a greater probability for a black defendant compared to a white defendant.

The reversal paradox of having a higher death penalty imposition for whites (of 12%) compared to blacks (of 10%) in the  $2 \times 2$  contingency table aggregated over the race of the victim, is represented by two vertical lines in the graphs. Because black defendants have 38% of their victims being white, the vertical line from the  $x$ -axis value of 38% intersects the black defendant line at 10%; similarly, because white defendants have 94% of their victims being white, the vertical line from the  $x$ -axis value of 94% intersects the white defendant line at (a higher value of) 12%. The reversal occurs because there is a much greater percentage of white victims for white defendants than for black defendants. (The two lines in the graph can be constructed readily by noting how the endpoints were obtained of 0% and 6%, and of 13% and 17%. When the percentage of white victims along the  $x$ -axis is 0%, that is the same as having a black victim [which immediately generates the graph values of 0% and 6%]; if the percent-

Figure 1: Graphical representation for the Florida death penalty data.



age of white victims is 100%, this is equivalent to the victim being white [and again, immediately provides the other two endpoints of 13% and 17%]).

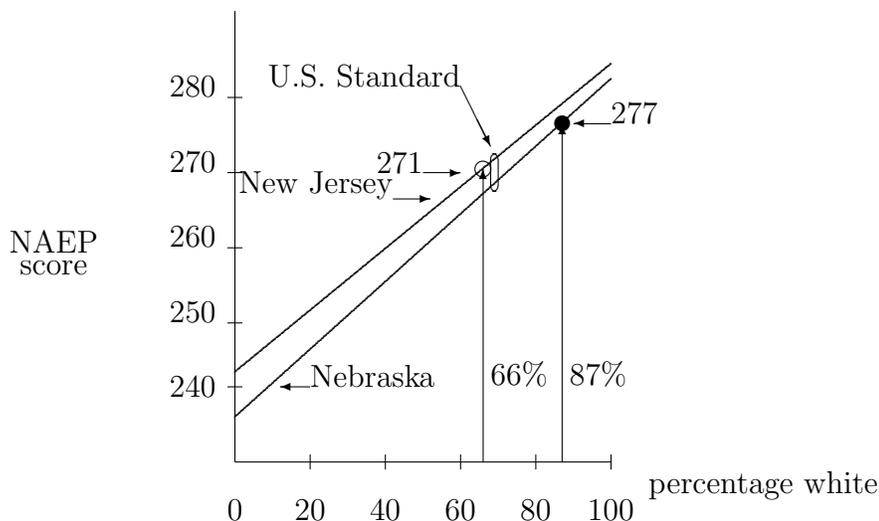
We conclude with yet another example of Simpson’s paradox (taken from Wainer, 2005, pp. 63–67) and a solution called standardization that makes the paradox disappear. Consider the results from the National Assessment of Educational Progress (NAEP) shown in Table 1. The 8th grade students in Nebraska scored 6 points higher in mathematics than their counterparts in New Jersey. White students do better in New Jersey, and so do black students; in fact, all students do better in New Jersey. How is this possible? Again, this is an example of Simpson’s paradox. Because a much greater proportion of Nebraska’s 8th grade students (87%) are from the higher scoring white population than in New Jersey (66%), their scores contribute more to the total.

Is ranking states on such an overall score sensible? It depends

on the question that these scores are being used to answer. If the question is “I want to open a business. In which state will I find a higher proportion of high-scoring math students to hire?”, the unadjusted score is sensible. If, however, the question of interest is: “I want to enroll my children in school. In which state are they likely to do better in math?”, a different answer is required. Irrespective of race, children are more likely to do better in New Jersey. When questions of this latter type are asked more frequently, it makes sense to adjust the total to reflect the correct answer. One way to do this is through the method of standardization, where each state’s score is based upon a common demographic mixture. In this instance, a sensible mixture to use is that of the nation overall. After standardization, the result obtained is the score we would expect each state to have if it had the same demographic mix as the nation. When this is done, New Jersey’s score is not affected much (273 instead of 271), but Nebraska’s score shrinks substantially (271 instead of 277).

Although Simpson’s paradox is subtle, experience has taught us that a graphic depiction often aids understanding. A graphic representation of Simpson’s paradox was provided by Baker and Kramer in 2001. Consider the graphic representation of the results from this table shown in Figure 2. A solid diagonal line shows the average NAEP math score for various proportions of white examinees in Nebraska. At the extreme left, if no whites took the test, the mean score would be that for nonwhites, 236. At the extreme right is what the mean score would be if only whites took the test, 281. The large black dot labeled “277” represents the observed score for the mixture that includes 87% whites. A second solid line above the one for Nebraska shows the same thing for New Jersey; the large open dot

Figure 2: A Baker–Kramer plot of the New Jersey–Nebraska average 8th grade National Assessment of Educational Progress (NAEP) mathematics scores.



labeled “271” denotes the score for a mixture in which 66% of those tested were white.

We see that for any fixed percentage of whites on the horizontal axis, the advantage of New Jersey over Nebraska is the same, two NAEP points. But because Nebraska has a much larger proportion of higher scoring white examinees, its mean score is higher than that of New Jersey. The small vertical box marks the percentage mixture representing the United States as a whole, and hence, encloses the standardized values. The graph makes clear how and why standardization works; it uses the same location on the horizontal axis for all groups being compared.

Simpson’s paradox generally occurs when data are aggregated. If data are collapsed across a subclassification (such as grades, race, or age), the overall difference observed might not represent what is

Table 1: National Assessment of Educational Progress (NAEP) 1992 8th grade mathematics scores.

	State	White	Black	Other Non-White	Stand- ardized
Nebraska	277	281	236	259	271
New Jersey	271	283	242	260	273
	% Population				
Nebraska		87%	5%	8%	
New Jersey		66%	15%	19%	
Nation		69%	16%	15%	

really occurring. Standardization can help correct this, but nothing will prevent the possibility of yet another subclassification, as yet unidentified, from changing things around. We believe, however, that knowing of the possibility helps contain the enthusiasm for what may be overly impulsive first inferences.<sup>3</sup>

Although Simpson’s paradox has been known by this name only rather recently (as coined by Colin Blyth in 1972), the phenomenon has been recognized and discussed for well over a hundred years; in fact, it has a complete textbook development in Yule’s *An Introduction to the Theory of Statistics*, first published in 1911.

In honor of Yule’s early contribution, we sometimes see the title of the Yule–Simpson effect. But most often, Stigler’s Law of Eponymy

---

<sup>3</sup>Fienberg (1988, p. 40) discusses an interesting example of Simpson’s paradox as it occurred in a court case involving alleged racial employment discrimination in the receipt of promotions. In this instance, blacks were being “under-promoted” in virtually every pay grade, but because of the differing numbers of blacks and whites in the various grades, blacks appeared to be “over-promoted” in the aggregate. As always, before an overall conclusion is reached based on data that have been aggregated over a variable (such as pay grade), it is always wise to “look under the hood.”

is operative (that is, “every scientific discovery is named after the last individual too ungenerous to give due credit to his predecessors.”), and Simpson is given sole naming credit for the phenomenon.<sup>4</sup>

## **2 Statistical Sleuthing and the Imposition of the Death Penalty: *McCleskey v. Kemp* (1987)**

The United States has had a troubled history with the imposition of the death penalty. Two amendments to the Constitution, the Eighth and the Fourteenth, operate as controlling guidelines for how death penalties are to be decided on and administered (if at all). The Eighth Amendment prevents “cruel and unusual punishment”; the Fourteenth Amendment contains the famous “equal protection” clause:

No State shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.

Various Supreme Court rulings over the years have relied on the Eighth Amendment to forbid some punishments entirely and to exclude others that are excessive in relation to the crime or the competence of the defendant. One of the more famous such rulings was in *Furman v. Georgia* (1972), which held that an arbitrary and inconsistent imposition of the death penalty violates both the Eighth and Fourteenth Amendments, and constitutes cruel and unusual pun-

---

<sup>4</sup>To get a better sense of the ubiquity of Simpson’s paradox in day-to-day reporting of economic statistics, see the article by Cari Tuna, *Wall Street Journal* (December 2, 2009), “When Combined Data Reveal the Flaw of Averages.”

ishment. This ruling led to a moratorium on capital punishment throughout the United States that extended to 1976 when another Georgia case was decided in *Gregg v. Georgia* (1976).

Although no majority opinion was actually written in the 5 to 4 decision in *Furman v. Georgia*, Justice Brennan writing separately in concurrence noted that

There are, then, four principles by which we may determine whether a particular punishment is ‘cruel and unusual’ . . . [the] essential predicate [is] that a punishment must not by its severity be degrading to human dignity . . . a severe punishment that is obviously inflicted in wholly arbitrary fashion . . . a severe punishment that is clearly and totally rejected throughout society . . . a severe punishment that is patently unnecessary.

Brennan went on to write that he expected that no state would pass laws obviously violating any one of these principles; and that court decisions involving the Eighth Amendment would use a “cumulative” analysis of the implication of each of the four principles.

The Supreme Court case of *Gregg v. Georgia* reaffirmed the use of the death penalty in the United States. It held that the imposition of the death penalty does not automatically violate the Eighth and Fourteenth Amendments. If the jury is furnished with standards to direct and limit the sentencing discretion, and the jury’s decision is subjected to meaningful appellate review, the death sentence may be constitutional. If, however, the death penalty is mandatory, so there is no provision for mercy based on the characteristics of the offender, then it is unconstitutional.

This short background on *Furman v. Georgia* and *Gregg v. Georgia* brings us to the case of *McCleskey v. Kemp* (1987), of primary interest in this section. For us, the main importance of *McCleskey v. Kemp* is the use and subsequent complete disregard of a monu-

mental statistical study by David C. Baldus, Charles Pulaski, and George G. Woodworth, “Comparative Review of Death Sentences: An Empirical Study of the Georgia Experience” (*Journal of Criminal Law and Criminology*, 1983, 74, 661–753). For a book length and extended version of this article, including an explicit discussion of *McCleskey v. Kemp*, see *Equal Justice and the Death Penalty: A Legal and Empirical Analysis*. David C. Baldus, George Woodworth, and Charles A. Pulaski, Jr., Boston: Northeastern University Press, 1990.

There are many analyses done by Baldus et al. and others on the interrelation between the race of the victim and of the defendant and the imposition of the death penalty. Most do not show an explicit Simpson’s paradox such as for the Radelet data of the last section, where a black defendant has a higher probability of receiving the death penalty compared to a white defendant. But universally, the race of the victim plays a crucial part in death penalty imposition – when the victim is white, the probability of receiving the death penalty is substantially higher than for black victims. The relative risks, for example, are all much greater than the value of 2.0 needed to legally assert specific causation.

In *McCleskey v. Kemp*, the Court held that despite statistical evidence of a profound racial disparity in application of the death penalty, such evidence is insufficient to invalidate a defendant’s death sentence. The syllabus of this ruling is given below. To see additional contemporary commentary, an article by Anthony Lewis lamenting this ruling appeared in the *New York Times* (April 28, 1987), entitled “Bowling To Racism.”

## 2.1 United States Supreme Court, *McCleskey v. Kemp* (1987): Syllabus

In 1978, petitioner, a black man, was convicted in a Georgia trial court of armed robbery and murder, arising from the killing of a white police officer during the robbery of a store. Pursuant to Georgia statutes, the jury at the penalty hearing considered the mitigating and aggravating circumstances of petitioner's conduct, and recommended the death penalty on the murder charge. The trial court followed the recommendation, and the Georgia Supreme Court affirmed. After unsuccessfully seeking post-conviction relief in state courts, petitioner sought habeas corpus relief in Federal District Court. His petition included a claim that the Georgia capital sentencing process was administered in a racially discriminatory manner in violation of the Eighth and Fourteenth Amendments. In support of the claim, petitioner proffered a statistical study (the Baldus study) that purports to show a disparity in the imposition of the death sentence in Georgia based on the murder victim's race and, to a lesser extent, the defendant's race. The study is based on over 2,000 murder cases that occurred in Georgia during the 1970's, and involves data relating to the victim's race, the defendant's race, and the various combinations of such persons' races. The study indicates that black defendants who killed white victims have the greatest likelihood of receiving the death penalty. Rejecting petitioner's constitutional claims, the court denied his petition insofar as it was based on the Baldus study, and the Court of Appeals affirmed the District Court's decision on this issue. It assumed the validity of the Baldus study, but found the statistics insufficient to demonstrate unconstitutional discrimination in the Fourteenth Amendment context or to show irrationality, arbitrariness, and capriciousness under Eighth Amendment analysis.

Held:

1. The Baldus study does not establish that the administration of the Georgia capital punishment system violates the Equal Protection Clause.

- (a) To prevail under that Clause, petitioner must prove that the decision makers in his case acted with discriminatory purpose. Petitioner offered no evidence specific to his own case that would support an inference that

racial considerations played a part in his sentence, and the Baldus study is insufficient to support an inference that any of the decision makers in his case acted with discriminatory purpose. This Court has accepted statistics as proof of intent to discriminate in the context of a State's selection of the jury venire, and in the context of statutory violations under Title VII of the Civil Rights Act of 1964. However, the nature of the capital sentencing decision and the relationship of the statistics to that decision are fundamentally different from the corresponding elements in the venire selection or Title VII cases. Petitioner's statistical proffer must be viewed in the context of his challenge to decisions at the heart of the State's criminal justice system. Because discretion is essential to the criminal justice process, exceptionally clear proof is required before this Court will infer that the discretion has been abused.

(b) There is no merit to petitioner's argument that the Baldus study proves that the State has violated the Equal Protection Clause by adopting the capital punishment statute and allowing it to remain in force despite its allegedly discriminatory application. For this claim to prevail, petitioner would have to prove that the Georgia Legislature enacted or maintained the death penalty statute because of an anticipated racially discriminatory effect. There is no evidence that the legislature either enacted the statute to further a racially discriminatory purpose or maintained the statute because of the racially disproportionate impact suggested by the Baldus study.

2. Petitioner's argument that the Baldus study demonstrates that the Georgia capital sentencing system violates the Eighth Amendment's prohibition of cruel and unusual punishment must be analyzed in the light of this Court's prior decisions under that Amendment. Decisions since *Furman v. Georgia*, have identified a constitutionally permissible range of discretion in imposing the death penalty. First, there is a required threshold below which the death penalty cannot be imposed, and the State must establish rational criteria that narrow the decision-maker's judgment as to whether the circumstances of a particular defendant's case meet the threshold. Second, States cannot limit the sentencer's consideration of any relevant circumstance that could cause it to decline to impose the death penalty. In this respect, the State cannot channel the sentencer's discretion, but must allow it to consider any relevant information offered by the defendant.

3. The Baldus study does not demonstrate that the Georgia capital sentencing system violates the Eighth Amendment.

(a) Petitioner cannot successfully argue that the sentence in his case is disproportionate to the sentences in other murder cases. On the one hand, he cannot base a constitutional claim on an argument that his case differs from other cases in which defendants did receive the death penalty. The Georgia Supreme Court found that his death sentence was not disproportionate to other death sentences imposed in the State. On the other hand, absent a showing that the Georgia capital punishment system operates in an arbitrary and capricious manner, petitioner cannot prove a constitutional violation by demonstrating that other defendants who may be similarly situated did not receive the death penalty. The opportunities for discretionary leniency under state law do not render the capital sentences imposed arbitrary and capricious. Because petitioner's sentence was imposed under Georgia sentencing procedures that focus discretion "on the particularized nature of the crime and the particularized characteristics of the individual defendant," it may be presumed that his death sentence was not "wantonly and freakishly" imposed, and thus that the sentence is not disproportionate within any recognized meaning under the Eighth Amendment.

(b) There is no merit to the contention that the Baldus study shows that Georgia's capital punishment system is arbitrary and capricious in application. The statistics do not prove that race enters into any capital sentencing decisions or that race was a factor in petitioner's case. The likelihood of racial prejudice allegedly shown by the study does not constitute the constitutional measure of an unacceptable risk of racial prejudice. The inherent lack of predictability of jury decisions does not justify their condemnation. On the contrary, it is the jury's function to make the difficult and uniquely human judgments that defy codification and that build discretion, equity, and flexibility into the legal system.

(c) At most, the Baldus study indicates a discrepancy that appears to correlate with race, but this discrepancy does not constitute a major systemic defect. Any mode for determining guilt or punishment has its weaknesses and the potential for misuse. Despite such imperfections, constitutional guarantees are met when the mode for determining guilt or punishment has been

surrounded with safeguards to make it as fair as possible.

4. Petitioner's claim, taken to its logical conclusion, throws into serious question the principles that underlie the entire criminal justice system. His claim easily could be extended to apply to other types of penalties and to claims based on unexplained discrepancies correlating to membership in other minority groups and even to gender. The Constitution does not require that a State eliminate any demonstrable disparity that correlates with a potentially irrelevant factor in order to operate a criminal justice system that includes capital punishment. Petitioner's arguments are best presented to the legislative bodies, not the courts.

We make a number of comments about the majority opinion in *McCleskey v. Kemp* just summarized in the syllabus and noted in the article by Anthony Lewis. First, it is rarely the case that a policy could be identified as the cause for an occurrence in one specific individual. The legal system in its dealings with epidemiology and toxicology has generally recognized that an agent can never be said to have been the specific cause of, say, a disease in a particular individual. This is the notion of specific causation, which is typically unprovable. As an alternative approach to causation, courts have commonly adopted a criterion of general causation defined by relative risk being greater than 2.0 (as discussed in Module 1) to infer that a toxic agent was more likely than not the cause of a specific person's disease (and thus open to compensation).<sup>5</sup> To require that a defendant prove that the decision makers in his particular case acted with discriminatory malice is to set an unreachable standard. So is an expectation that statistics could ever absolutely prove "that

---

<sup>5</sup>In his dissent, Justice Brennan makes this exact point when he states: "For this reason, we have demanded a uniquely high degree of rationality in imposing the death penalty. A capital sentencing system in which race more likely than not plays a role does not meet this standard."

race enters into any capital sentencing decisions or that race was a factor in petitioner's case." Statistical sleuthing can at best identify anomalies that need further study; but irrespective, the anomalies cannot be just willed away as if they never existed.

The statement that "petitioner cannot successfully argue that the sentence in his case is disproportionate to the sentences in other murder cases" again assigns an impossible personal standard. It will always be impossible to define unequivocally what the "comparables" are that might be used in such comparisons. The operation of confirmation biases would soon overwhelm any attempt to define a set of comparables. Even realtors have huge difficulties in assigning comparable sales to a given property when deciding on an asking or selling price. Usually, realtors just fall back on a simple linear rule of dollars per square foot. But unfortunately, nothing so simple exists in defining comparables in imposing (or not) death sentences in Georgia.

If it can be shown that an enacted (legislative or legal) policy has the effect of denying constitutional rights for an identifiable group of individuals, then that policy should be declared discriminatory and changed. It should never be necessary to show that the enactors of such a policy consciously meant for that effect to occur—the law of unintended consequences again rears its ugly head—or that in one specific case it was operative. When policies must be carried out through human judgment, any number of subjective biases may be present at any given moment, and without any possibility of identifying which ones are at work and which ones are not.

In various places throughout the majority opinion, there appears to be argument by sheer assertion with no other supporting evidence

at all. We all need to repeat to ourselves the admonition that just saying so doesn't necessarily make it so. Thus, we have the admission that there appears to be discriminatory effects correlated with race, with the empty assertion that "this discrepancy does not constitute a major systemic weakness" or "despite such imperfection, constitutional guarantees are met." To us, this seems like nonsense, pure and simple.

The final point in the syllabus is that "if the Petitioner's claim is taken to its logical conclusion, questions arise about the principles underlying the entire criminal justice system." Or in Justice Brennan's dissent, the majority opinion is worried about "too much justice." God forbid that other anomalies be identified that correlate with membership in other groups (for example, sex, age, other minorities) that would then have to be dealt with.

The *New York Review of Books* in its December 23, 2010 issue scored a coup by having a lead article entitled "On the Death Sentence," by retired Supreme Court Justice John Paul Stevens. Stevens was reviewing the book, *Peculiar Institution: America's Death Penalty in an Age of Abolition* (by David Garland). In the course of his essay, Stevens comments on *McCleskey v. Kemp* and notes that Justice Powell (who wrote the majority opinion) in remarks he made to his biographer, said that he should have voted the other way in the *McCleskey* 5 to 4 decision. It's too bad we cannot retroactively reverse Supreme Court rulings, particularly given the doctrine of *stare decisis*, according to which judges are obliged to respect the precedents set by prior decisions. The doctrine of *stare decisis* suggests that no amount of statistical evidence will ever be sufficient to declare the death penalty in violation of the "equal protection"

clause of the Fourteenth Amendment. The relevant quotation from the Stevens review follows:

In 1987, the Court held in *McCleskey v. Kemp* that it did not violate the Constitution for a state to administer a criminal justice system under which murderers of victims of one race received death sentences much more frequently than murderers of victims of another race. The case involved a study by Iowa law professor David Baldus and his colleagues demonstrating that in Georgia murderers of white victims were eleven times more likely to be sentenced to death than were murderers of black victims. Controlling for race-neutral factors and focusing solely on decisions by prosecutors about whether to seek the death penalty, Justice Blackmun observed in dissent, the effect of race remained “readily identifiable” and “statistically significant” across a sample of 2,484 cases.

That the murder of black victims is treated as less culpable than the murder of white victims provides a haunting reminder of once-prevalent Southern lynchings. Justice Stewart, had he remained on the Court, surely would have voted with the four dissenters. That conclusion is reinforced by Justice Powell’s second thoughts; he later told his biographer that he regretted his vote in *McCleskey*.

We give redactions of the majority opinion and dissent in an appendix (by Justice Brennan) for *McCleskey v. Kemp*. It is a pity that Brennan’s dissent did not form the majority opinion as it would have but for Justice Powell’s vote that in hindsight he wished he could change. It also would have given greater legitimacy and importance to such landmark statistical studies as done by Baldus, et al. (1983). We will leave readers to peruse the majority and dissenting opinions and arrive at their own identification of outrageous argumentation on either side. In reading the majority and dissenting opinions, it is best to keep in mind the word “opinion.” Such opinions include disregarding incontrovertible statistical evidence that something is amiss in the administration of the Georgia death penalty, wherever

that may arise from. Although the cause may be ambiguous, there is no doubt that it results from all the various actors in the legal system who make the series of decisions necessary in determining who lives and who dies.

## References

- [1] Baker, S. G., & Kramer, B. S. (2001). Good for women, good for men, bad for people: Simpson's paradox and the importance of sex-specific analysis in observational studies. *Journal of Women's Health & Gender-Based Medicine*, *10*, 867–872.
- [2] Bickel, P. J., Hammel, E. A., & O'Connell, J. W. (1975). Sex bias in graduate admissions: Data from Berkeley. *Science*, *187*, 398–404.
- [3] Blyth, C. R. (1972). On Simpson's paradox and the sure-thing principle. *Journal of the American Statistical Association*, *67*, 364–366.
- [4] Fienberg, S. E. (Ed.) (1988). *The evolving role of statistical assessments as evidence in the courts*. New York: Springer-Verlag.
- [5] Radelet, M. L. (1981). Racial characteristics and the composition of the death penalty. *American Sociological Review*, *46*, 918–927.
- [6] Wainer, H. (2005). *Graphic discovery: A trout in the milk and other visual adventures*. Princeton, N.J.: Princeton University Press.

### **3 Appendix: United States Supreme Court, McCleskey v. Kemp (1987): Majority Opinion and Dissent**

Justice Powell delivered the opinion of the Court.

This case presents the question whether a complex statistical study that indicates a risk that racial considerations enter into capital sentencing determinations proves that petitioner McCleskey's capital sentence is unconstitutional under the Eighth or Fourteenth Amendment.

...

McCleskey next filed a petition for a writ of habeas corpus in the Federal District Court for the Northern District of Georgia. His petition raised 18 claims, one of which was that the Georgia capital sentencing process is administered in a racially discriminatory manner in violation of the Eighth and Fourteenth Amendments to the United States Constitution. In support of his claim, McCleskey proffered a statistical study performed by Professors David C. Baldus, Charles Pulaski, and George Woodworth (the Baldus study) that purports to show a disparity in the imposition of the death sentence in Georgia based on the race of the murder victim and, to a lesser extent, the race of the defendant. The Baldus study is actually two sophisticated statistical studies that examine over 2,000 murder cases that occurred in Georgia during the 1970's. The raw numbers collected by Professor Baldus indicate that defendants charged with killing white persons received the death penalty in 11% of the cases, but defendants charged with killing blacks received the death penalty in only 1% of the cases. The raw numbers also indicate a reverse racial disparity according to the race of the defendant: 4% of the black defendants received the death penalty, as opposed to 7% of the white defendants.

Baldus also divided the cases according to the combination of the race of the defendant and the race of the victim. He found that the death penalty was assessed in 22% of the cases involving black defendants and white victims; 8% of the cases involving white defendants and white victims; 1% of the cases involving black defendants and black victims; and 3% of the cases involving white defendants and black victims. Similarly, Baldus found that prosecutors sought the death penalty in 70% of the cases involving black defendants and

white victims; 32% of the cases involving white defendants and white victims; 15% of the cases involving black defendants and black victims; and 19% of the cases involving white defendants and black victims.

Baldus subjected his data to an extensive analysis, taking account of 230 variables that could have explained the disparities on nonracial grounds. One of his models concludes that, even after taking account of 39 nonracial variables, defendants charged with killing white victims were 4.3 times as likely to receive a death sentence as defendants charged with killing blacks. According to this model, black defendants were 1.1 times as likely to receive a death sentence as other defendants. Thus, the Baldus study indicates that black defendants, such as McCleskey, who kill white victims have the greatest likelihood of receiving the death penalty.

The District Court held an extensive evidentiary hearing on McCleskey's petition. . . . It concluded that McCleskey's statistics do not demonstrate a prima facie case in support of the contention that the death penalty was imposed upon him because of his race, because of the race of the victim, or because of any Eighth Amendment concern.

As to McCleskey's Fourteenth Amendment claim, the court found that the methodology of the Baldus study was flawed in several respects. Because of these defects, the court held that the Baldus study "fail[ed] to contribute anything of value" to McCleskey's claim. Accordingly, the court denied the petition insofar as it was based upon the Baldus study.<sup>6</sup>

---

<sup>6</sup>Baldus, among other experts, testified at the evidentiary hearing. The District Court "was impressed with the learning of all of the experts." Nevertheless, the District Court noted that, in many respects, the data were incomplete. In its view, the questionnaires used to obtain the data failed to capture the full degree of the aggravating or mitigating circumstances. The court criticized the researcher's decisions regarding unknown variables. The researchers could not discover whether penalty trials were held in many of the cases, thus undercutting the value of the study's statistics as to prosecutorial decisions. In certain cases, the study lacked information on the race of the victim in cases involving multiple victims, on whether or not the prosecutor offered a plea bargain, and on credibility problems with witnesses. The court concluded that McCleskey had failed to establish by a preponderance of the

The Court of Appeals for the Eleventh Circuit, sitting en banc, carefully reviewed the District Court's decision on McCleskey's claim. It assumed the validity of the study itself, and addressed the merits of McCleskey's Eighth and Fourteenth Amendment claims. That is, the court assumed that the study showed that systematic and substantial disparities existed in the penalties imposed upon homicide defendants in Georgia based on race of the homicide victim, that the disparities existed at a less substantial rate in death sentencing based on race of defendants, and that the factors of race of the victim and defendant were at work in Fulton County.

Even assuming the study's validity, the Court of Appeals found the statistical evidence that the data were trustworthy.

It is a major premise of a statistical case that the database numerically mirrors reality. If it does not in substantial degree mirror reality, any inferences empirically arrived at are untrustworthy.

The District Court noted other problems with Baldus' methodology. First, the researchers assumed that all of the information available from the questionnaires was available to the juries and prosecutors when the case was tried. The court found this assumption "questionable." Second, the court noted the instability of the various models. Even with the 230-variable model, consideration of 20 further variables caused a significant drop in the statistical significance of race. In the court's view, this undermined the persuasiveness of the model that showed the greatest racial disparity, the 39-variable model. Third, the court found that the high correlation between race and many of the nonracial variables diminished the weight to which the study was entitled.

Finally, the District Court noted the inability of any of the models to predict the outcome of actual cases. As the court explained, statisticians use a measure called an "r-squared" to measure what portion of the variance in the dependent variable (death sentencing rate, in this case) is accounted for by the independent variables of the model. A perfectly predictive model would have an r-squared value of 1.0. A model with no predictive power would have an r-squared value of 0. The r-squared value of Baldus' most complex model, the 230-variable model, was between .46 and .48. Thus, as the court explained, "the 230-variable model does not predict the outcome in half of the cases."

tics insufficient to demonstrate discriminatory intent or unconstitutional discrimination in the Fourteenth Amendment context, [and] insufficient to show irrationality, arbitrariness and capriciousness under any kind of Eighth Amendment analysis.

The court noted:

The very exercise of discretion means that persons exercising discretion may reach different results from exact duplicates. Assuming each result is within the range of discretion, all are correct in the eyes of the law. It would not make sense for the system to require the exercise of discretion in order to be facially constitutional, and at the same time hold a system unconstitutional in application where that discretion achieved different results for what appear to be exact duplicates, absent the state showing the reasons for the difference.

The Baldus approach ... would take the cases with different results on what are contended to be duplicate facts, where the differences could not be otherwise explained, and conclude that the different result was based on race alone. ... This approach ignores the realities. ... There are, in fact, no exact duplicates in capital crimes and capital defendants. The type of research submitted here tends to show which of the directed factors were effective, but is of restricted use in showing what undirected factors control the exercise of constitutionally required discretion.

The court concluded:

Viewed broadly, it would seem that the statistical evidence presented here, assuming its validity, confirms, rather than condemns, the system. ... The marginal disparity based on the race of the victim tends to support the state's contention that the system is working far differently from the one which *Furman v. Georgia*, condemned. In pre-*Furman* days, there was no rhyme or reason as to who got the death penalty and who did not. But now, in the vast majority of cases, the reasons for a difference are well documented. That they are not so clear in a small percentage of the cases is no reason to declare the entire system unconstitutional.

The Court of Appeals affirmed the denial by the District Court of McCleskey's petition for a writ of habeas corpus insofar as the petition was based upon the Baldus study, with three judges dissenting as to McCleskey's

claims based on the Baldus study. We granted certiorari, and now affirm.

...

McCleskey's first claim is that the Georgia capital punishment statute violates the Equal Protection Clause of the Fourteenth Amendment.<sup>7</sup> He argues that race has infected the administration of Georgia's statute in two ways: persons who murder whites are more likely to be sentenced to death than persons who murder blacks, and black murderers are more likely to be sentenced to death than white murderers. As a black defendant who killed a white victim, McCleskey claims that the Baldus study demonstrates that he was discriminated against because of his race and because of the race of his victim. In its broadest form, McCleskey's claim of discrimination extends to every actor in the Georgia capital sentencing process, from the prosecutor who sought the death penalty and the jury that imposed the sentence to the State itself that enacted the capital punishment statute and allows it to remain in effect despite its allegedly discriminatory application. We agree with the Court of Appeals, and every other court that has considered such a challenge, that this claim must fail.

Our analysis begins with the basic principle that a defendant who alleges an equal protection violation has the burden of proving "the existence of purposeful discrimination." A corollary to this principle is that a criminal defendant must prove that the purposeful discrimination "had a discriminatory effect" on him. Thus, to prevail under the Equal Protection Clause, McCleskey must prove that the decision-makers in his case acted with dis-

---

<sup>7</sup>Although the District Court rejected the findings of the Baldus study as flawed, the Court of Appeals assumed that the study is valid, and reached the constitutional issues. Accordingly, those issues are before us. As did the Court of Appeals, we assume the study is valid statistically, without reviewing the factual findings of the District Court. Our assumption that the Baldus study is statistically valid does not include the assumption that the study shows that racial considerations actually enter into any sentencing decisions in Georgia. Even a sophisticated multiple-regression analysis such as the Baldus study can only demonstrate a risk that the factor of race entered into some capital sentencing decisions, and a necessarily lesser risk that race entered into any particular sentencing decision.

criminy purpose. He offers no evidence specific to his own case that would support an inference that racial considerations played a part in his sentence. Instead, he relies solely on the Baldus study. McCleskey argues that the Baldus study compels an inference that his sentence rests on purposeful discrimination. McCleskey's claim that these statistics are sufficient proof of discrimination, without regard to the facts of a particular case, would extend to all capital cases in Georgia, at least where the victim was white and the defendant is black.

The Court has accepted statistics as proof of intent to discriminate in certain limited contexts. First, this Court has accepted statistical disparities as proof of an equal protection violation in the selection of the jury venire in a particular district. Although statistical proof normally must present a "stark" pattern to be accepted as the sole proof of discriminatory intent under the Constitution, [b]ecause of the nature of the jury-selection task, . . . we have permitted a finding of constitutional violation even when the statistical pattern does not approach [such] extremes.

Second, this Court has accepted statistics in the form of multiple-regression analysis to prove statutory violations under Title VII of the Civil Rights Act of 1964.

But the nature of the capital sentencing decision, and the relationship of the statistics to that decision, are fundamentally different from the corresponding elements in the venire selection or Title VII cases. Most importantly, each particular decision to impose the death penalty is made by a petit jury selected from a properly constituted venire. Each jury is unique in its composition, and the Constitution requires that its decision rest on consideration of innumerable factors that vary according to the characteristics of the individual defendant and the facts of the particular capital offense. Thus, the application of an inference drawn from the general statistics to a specific decision in a trial and sentencing simply is not comparable to the application of an inference drawn from general statistics to a specific venire-selection or Title VII case. In those cases, the statistics relate to fewer entities, and fewer variables are relevant to the challenged decisions.

Another important difference between the cases in which we have accepted statistics as proof of discriminatory intent and this case is that, in the venire-

selection and Title VII contexts, the decision-maker has an opportunity to explain the statistical disparity. Here, the State has no practical opportunity to rebut the Baldus study. “[C]ontrolling considerations of . . . public policy,” dictate that jurors “cannot be called . . . to testify to the motives and influences that led to their verdict.” Similarly, the policy considerations behind a prosecutor’s traditionally “wide discretion” suggest the impropriety of our requiring prosecutors to defend their decisions to seek death penalties, “often years after they were made.” Moreover, absent far stronger proof, it is unnecessary to seek such a rebuttal, because a legitimate and unchallenged explanation for the decision is apparent from the record: McCleskey committed an act for which the United States Constitution and Georgia laws permit imposition of the death penalty.

Finally, McCleskey’s statistical proffer must be viewed in the context of his challenge. McCleskey challenges decisions at the heart of the State’s criminal justice system.

[O]ne of society’s most basic tasks is that of protecting the lives of its citizens, and one of the most basic ways in which it achieves the task is through criminal laws against murder.

Implementation of these laws necessarily requires discretionary judgments. Because discretion is essential to the criminal justice process, we would demand exceptionally clear proof before we would infer that the discretion has been abused. The unique nature of the decisions at issue in this case also counsels against adopting such an inference from the disparities indicated by the Baldus study. Accordingly, we hold that the Baldus study is clearly insufficient to support an inference that any of the decision-makers in McCleskey’s case acted with discriminatory purpose.

...

McCleskey also suggests that the Baldus study proves that the State as a whole has acted with a discriminatory purpose. He appears to argue that the State has violated the Equal Protection Clause by adopting the capital punishment statute and allowing it to remain in force despite its allegedly discriminatory application. But “[d]iscriminatory purpose” . . . implies more than intent as volition or intent as awareness of consequences. It implies that the decision-maker, in this case a state legislature, selected or reaffirmed a

particular course of action at least in part “because of,” not merely “in spite of,” its adverse effects upon an identifiable group.

For this claim to prevail, McCleskey would have to prove that the Georgia Legislature enacted or maintained the death penalty statute because of an anticipated racially discriminatory effect. In *Gregg v. Georgia*, this Court found that the Georgia capital sentencing system could operate in a fair and neutral manner. There was no evidence then, and there is none now, that the Georgia Legislature enacted the capital punishment statute to further a racially discriminatory purpose. Nor has McCleskey demonstrated that the legislature maintains the capital punishment statute because of the racially disproportionate impact suggested by the Baldus study. As legislatures necessarily have wide discretion in the choice of criminal laws and penalties, and as there were legitimate reasons for the Georgia Legislature to adopt and maintain capital punishment, we will not infer a discriminatory purpose on the part of the State of Georgia. Accordingly, we reject McCleskey’s equal protection claims.

...

Although our decision in *Gregg* as to the facial validity of the Georgia capital punishment statute appears to foreclose McCleskey’s disproportionality argument, he further contends that the Georgia capital punishment system is arbitrary and capricious in application, and therefore his sentence is excessive, because racial considerations may influence capital sentencing decisions in Georgia. We now address this claim.

To evaluate McCleskey’s challenge, we must examine exactly what the Baldus study may show. Even Professor Baldus does not contend that his statistics prove that race enters into any capital sentencing decisions, or that race was a factor in McCleskey’s particular case. Statistics, at most, may show only a likelihood that a particular factor entered into some decisions. There is, of course, some risk of racial prejudice influencing a jury’s decision in a criminal case. There are similar risks that other kinds of prejudice will influence other criminal trials. The question “is at what point that risk becomes constitutionally unacceptable,” McCleskey asks us to accept the likelihood allegedly shown by the Baldus study as the constitutional measure of an unacceptable risk of racial prejudice influencing capital sentencing de-

cisions. This we decline to do. Because of the risk that the factor of race may enter the criminal justice process, we have engaged in “unceasing efforts” to eradicate racial prejudice from our criminal justice system. Our efforts have been guided by our recognition that the inestimable privilege of trial by jury . . . is a vital principle, underlying the whole administration of criminal justice system. Thus, it is the jury that is a criminal defendant’s fundamental “protection of life and liberty against race or color prejudice.” Specifically, a capital sentencing jury representative of a criminal defendant’s community assures a ‘diffused impartiality,’ in the jury’s task of “express[ing] the conscience of the community on the ultimate question of life or death.

Individual jurors bring to their deliberations “qualities of human nature and varieties of human experience, the range of which is unknown and perhaps unknowable.” The capital sentencing decision requires the individual jurors to focus their collective judgment on the unique characteristics of a particular criminal defendant. It is not surprising that such collective judgments often are difficult to explain. But the inherent lack of predictability of jury decisions does not justify their condemnation. On the contrary, it is the jury’s function to make the difficult and uniquely human judgments that defy codification, and that “buil[d] discretion, equity, and flexibility into a legal system.”

McCleskey’s argument that the Constitution condemns the discretion allowed decision-makers in the Georgia capital sentencing system is antithetical to the fundamental role of discretion in our criminal justice system. Discretion in the criminal justice system offers substantial benefits to the criminal defendant. Not only can a jury decline to impose the death sentence, it can decline to convict or choose to convict of a lesser offense. Whereas decisions against a defendant’s interest may be reversed by the trial judge or on appeal, these discretionary exercises of leniency are final and unreviewable. Similarly, the capacity of prosecutorial discretion to provide individualized justice is “only entrenched in American law.” As we have noted, a prosecutor can decline to charge, offer a plea bargain, or decline to seek a death sentence in any particular case. Of course, “the power to be lenient [also] is the power to discriminate,” but a capital punishment system that did not allow for discretionary acts of leniency “would be totally alien to our notions of criminal justice.”

...

At most, the Baldus study indicates a discrepancy that appears to correlate with race. Apparent disparities in sentencing are an inevitable part of our criminal justice system. The discrepancy indicated by the Baldus study is “a far cry from the major systemic defects identified in Furman.” As this Court has recognized, any mode for determining guilt or punishment “has its weaknesses and the potential for misuse.” Specifically, “there can be ‘no perfect procedure for deciding in which cases governmental authority should be used to impose death.’” Despite these imperfections, our consistent rule has been that constitutional guarantees are met when “the mode [for determining guilt or punishment] itself has been surrounded with safeguards to make it as fair as possible.” Where the discretion that is fundamental to our criminal process is involved, we decline to assume that what is unexplained is invidious. In light of the safeguards designed to minimize racial bias in the process, the fundamental value of jury trial in our criminal justice system, and the benefits that discretion provides to criminal defendants, we hold that the Baldus study does not demonstrate a constitutionally significant risk of racial bias affecting the Georgia capital sentencing process.

...

Two additional concerns inform our decision in this case. First, McCleskey’s claim, taken to its logical conclusion, throws into serious question the principles that underlie our entire criminal justice system. The Eighth Amendment is not limited in application to capital punishment, but applies to all penalties. Thus, if we accepted McCleskey’s claim that racial bias has impermissibly tainted the capital sentencing decision, we could soon be faced with similar claims as to other types of penalty. Moreover, the claim that his sentence rests on the irrelevant factor of race easily could be extended to apply to claims based on unexplained discrepancies that correlate to membership in other minority groups, and even to gender. Similarly, since McCleskey’s claim relates to the race of his victim, other claims could apply with equally logical force to statistical disparities that correlate with the race or sex of other actors in the criminal justice system, such as defense attorneys or judges. Also, there is no logical reason that such a claim need be limited to racial or sexual bias. If arbitrary and capricious punishment is the touchstone under the

Eighth Amendment, such a claim could—at least in theory—be based upon any arbitrary variable, such as the defendant’s facial characteristics, or the physical attractiveness of the defendant or the victim, that some statistical study indicates may be influential in jury decision-making. As these examples illustrate, there is no limiting principle to the type of challenge brought by McCleskey. The Constitution does not require that a State eliminate any demonstrable disparity that correlates with a potentially irrelevant factor in order to operate a criminal justice system that includes capital punishment. As we have stated specifically in the context of capital punishment, the Constitution does not “plac[e] totally unrealistic conditions on its use.”

Second, McCleskey’s arguments are best presented to the legislative bodies. It is not the responsibility—or indeed even the right—of this Court to determine the appropriate punishment for particular crimes. It is the legislatures, the elected representatives of the people, that are “constituted to respond to the will and consequently the moral values of the people.” Legislatures also are better qualified to weigh and evaluate the results of statistical studies in terms of their own local conditions and with a flexibility of approach that is not available to the courts,

Capital punishment is now the law in more than two-thirds of our States. It is the ultimate duty of courts to determine on a case-by-case basis whether these laws are applied consistently with the Constitution. Despite McCleskey’s wide-ranging arguments that basically challenge the validity of capital punishment in our multiracial society, the only question before us is whether, in his case, the law of Georgia was properly applied. We agree with the District Court and the Court of Appeals for the Eleventh Circuit that this was carefully and correctly done in this case.

Accordingly, we affirm the judgment of the Court of Appeals for the Eleventh Circuit.

It is so ordered.

---

Justice Brennan, Dissenting Opinion

...

At some point in this case, Warren McCleskey doubtless asked his lawyer whether a jury was likely to sentence him to die. A candid reply to this

question would have been disturbing. First, counsel would have to tell McCleskey that few of the details of the crime or of McCleskey's past criminal conduct were more important than the fact that his victim was white. Furthermore, counsel would feel bound to tell McCleskey that defendants charged with killing white victims in Georgia are 4.3 times as likely to be sentenced to death as defendants charged with killing blacks. In addition, frankness would compel the disclosure that it was more likely than not that the race of McCleskey's victim would determine whether he received a death sentence: 6 of every 11 defendants convicted of killing a white person would not have received the death penalty if their victims had been black, while, among defendants with aggravating and mitigating factors comparable to McCleskey's, 20 of every 34 would not have been sentenced to die if their victims had been black. Finally, the assessment would not be complete without the information that cases involving black defendants and white victims are more likely to result in a death sentence than cases featuring any other racial combination of defendant and victim. The story could be told in a variety of ways, but McCleskey could not fail to grasp its essential narrative line: there was a significant chance that race would play a prominent role in determining if he lived or died.

The Court today holds that Warren McCleskey's sentence was constitutionally imposed. It finds no fault in a system in which lawyers must tell their clients that race casts a large shadow on the capital sentencing process. The Court arrives at this conclusion by stating that the Baldus study cannot "prove that race enters into any capital sentencing decisions or that race was a factor in McCleskey's particular case." Since, according to Professor Baldus, we cannot say "to a moral certainty" that race influenced a decision, we can identify only "a likelihood that a particular factor entered into some decisions," and "a discrepancy that appears to correlate with race." This "likelihood" and "discrepancy," holds the Court, is insufficient to establish a constitutional violation. The Court reaches this conclusion by placing four factors on the scales opposite McCleskey's evidence: the desire to encourage sentencing discretion, the existence of "statutory safeguards" in the Georgia scheme, the fear of encouraging widespread challenges to other sentencing decisions, and the limits of the judicial role. The Court's evaluation of the

significance of petitioner’s evidence is fundamentally at odds with our consistent concern for rationality in capital sentencing, and the considerations that the majority invokes to discount that evidence cannot justify ignoring its force.

...

It is important to emphasize at the outset that the Court’s observation that McCleskey cannot prove the influence of race on any particular sentencing decision is irrelevant in evaluating his Eighth Amendment claim. Since *Furman v. Georgia*, the Court has been concerned with the risk of the imposition of an arbitrary sentence, rather than the proven fact of one. *Furman* held that the death penalty may not be imposed under sentencing procedures that create a substantial risk that the punishment will be inflicted in an arbitrary and capricious manner.

As Justice O’Connor observed in *Caldwell v. Mississippi*, a death sentence must be struck down when the circumstances under which it has been imposed creat[e] an unacceptable risk that “the death penalty [may have been] meted out arbitrarily or capriciously,” or through “whim or mistake.” This emphasis on risk acknowledges the difficulty of divining the jury’s motivation in an individual case. In addition, it reflects the fact that concern for arbitrariness focuses on the rationality of the system as a whole, and that a system that features a significant probability that sentencing decisions are influenced by impermissible considerations cannot be regarded as rational. As we said in *Gregg v. Georgia*, “the petitioner looks to the sentencing system as a whole (as the Court did in *Furman* and we do today)”: a constitutional violation is established if a plaintiff demonstrates a “pattern of arbitrary and capricious sentencing.”

As a result, our inquiry under the Eighth Amendment has not been directed to the validity of the individual sentences before us. In *Godfrey*, for instance, the Court struck down the petitioner’s sentence because the vagueness of the statutory definition of heinous crimes created a risk that prejudice or other impermissible influences might have infected the sentencing decision. In vacating the sentence, we did not ask whether it was likely that *Godfrey*’s own sentence reflected the operation of irrational considerations. Nor did we demand a demonstration that such considerations had actually entered into

other sentencing decisions involving heinous crimes. Similarly, in *Roberts v. Louisiana*, and *Woodson v. North Carolina*, we struck down death sentences in part because mandatory imposition of the death penalty created the risk that a jury might rely on arbitrary considerations in deciding which persons should be convicted of capital crimes. Such a risk would arise, we said, because of the likelihood that jurors, reluctant to impose capital punishment on a particular defendant, would refuse to return a conviction, so that the effect of mandatory sentencing would be to recreate the unbounded sentencing discretion condemned in *Furman*. We did not ask whether the death sentences in the cases before us could have reflected the jury's rational consideration and rejection of mitigating factors. Nor did we require proof that juries had actually acted irrationally in other cases.

Defendants challenging their death sentences thus never have had to prove that impermissible considerations have actually infected sentencing decisions. We have required instead that they establish that the system under which they were sentenced posed a significant risk of such an occurrence. *McCleskey's* claim does differ, however, in one respect from these earlier cases: it is the first to base a challenge not on speculation about how a system might operate, but on empirical documentation of how it does operate.

The Court assumes the statistical validity of the *Baldus* study, and acknowledges that *McCleskey* has demonstrated a risk that racial prejudice plays a role in capital sentencing in Georgia. Nonetheless, it finds the probability of prejudice insufficient to create constitutional concern. Close analysis of the *Baldus* study, however, in light of both statistical principles and human experience, reveals that the risk that race influenced *McCleskey's* sentence is intolerable by any imaginable standard.

...

The *Baldus* study indicates that, after taking into account some 230 non-racial factors that might legitimately influence a sentencer, the jury more likely than not would have spared *McCleskey's* life had his victim been black. The study distinguishes between those cases in which (1) the jury exercises virtually no discretion because the strength or weakness of aggravating factors usually suggests that only one outcome is appropriate; and (2) cases reflecting an "intermediate" level of aggravation, in which the jury has con-

siderable discretion in choosing a sentence. McCleskey's case falls into the intermediate range. In such cases, death is imposed in 34% of white-victim crimes and 14% of black-victim crimes, a difference of 139% in the rate of imposition of the death penalty. In other words, just under 59%—almost 6 in 10—defendants comparable to McCleskey would not have received the death penalty if their victims had been black.

Furthermore, even examination of the sentencing system as a whole, factoring in those cases in which the jury exercises little discretion, indicates the influence of race on capital sentencing. For the Georgia system as a whole, race accounts for a six percentage point difference in the rate at which capital punishment is imposed. Since death is imposed in 11% of all white-victim cases, the rate in comparably aggravated black-victim cases is 5%. The rate of capital sentencing in a white-victim case is thus 120% greater than the rate in a black-victim case. Put another way, over half—55%—of defendants in white-victim crimes in Georgia would not have been sentenced to die if their victims had been black. Of the more than 200 variables potentially relevant to a sentencing decision, race of the victim is a powerful explanation for variation in death sentence rates—as powerful as nonracial aggravating factors such as a prior murder conviction or acting as the principal planner of the homicide.

These adjusted figures are only the most conservative indication of the risk that race will influence the death sentences of defendants in Georgia. Data unadjusted for the mitigating or aggravating effect of other factors show an even more pronounced disparity by race. The capital sentencing rate for all white-victim cases was almost 11 times greater than the rate for black-victim cases. Furthermore, blacks who kill whites are sentenced to death at nearly 22 times the rate of blacks who kill blacks, and more than 7 times the rate of whites who kill blacks. In addition, prosecutors seek the death penalty for 70% of black defendants with white victims, but for only 15% of black defendants with black victims, and only 19% of white defendants with black victims. Since our decision upholding the Georgia capital sentencing system in *Gregg*, the State has executed seven persons. All of the seven were convicted of killing whites, and six of the seven executed were black. Such execution figures are especially striking in light of the fact that,

during the period encompassed by the Baldus study, only 9.2% of Georgia homicides involved black defendants and white victims, while 60.7% involved black victims.

McCleskey's statistics have particular force because most of them are the product of sophisticated multiple-regression analysis. Such analysis is designed precisely to identify patterns in the aggregate, even though we may not be able to reconstitute with certainty any individual decision that goes to make up that pattern. Multiple-regression analysis is particularly well suited to identify the influence of impermissible considerations in sentencing, since it is able to control for permissible factors that may explain an apparent arbitrary pattern. While the decision-making process of a body such as a jury may be complex, the Baldus study provides a massive compilation of the details that are most relevant to that decision. As we held in the context of Title VII of the Civil Rights Act of 1964 last Term in *Bazemore v. Friday*, a multiple-regression analysis need not include every conceivable variable to establish a party's case, as long as it includes those variables that account for the major factors that are likely to influence decisions. In this case, Professor Baldus in fact conducted additional regression analyses in response to criticisms and suggestions by the District Court, all of which confirmed, and some of which even strengthened, the study's original conclusions.

The statistical evidence in this case thus relentlessly documents the risk that McCleskey's sentence was influenced by racial considerations. This evidence shows that there is a better than even chance in Georgia that race will influence the decision to impose the death penalty: a majority of defendants in white-victim crimes would not have been sentenced to die if their victims had been black. In determining whether this risk is acceptable, our judgment must be shaped by the awareness that [t]he risk of racial prejudice infecting a capital sentencing proceeding is especially serious in light of the complete finality of the death sentence, and that [i]t is of vital importance to the defendant and to the community that any decision to impose the death sentence be, and appear to be, based on reason rather than caprice or emotion. In determining the guilt of a defendant, a State must prove its case beyond a reasonable doubt. That is, we refuse to convict if the chance of error is simply less likely than not. Surely, we should not be willing to take a person's

life if the chance that his death sentence was irrationally imposed is more likely than not. In light of the gravity of the interest at stake, petitioner's statistics, on their face, are a powerful demonstration of the type of risk that our Eighth Amendment jurisprudence has consistently condemned.

...

Evaluation of McCleskey's evidence cannot rest solely on the numbers themselves. We must also ask whether the conclusion suggested by those numbers is consonant with our understanding of history and human experience. Georgia's legacy of a race-conscious criminal justice system, as well as this Court's own recognition of the persistent danger that racial attitudes may affect criminal proceedings, indicates that McCleskey's claim is not a fanciful product of mere statistical artifice.

For many years, Georgia operated openly and formally precisely the type of dual system the evidence shows is still effectively in place. The criminal law expressly differentiated between crimes committed by and against blacks and whites, distinctions whose lineage traced back to the time of slavery. During the colonial period, black slaves who killed whites in Georgia, regardless of whether in self-defense or in defense of another, were automatically executed.

By the time of the Civil War, a dual system of crime and punishment was well established in Georgia. The state criminal code contained separate sections for "Slaves and Free Persons of Color," and for all other persons. The code provided, for instance, for an automatic death sentence for murder committed by blacks, but declared that anyone else convicted of murder might receive life imprisonment if the conviction were founded solely on circumstantial testimony or simply if the jury so recommended. The code established that the rape of a free white female by a black "shall be punishable by death. However, rape by anyone else of a free white female was punishable by a prison term not less than 2 nor more than 20 years. The rape of blacks was punishable "by fine and imprisonment, at the discretion of the court." A black convicted of assaulting a free white person with intent to murder could be put to death at the discretion of the court, but the same offense committed against a black, slave or free, was classified as a "minor" offense whose punishment lay in the discretion of the court, as long as such punishment did not "extend to life, limb, or health." Assault with intent to murder by a

white person was punishable by a prison term of from 2 to 10 years. While sufficient provocation could reduce a charge of murder to manslaughter, the code provided that [o]bedience and submission being the duty of a slave, much greater provocation is necessary to reduce a homicide of a white person by him to voluntary manslaughter, than is prescribed for white persons.

In more recent times, some 40 years ago, Gunnar Myrdal's epochal study of American race relations produced findings mirroring McCleskey's evidence:

As long as only Negroes are concerned and no whites are disturbed, great leniency will be shown in most cases. . . . The sentences for even major crimes are ordinarily reduced when the victim is another Negro.

...

For offenses which involve any actual or potential danger to whites, however, Negroes are punished more severely than whites.

...

On the other hand, it is quite common for a white criminal to be set free if his crime was against a Negro.

...

This Court has invalidated portions of the Georgia capital sentencing system three times over the past 15 years. The specter of race discrimination was acknowledged by the Court in striking down the Georgia death penalty statute in *Furman*. Justice Douglas cited studies suggesting imposition of the death penalty in racially discriminatory fashion, and found the standard-less statutes before the Court "pregnant with discrimination." Justice Marshall pointed to statistics indicating that Negroes [have been] executed far more often than whites in proportion to their percentage of the population. Studies indicate that, while the higher rate of execution among Negroes is partially due to a higher rate of crime, there is evidence of racial discrimination. Although Justice Stewart declined to conclude that racial discrimination had been plainly proved, he stated that [m]y concurring Brothers have demonstrated that, if any basis can be discerned for the selection of these few to be sentenced to die, it is the constitutionally impermissible basis of race. In dissent, Chief Justice Burger acknowledged that statistics suggest, at least as a historical matter, that Negroes have been sentenced to death with greater frequency than whites in several States, particularly for the crime of interracial

rape. Finally, also in dissent, Justice Powell intimated that an Equal Protection Clause argument would be available for a black who could demonstrate that members of his race were being singled out for more severe punishment than others charged with the same offense. He noted that, although the Eighth Circuit had rejected a claim of discrimination in *Maxwell v. Bishop*, vacated and remanded on other grounds, the statistical evidence in that case tend[ed] to show a pronounced disproportion in the number of Negroes receiving death sentences for rape in parts of Arkansas and elsewhere in the South. It is clear that the Court regarded the opportunity for the operation of racial prejudice a particularly troublesome aspect of the unbounded discretion afforded by the Georgia sentencing scheme. Five years later, the Court struck down the imposition of the death penalty in Georgia for the crime of rape. Although the Court did not explicitly mention race, the decision had to have been informed by the specific observations on rape by both the Chief Justice and Justice Powell in *Furman*. Furthermore, evidence submitted to the Court indicated that black men who committed rape, particularly of white women, were considerably more likely to be sentenced to death than white rapists. For instance, by 1977, Georgia had executed 62 men for rape since the Federal Government began compiling statistics in 1930. Of these men, 58 were black and 4 were white. Three years later, the Court in *Godfrey* found one of the State's statutory aggravating factors unconstitutionally vague, since it resulted in "standard-less and unchanneled imposition of death sentences in the uncontrolled discretion of a basically uninstructed jury. . . ." Justice Marshall, concurring in the judgment, noted that [t]he disgraceful distorting effects of racial discrimination and poverty continue to be painfully visible in the imposition of death sentences.

This historical review of Georgia criminal law is not intended as a bill of indictment calling the State to account for past transgressions. Citation of past practices does not justify the automatic condemnation of current ones. But it would be unrealistic to ignore the influence of history in assessing the plausible implications of McCleskey's evidence. [A]mericans share a historical experience that has resulted in individuals within the culture ubiquitously attaching a significance to race that is irrational and often outside their awareness. As we said in *Rose v. Mitchell*: [W]e . . . cannot deny that,

114 years after the close of the War Between the States and nearly 100 years after *Strauder*, racial and other forms of discrimination still remain a fact of life, in the administration of justice as in our society as a whole. Perhaps today that discrimination takes a form more subtle than before. But it is not less real or pernicious.

The ongoing influence of history is acknowledged, as the majority observes, by our “unceasing efforts to eradicate racial prejudice from our criminal justice system.” These efforts, however, signify not the elimination of the problem, but its persistence. Our cases reflect a realization of the myriad of opportunities for racial considerations to influence criminal proceedings: in the exercise of peremptory challenges, in the selection of the grand jury, in the selection of the petit jury, in the exercise of prosecutorial discretion, in the conduct of argument, and in the conscious or unconscious bias of jurors.

The discretion afforded prosecutors and jurors in the Georgia capital sentencing system creates such opportunities. No guidelines govern prosecutorial decisions to seek the death penalty, and Georgia provides juries with no list of aggravating and mitigating factors, nor any standard for balancing them against one another. Once a jury identifies one aggravating factor, it has complete discretion in choosing life or death, and need not articulate its basis for selecting life imprisonment. The Georgia sentencing system therefore provides considerable opportunity for racial considerations, however subtle and unconscious, to influence charging and sentencing decisions.

History and its continuing legacy thus buttress the probative force of McCleskey’s statistics. Formal dual criminal laws may no longer be in effect, and intentional discrimination may no longer be prominent. Nonetheless, as we acknowledged in *Turner*, “subtle, less consciously held racial attitudes” continue to be of concern, and the Georgia system gives such attitudes considerable room to operate. The conclusions drawn from McCleskey’s statistical evidence are therefore consistent with the lessons of social experience.

The majority thus misreads our Eighth Amendment jurisprudence in concluding that McCleskey has not demonstrated a degree of risk sufficient to raise constitutional concern. The determination of the significance of his evidence is at its core an exercise in human moral judgment, not a mechanical statistical analysis. It must first and foremost be informed by awareness of

the fact that death is irrevocable, and that, as a result, the qualitative difference of death from all other punishments requires a greater degree of scrutiny of the capital sentencing determination. For this reason, we have demanded a uniquely high degree of rationality in imposing the death penalty. A capital sentencing system in which race more likely than not plays a role does not meet this standard. It is true that every nuance of decision cannot be statistically captured, nor can any individual judgment be plumbed with absolute certainty. Yet the fact that we must always act without the illumination of complete knowledge cannot induce paralysis when we confront what is literally an issue of life and death. Sentencing data, history, and experience all counsel that Georgia has provided insufficient assurance of the heightened rationality we have required in order to take a human life.

...

The Court cites four reasons for shrinking from the implications of McCleskey's evidence: the desirability of discretion for actors in the criminal justice system, the existence of statutory safeguards against abuse of that discretion, the potential consequences for broader challenges to criminal sentencing, and an understanding of the contours of the judicial role. While these concerns underscore the need for sober deliberation, they do not justify rejecting evidence as convincing as McCleskey has presented.

The Court maintains that petitioner's claim "is antithetical to the fundamental role of discretion in our criminal justice system." It states that "[w]here the discretion that is fundamental to our criminal process is involved, we decline to assume that what is unexplained is invidious."

Reliance on race in imposing capital punishment, however, is antithetical to the very rationale for granting sentencing discretion. Discretion is a means, not an end. It is bestowed in order to permit the sentencer to "trea[t] each defendant in a capital case with that degree of respect due the uniqueness of the individual." The decision to impose the punishment of death must be based on a "particularized consideration of relevant aspects of the character and record of each convicted defendant." Failure to conduct such an individualized moral inquiry treats all persons convicted of a designated offense not as unique individual human beings, but as members of a faceless, undifferentiated mass to be subjected to the blind infliction of the penalty of

death.

Considering the race of a defendant or victim in deciding if the death penalty should be imposed is completely at odds with this concern that an individual be evaluated as a unique human being. Decisions influenced by race rest in part on a categorical assessment of the worth of human beings according to color, insensitive to whatever qualities the individuals in question may possess. Enhanced willingness to impose the death sentence on black defendants, or diminished willingness to render such a sentence when blacks are victims, reflects a devaluation of the lives of black persons. When confronted with evidence that race more likely than not plays such a role in a capital sentencing system, it is plainly insufficient to say that the importance of discretion demands that the risk be higher before we will act—for, in such a case, the very end that discretion is designed to serve is being undermined.

Our desire for individualized moral judgments may lead us to accept some inconsistencies in sentencing outcomes. Since such decisions are not reducible to mathematical formulae, we are willing to assume that a certain degree of variation reflects the fact that no two defendants are completely alike. There is thus a presumption that actors in the criminal justice system exercise their discretion in responsible fashion, and we do not automatically infer that sentencing patterns that do not comport with ideal rationality are suspect.

As we made clear in *Batson v. Kentucky*, however, that presumption is rebuttable. *Batson* dealt with another arena in which considerable discretion traditionally has been afforded, the exercise of peremptory challenges. Those challenges are normally exercised without any indication whatsoever of the grounds for doing so. The rationale for this deference has been a belief that the unique characteristics of particular prospective jurors may raise concern on the part of the prosecution or defense, despite the fact that counsel may not be able to articulate that concern in a manner sufficient to support exclusion for cause. As with sentencing, therefore, peremptory challenges are justified as an occasion for particularized determinations related to specific individuals, and, as with sentencing, we presume that such challenges normally are not made on the basis of a factor such as race. As we said in *Batson*, however, such features do not justify imposing a “crippling burden of proof,” in order to rebut that presumption. The Court in this case apparently seeks

to do just that. On the basis of the need for individualized decisions, it rejects evidence, drawn from the most sophisticated capital sentencing analysis ever performed, that reveals that race more likely than not infects capital sentencing decisions. The Court's position converts a rebuttable presumption into a virtually conclusive one.

The Court also declines to find McCleskey's evidence sufficient in view of "the safeguards designed to minimize racial bias in the [capital sentencing] process." *Gregg v. Georgia*, upheld the Georgia capital sentencing statute against a facial challenge which Justice White described in his concurring opinion as based on "simply an assertion of lack of faith" that the system could operate in a fair manner (opinion concurring in judgment). Justice White observed that the claim that prosecutors might act in an arbitrary fashion was "unsupported by any facts," and that prosecutors must be assumed to exercise their charging duties properly "[a]bsent facts to the contrary." It is clear that *Gregg* bestowed no permanent approval on the Georgia system. It simply held that the State's statutory safeguards were assumed sufficient to channel discretion without evidence otherwise.

It has now been over 13 years since Georgia adopted the provisions upheld in *Gregg*. Professor Baldus and his colleagues have compiled data on almost 2,500 homicides committed during the period 1973-1979. They have taken into account the influence of 230 nonracial variables, using a multitude of data from the State itself, and have produced striking evidence that the odds of being sentenced to death are significantly greater than average if a defendant is black or his or her victim is white. The challenge to the Georgia system is not speculative or theoretical; it is empirical. As a result, the Court cannot rely on the statutory safeguards in discounting McCleskey's evidence, for it is the very effectiveness of those safeguards that such evidence calls into question. While we may hope that a model of procedural fairness will curb the influence of race on sentencing, "we cannot simply assume that the model works as intended; we must critique its performance in terms of its results."

The Court next states that its unwillingness to regard petitioner's evidence as sufficient is based in part on the fear that recognition of McCleskey's claim would open the door to widespread challenges to all aspects of criminal sentencing. Taken on its face, such a statement seems to suggest a fear

of too much justice. Yet surely the majority would acknowledge that, if striking evidence indicated that other minority groups, or women, or even persons with blond hair, were disproportionately sentenced to death, such a state of affairs would be repugnant to deeply rooted conceptions of fairness. The prospect that there may be more widespread abuse than McCleskey documents may be dismaying, but it does not justify complete abdication of our judicial role. The Constitution was framed fundamentally as a bulwark against governmental power, and preventing the arbitrary administration of punishment is a basic ideal of any society that purports to be governed by the rule of law.

In fairness, the Court's fear that McCleskey's claim is an invitation to descend a slippery slope also rests on the realization that any humanly imposed system of penalties will exhibit some imperfection. Yet to reject McCleskey's powerful evidence on this basis is to ignore both the qualitatively different character of the death penalty and the particular repugnance of racial discrimination, considerations which may properly be taken into account in determining whether various punishments are "cruel and unusual." Furthermore, it fails to take account of the unprecedented refinement and strength of the Baldus study.

It hardly needs reiteration that this Court has consistently acknowledged the uniqueness of the punishment of death. Death, in its finality, differs more from life imprisonment than a 100-year prison term differs from one of only a year or two. Because of that qualitative difference, there is a corresponding difference in the need for reliability in the determination that death is the appropriate punishment. Furthermore, the relative interests of the state and the defendant differ dramatically in the death penalty context. The marginal benefits accruing to the state from obtaining the death penalty, rather than life imprisonment, are considerably less than the marginal difference to the defendant between death and life in prison. Such a disparity is an additional reason for tolerating scant arbitrariness in capital sentencing. Even those who believe that society can impose the death penalty in a manner sufficiently rational to justify its continuation must acknowledge that the level of rationality that is considered satisfactory must be uniquely high. As a result, the degree of arbitrariness that may be adequate to render the death

penalty “cruel and unusual” punishment may not be adequate to invalidate lesser penalties. What these relative degrees of arbitrariness might be in other cases need not concern us here; the point is that the majority’s fear of wholesale invalidation of criminal sentences is unfounded.

The Court also maintains that accepting McCleskey’s claim would pose a threat to all sentencing because of the prospect that a correlation might be demonstrated between sentencing outcomes and other personal characteristics. Again, such a view is indifferent to the considerations that enter into a determination whether punishment is “cruel and unusual.” Race is a consideration whose influence is expressly constitutionally proscribed. We have expressed a moral commitment, as embodied in our fundamental law, that this specific characteristic should not be the basis for allotting burdens and benefits. Three constitutional amendments, and numerous statutes, have been prompted specifically by the desire to address the effects of racism.

Over the years, this Court has consistently repudiated “[d]istinctions between citizens solely because of their ancestry” as being “odious to a free people whose institutions are founded upon the doctrine of equality.”

Furthermore, we have explicitly acknowledged the illegitimacy of race as a consideration in capital sentencing. That a decision to impose the death penalty could be influenced by race is thus a particularly repugnant prospect, and evidence that race may play even a modest role in levying a death sentence should be enough to characterize that sentence as “cruel and unusual.”

Certainly, a factor that we would regard as morally irrelevant, such as hair color, at least theoretically could be associated with sentencing results to such an extent that we would regard as arbitrary a system in which that factor played a significant role. As I have said above, however, the evaluation of evidence suggesting such a correlation must be informed not merely by statistics, but by history and experience. One could hardly contend that this Nation has, on the basis of hair color, inflicted upon persons deprivation comparable to that imposed on the basis of race. Recognition of this fact would necessarily influence the evaluation of data suggesting the influence of hair color on sentencing, and would require evidence of statistical correlation even more powerful than that presented by the Baldus study.

Furthermore, the Court’s fear of the expansive ramifications of a holding

for McCleskey in this case is unfounded, because it fails to recognize the uniquely sophisticated nature of the Baldus study. McCleskey presents evidence that is far and away the most refined data ever assembled on any system of punishment, data not readily replicated through casual effort. Moreover, that evidence depicts not merely arguable tendencies, but striking correlations, all the more powerful because nonracial explanations have been eliminated. Acceptance of petitioner's evidence would therefore establish a remarkably stringent standard of statistical evidence unlikely to be satisfied with any frequency.

The Court's projection of apocalyptic consequences for criminal sentencing is thus greatly exaggerated. The Court can indulge in such speculation only by ignoring its own jurisprudence demanding the highest scrutiny on issues of death and race. As a result, it fails to do justice to a claim in which both those elements are intertwined—an occasion calling for the most sensitive inquiry a court can conduct. Despite its acceptance of the validity of Warren McCleskey's evidence, the Court is willing to let his death sentence stand because it fears that we cannot successfully define a different standard for lesser punishments. This fear is baseless.

Finally, the Court justifies its rejection of McCleskey's claim by cautioning against usurpation of the legislatures' role in devising and monitoring criminal punishment. The Court is, of course, correct to emphasize the gravity of constitutional intervention, and the importance that it be sparingly employed. The fact that “[c]apital punishment is now the law in more than two thirds of our States,” however, does not diminish the fact that capital punishment is the most awesome act that a State can perform. The judiciary's role in this society counts for little if the use of governmental power to extinguish life does not elicit close scrutiny. It is true that society has a legitimate interest in punishment. Yet, as Alexander Bickel wrote:

It is a premise we deduce not merely from the fact of a written constitution but from the history of the race, and ultimately as a moral judgment of the good society, that government should serve not only what we conceive from time to time to be our immediate material needs, but also certain enduring values. This in part is what is meant by government under law.

Our commitment to these values requires fidelity to them even when there

is temptation to ignore them. Such temptation is especially apt to arise in criminal matters, for those granted constitutional protection in this context are those whom society finds most menacing and opprobrious. Even less sympathetic are those we consider for the sentence of death, for execution “is a way of saying, ‘You are not fit for this world, take your chance elsewhere.’” For these reasons, [t]he methods we employ in the enforcement of our criminal law have aptly been called the measures by which the quality of our civilization may be judged. Those whom we would banish from society or from the human community itself often speak in too faint a voice to be heard above society’s demand for punishment. It is the particular role of courts to hear these voices, for the Constitution declares that the majoritarian chorus may not alone dictate the conditions of social life. The Court thus fulfills, rather than disrupts, the scheme of separation of powers by closely scrutinizing the imposition of the death penalty, for no decision of a society is more deserving of “sober second thought.”

...

At the time our Constitution was framed 200 years ago this year, blacks had for more than a century before been regarded as beings of an inferior order, and altogether unfit to associate with the white race, either in social or political relations; and so far inferior that they had no rights which the white man was bound to respect. Only 130 years ago, this Court relied on these observations to deny American citizenship to blacks. A mere three generations ago, this Court sanctioned racial segregation, stating that “[i]f one race be inferior to the other socially, the Constitution of the United States cannot put them upon the same plane.”

In more recent times, we have sought to free ourselves from the burden of this history. Yet it has been scarcely a generation since this Court’s first decision striking down racial segregation, and barely two decades since the legislative prohibition of racial discrimination in major domains of national life. These have been honorable steps, but we cannot pretend that, in three decades, we have completely escaped the grip of a historical legacy spanning centuries. Warren McCleskey’s evidence confronts us with the subtle and persistent influence of the past. His message is a disturbing one to a society that has formally repudiated racism, and a frustrating one to a Nation accus-

tomed to regarding its destiny as the product of its own will. Nonetheless, we ignore him at our peril, for we remain imprisoned by the past as long as we deny its influence in the present.

It is tempting to pretend that minorities on death row share a fate in no way connected to our own, that our treatment of them sounds no echoes beyond the chambers in which they die. Such an illusion is ultimately corrosive, for the reverberations of injustice are not so easily confined. “The destinies of the two races in this country are indissolubly linked together,” and the way in which we choose those who will die reveals the depth of moral commitment among the living.

The Court’s decision today will not change what attorneys in Georgia tell other Warren McCleskeys about their chances of execution. Nothing will soften the harsh message they must convey, nor alter the prospect that race undoubtedly will continue to be a topic of discussion. McCleskey’s evidence will not have obtained judicial acceptance, but that will not affect what is said on death row. However many criticisms of today’s decision may be rendered, these painful conversations will serve as the most eloquent dissents of all.