Social Science and the Evolving Standards of Death Penalty Law

Samuel R. Gross  
*University of Michigan Law School*

Phoebe C. Ellsworth  
*University of Michigan Law School*

Available at: https://repository.law.umich.edu/book_chapters/254

Follow this and additional works at: https://repository.law.umich.edu/book_chapters

Part of the Criminal Law Commons

Publication Information & Recommended Citation


This Book Chapter is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Book Chapters by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.
Introduction

Unlike many of the topics covered in this book, death penalty litigation involves a wide variety of empirical issues. The Eighth Amendment of the U.S. Constitution provides that “Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishment inflicted.” But what is a “cruel and unusual punishment?” It could be a punishment that is morally unacceptable to the American people, like cutting off noses or hands. Following the other clauses of the amendment, it could be a punishment that is excessive, in that a lesser penalty would achieve the same ends. For example, if a death sentence served no penological purpose that was not served just as well by life imprisonment, the death penalty might be seen as excessive. Or the death penalty could be cruel and unusual in practice if it is rarely imposed and if the decision to take or spare the criminal’s life is unprincipled. We would like to believe that the people who are executed are those who have committed the most monstrous crimes. If instead there is no rational means of differentiating the few who are sentenced to die from other killers, except perhaps for impermissible criteria like race or poverty, that could be cruel and unusual. All of these definitions have been proposed, and each has produced various lines of empirical research that have played a role in death penalty litigation.

The constitutionality of the death penalty, and of the procedures that are used to impose it, are questions of law. Although much of the research involves juries and the public at large, on these issues the ultimate fact finders and evaluators of the research are judges, usually Supreme Court justices, and it is their “common knowledge” and their ability to evaluate the research that matters. The evidence comes to
them summarized in briefs written by the parties or by "friends of the court" – amici curiae. These amici may include professional organizations such as the American Psychological Association, or ad hoc groups of social scientists. The justices may evaluate the quality of the research, but initially they also decide whether the research is even relevant to the constitutional question before them, as they understand it.

It would be impossible in a chapter of this length to provide a detailed description of all the varieties of empirical research related to the death penalty, the ways in which they have been presented to the courts, and the courts' responses to them. Instead, this chapter provides a general framework for classifying and understanding the kinds of empirical questions that have been raised and their constitutional implications, the kinds of data that have been presented to the courts, and the courts' response.

It is important to remember that public policy on the death penalty does not just involve knowledge; it also involves values. Over the last 40 years capital punishment has been an emotionally charged issue in the United States, central to people's ideological self-image (Ellsworth & Ross, 1983; Ellsworth & Gross, 1994; Gross, 1998). Most of the important cases on capital punishment were decided during this period, and all of the cases that involved empirical research. The values of the Supreme Court – like those of the public at large – were decidedly pro-death penalty. As a result their evaluations of the evidence may have been less objective than their evaluation of social science evidence on less emotionally charged topics.

Public Acceptance of Capital Punishment: “Evolving Standards of Decency”

The most direct definition of a cruel and unusual punishment is a punishment that is uncommon and morally unacceptable to the public. No one, not even the strict constructionists on the Court, believes that all of the punishments commonly used in the late eighteenth century, when the Constitution was drafted, would be acceptable under the Eighth Amendment today. Mutilation and flogging were abandoned long ago. Instead, the Supreme Court has held that the Eighth Amendment “draws its meaning from the evolving standards of decency that mark the progress of a maturing society” (Trop v. Dulles, 1958). The implication is that there is a trend toward a penal system that is more enlightened and humane, and that some punishments that once seemed normal may eventually come to be seen as
barbarous. The empirical questions are whether such a trend exists and whether we have evolved to a point where the death penalty is no longer tolerable in our society.

In Europe, this trend is clear. Since the Second World War every European country has abolished capital punishment. By the turn of the twenty-first century the policy of the European Union was that "there is no case where capital punishment can be justified under the international human rights standard" (Zimring, 2003, p. 27), and no nation that retained the death penalty was permitted to be a member. The European governments' standards of decency seem to have evolved more rapidly than those of their citizens. In most European countries the death penalty was abolished despite public support, and it was only later that popular opinion fell in line with official policy.

In the United States the primary responsibility for criminal justice rests with the states. Neither Congress nor the president has any obvious authority to abolish the death penalty at the state level. The Supreme Court could do so, however, because the Fourteenth Amendment requires states to provide "due process of law" in all criminal cases, and due process has been interpreted to include the Eighth Amendment prohibition of cruel and unusual punishment. In considering whether executions violate that prohibition, the Supreme Court has looked to the ethical standards of contemporary society.

Public support for the death penalty in the United States fell steadily from the early 1950s until the mid-1960s, when opponents were at least as numerous as proponents. The number of executions also declined during this period, from over 100 in 1951 to nearly zero in the mid- to late 1960s, followed by a judicial moratorium on executions that lasted from 1967 until 1976. In 1972, in Furman v. Georgia, the Supreme Court held that all death penalty laws, as administered at the time, were unconstitutional. But the Court left open the possibility that a fairer, less arbitrary system of capital punishment might pass constitutional muster. In 1966 public support for the death penalty reached its lowest point, with only 47% of Americans favoring it (Ellsworth & Gross, 1994); after that, the trend reversed, and support for the death penalty increased sharply between 1966 and 1982, and moderately between 1982 and the mid-1990s, when one poll (Gallup, 1994) found that 80% of Americans favored the death penalty. Since then, support has declined to around 65% in 2005. By 1976, when the Supreme Court reconsidered the constitutionality of the death penalty in Gregg v. Georgia, support for the death penalty had already risen to 66% (NORC, 1976), and the Court concluded that "a large proportion of American society continues to regard death
as an appropriate and necessary criminal sanction" (Gregg v. Georgia, 1976, p. 878), and thus that capital punishment was not cruel and unusual by contemporary standards.

In general, the Supreme Court's opinions have tracked the empirical data on public opinion towards the death penalty fairly well. The death penalty was abolished not long after public support reached an all-time low, and was reinstated following the dramatic increase in support through the mid-1970's (Gregg v. Georgia, 1976). This is not to say that the Court analyzed the public opinion poll data carefully, or that the justices actually based their decisions on the empirical research. Their "empirical evidence" was broader and less systematic. Occasionally the opinions refer to a public opinion poll or a scholarly article, but they rely more heavily on state referenda, new legislation, and patterns of jury verdicts. On the whole, the Court has been right about where America stands in the progress of a maturing society.

This is not surprising. The results of public opinion polls on the death penalty are widely publicized, easy to understand, obviously related to the constitutional question at issue, and rarely disputed. Death penalty attitudes are relatively immune to the kinds of changes in question wording and order that raise doubts about the validity of surveys on some other topics (Schuman & Presser, 1981; Schwarz, Groves, & Schuman, 1998; Ellsworth & Gross, 1994; Gross & Ellsworth, 2003). There is quite a bit of research on death penalty attitudes that is more complex than the simple favor–oppose polls, research exploring the bases of people's attitudes, regional and demographic differences, and the reasons for attitude change. Few Justices have paid attention to this more nuanced research, but it is not clear that there is any reason that they should. All of the polls since the death penalty was reinstated show that a majority of Americans favor it. After 1995 this majority fell to 65%. It hovers around 50%, if respondents are explicitly given the choice between death and life in prison without parole, but it has not yet fallen to a level that could plausibly be called public rejection, and the Court has not considered the possibility of abolishing the death penalty since 1976.

The Court has attended to differences in public support of the death penalty for different kinds of capital defendant. Just a year after the death penalty was reinstated in Gregg, the Court held that the penalty of death was disproportionate and excessive for rapists, and therefore unconstitutional, and justified their decision with evidence that the public no longer felt that the death penalty was appropriate for the crime of rape (Coker v. Georgia, 1977). (The Court mostly cited legislative decisions and jury verdicts, rather than opinion polls, and
completely ignored the powerful empirical evidence of racial bias in executions for rape.)

In the late 1990s, after a 15-year plateau of general enthusiasm for the death penalty, support dropped significantly, from 75% or higher to about 65% (Gross & Ellsworth, 2003; Death Penalty Information Center). Correspondingly, after a short lag, the Supreme Court has begun to restrict the use of the death penalty. In two recent cases, the Court relied heavily on changes in public attitudes. In *Atkins v. Virginia* (2002) the Court held that the execution of people who are mentally retarded is unconstitutional, relying primarily on the fact that a majority of state legislatures had outlawed the practice; but also citing briefs by the APA and other professional organizations, as well as national opinion polls, to reach the conclusion that the "legislative judgment reflects a much broader social and professional consensus" (*Atkins v. Virginia*, 2002, FN 21). This is an accurate reflection of the research on public opinion. Even when levels of overall support were at their highest, only about 25% of the population favored the death penalty for mentally retarded killers. For example in 1988, when 71% of the population favored execution of murderers and 51% favored the execution of rapists, only 21% favored the death penalty for mentally retarded defendants (Gross & Ellsworth, 2003). By 2002 support was down to 13% (Gallup, 2002).

In the second case, *Roper v. Simmons* (2005), the Court held that it was unconstitutional to execute people for crimes they committed before the age of 18. The opinion closely follows the holding in *Atkins*, using evolving standards of decency as the rationale, and relying primarily on state legislative changes for evidence, although the movement to abolish capital punishment for defendants under 18 was more sluggish than it was for mentally retarded defendants. The Court used social science data not for the proposition that the public has come to repudiate the execution of juveniles, but for evidence that the juvenile mind is not the same as the adult mind: It is less aware of responsibility, less able to resist peer pressure, and more susceptible to further development. The Court also emphasized the rejection of the death penalty for juveniles by every other nation in the world (except Somalia, which had no functioning government.)

On the surface, the *Roper* opinion looks like the *Atkins* opinion bolstered by two additional lines of supporting evidence. A cynic might conclude that these two new lines of evidence were trotted out because the core argument – that the American public rejects the death penalty for young offenders – is somewhat weaker than the parallel argument in *Atkins*. The legislative evidence is less compelling, and
public opinion surveys — which the Court does not mention in Roper — are less overwhelmingly one sided. In 1989, when the Court held in Stanford v. Kentucky that the execution of 16- and 17-year-olds was acceptable to Americans, 57% of the public favored that position (Time\CNN\Yankelovich, 1989). In 1994 61% favored it (Gallup, 1994). By 2001, with overall support for the death penalty declining, support for juvenile executions dropped to about 36% (NORC, 2001 [34%]; Princeton Survey Research Associates, 2001 [38%]), and remained at that level (Princeton Survey Research Associates, 2003 [35%], and 2005 [37%]). Still, while rejection of the death penalty for juvenile defendants is more recent and less overwhelming than for retarded defendants, the Court’s decision accurately tracks changes in public opinion.

**Excessiveness of Capital Punishment: The Question of Deterrence**

The argument that the death penalty is excessive has two prongs. The simpler one is that it is grossly disproportionate to the crime for which it was imposed. This was the specific legal ground the Court relied on in Coker v. Georgia (1977) to outlaw the death penalty for rape. Evidence of public rejection of death as a punishment for rape (and other nonhomicidal crimes) was used to support the claim that the death penalty was excessive because it was categorically disproportionate to the crime. The Court has never entertained the argument that death is a disproportionate punishment for murder. But it did consider a second type of excessiveness in murder cases: that the death penalty is excessive because it serves no legitimate purpose that is not equally well served by life imprisonment. In Furman (1972) some of the justices in the majority took the position that death sentences as rare and unpredictable as the ones before the Court served no useful retributive purpose. The main question, however, was the marginal deterrence of capital punishment: Do executions deter homicide more effectively than imprisonment for life?

When Furman was decided the main source of systematic evidence on the deterrent effect of the death penalty was the seminal work of Thorsten Sellin (1967). Sellin conducted a series of comparisons of homicide rates in the United States between 1920 and 1955: Between states that had the death penalty and neighboring states that did not; within states that had had the death penalty and abolished it, and then sometimes restored it; and between states that abolished or restored
the death penalty and neighboring states that did not. None of these comparisons indicated that the death penalty lowered the homicide rate. Justice Marshall, in his concurring opinion in *Furman*, discussed this evidence in detail in support of his position that the death penalty is inherently unconstitutional.

In *Gregg* in 1976 the Court was presented with a new study, by Isaac Ehrlich, that analyzed the relationship between executions and homicide rates in the United States as a whole, by year, from 1933 through 1969. Using econometric models and multiple regression analyses, Ehrlich claimed to find that each execution reduced the homicide rate to the extent of saving eight lives (1975). By the time *Gregg* got to the Court, Ehrlich’s study had been subjected to extensive and withering criticism (see Blumstein, Cohen & Nagen, 1978; Dike, 1982; Lempert, 1981, for reviews). Among other problems, critics pointed out that he had not controlled for the severity of the alternative, noncapital sanctions for murder; and that the effect he found disappears entirely if the years from 1965 through 1969 are removed from the analysis – a period in which a very low rate of executions coincided with the beginning of a long-term rise in homicide rates. The plurality of the Court in the decisive opinion in *Gregg* did not attempt to address any of these methodological questions. Instead, it wrote the issue of deterrence out of the legal debate over the constitutionality of capital punishment. It noted that there were conflicting studies – which was literally true, ignoring the value of those studies – and concluded that “there is no convincing empirical evidence either supporting or refuting” the claim that the death penalty is a unique deterrent, and that this is “a complex factual issue the resolution of which properly rests with the legislatures” (*Gregg*, 1976, p. 186).

Having kicked the issue over to legislatures, the Supreme Court has never revisited the question of deterrence. Nor is it likely to. In *Furman* there was an active debate over the appropriateness of retribution as a justification for capital punishment (or punishment in general). For example, Justice Powell, writing for the four pro-death penalty dissenters, quoted an earlier Supreme Court statement that “Retribution is no longer the dominant objective of the criminal law” and explained: “It is clear, however, that the Court did not reject retribution altogether” (*Furman*, 1972, p. 452). In *Gregg*, four years later, the Court explicitly upheld retribution as a justification for the death penalty because a legislature could legitimately conclude “that certain crimes are themselves so grievous an affront to humanity that the only adequate response may be the penalty of death” (*Gregg*,
1976, p. 184). So deterrence is no longer a necessary argument for the punishment. The Court’s position here tracks public opinion. In the past forty years retribution has become accepted and popular in the United States as justification for punishment in general (Ellsworth & Gross, 1994), and it has overtaken deterrence and become the dominant reason people give for supporting for the death penalty. At the same time, a strong majority of the public has come to believe that the death penalty does not deter homicide better than life imprisonment (Gross & Ellsworth, 2003).

Academic debate over the deterrent effect of capital punishment continued after Gregg. In 1978 the National Academy issued a report reviewing the evidence in support of a deterrent effect of capital punishment – basically, the Ehrlich study – and found it unpersuasive (Blumstein et al., 1978). In the 20 years that followed there were periodic new studies on the issue, most of which found no deterrent effect (Bailey, 1998; Sorenson, Wrinkle, Brewer, & Marquart, 1999).

In the past several years there has been a spate of new econometric studies of deterrence and the death penalty, most of which claim to find a deterrent effect, but not all. As with Ehrlich’s, these studies have been reviewed and heavily criticized. Independent researchers have demonstrated that the findings of deterrence are unstable and depend on arbitrary methodological choices (Berk, 2005; Donohue and Wolfers, 2006), and have concluded that, as the National Academy found 28 years earlier, “We can be sure that the death penalty does not cause or eliminate large numbers of homicides, but we learn little else from the data” (Donohue & Wolfers, 2006, p. 844). The new claims of a deterrent effect – whatever their worth – are not likely to figure directly in court decisions one way or the other, but might be used in policy debates to bolster support for capital punishment now that it is in decline.

Arbitrariness and Discrimination: Capital Punishment in Practice

In 1972, when the Supreme Court held that the death penalty as then administered was unconstitutional (Furman v. Georgia, 1972), neither deterrence nor contemporary standards figured importantly in the decision. Furman is a particularly difficult case to interpret, because it was a 5 to 4 decision in which each Justice wrote a separate opinion. Brennan and Marshall, the two Justices who were categorically opposed to capital punishment, discussed both of these issues, but the only
issue that was common to the five majority Justices was that the death penalty statutes of the time resulted in decisions that were at best arbitrary, and possibly discriminatory against racial minorities, particularly African Americans. Those who were executed were not the ones who had committed the most awful murders, but a haphazard, unlucky selection — “like being struck by lightning” in the words of Justice Stewart (Furman, 1972, p. 309) — or, worse, a systematic selection of people without money, friends, or white skin. As Justice Douglas put it, “one searches our chronicles in vain for the execution of any member of the affluent strata of this society” (Furman, 1972, pp. 251-252).

Jury instructions

At the time of the Furman decision, capital juries were typically given no instructions at all on how to decide who should be executed and who should not. They were simply sent into the jury room and told to decide whether the defendant was guilty, and, if so, whether the penalty should be life or death.

By the time the Furman decision was handed down, public support for the death penalty was on the rise. Most states responded quickly and enacted new death penalty statutes designed to ensure that the unfettered jury discretion of the past would be suitably directed so that death sentences would be calibrated to the seriousness of the crime. In 1976, in Gregg v. Georgia and its companion cases, the Supreme Court declared that the death penalty was constitutional if states provided a bifurcated trial, in which the penalty was decided in a separate proceeding after guilt had been decided; there were specific instructions limiting the aggravating factors that could be considered in deciding for the death penalty, and requiring that the jury balance these against any mitigating factors. The idea was that these “guided discretion” statutes would eliminate the arbitrariness and possible discrimination that had existed in the pre-Furman days. As the plurality opinion stated, “No longer can a jury wantonly or freakishly impose the death sentence; it is always circumscribed by the legislative guidelines” (Gregg, 1976, pp. 206-207).

If the Court had hoped to settle the issue once and for all in Gregg, it certainly failed. Since then, capital cases have been a constant item on the Court’s agenda, and there is no sign that the number of cases is diminishing. Most of these cases have concerned the administration of the death penalty and the effectiveness of the post-Gregg laws in eliminating arbitrariness and discrimination.
This is not surprising. Decisions about the appropriate penalty – particularly when the choice is life or death – inevitably implicate values and moral principles more deeply than decisions about guilt or innocence. In most states the list of aggravating factors that the jury is supposed to consider include vague value-laden considerations – such as whether the murder was especially heinous, atrocious, and cruel – and sometimes considerations that invite race or class bias, such as whether the defendant would be dangerous or a threat to society in the future. One of the fundamental arguments for trial by jury is that the jury represents the “conscience of the community,” the values of the public, rather than the views of the government or an elite ruling class. Efforts to circumscribe community morality by formulaic rules may be incompatible not only with the purpose of the jury but with human psychology itself. In addition to the motivational and emotional difficulties of deciding good and evil according to a prescribed recipe, the cognitive challenge is substantial. Most researchers agree that jurors have a very hard time understanding legal instructions in general (Ellsworth & Reifman, 2000; Lieberman & Sales, 1997), and the instructions that they are given at the penalty phase of a capital trial are especially difficult (Diamond & Levi, 1996; Eisenberg & Wells, 1993; Haney & Lynch, 1994, 1997). Even before Furman, in McGautha v. California (1971) Justice Harlan foresaw this dilemma: “To identify before the fact those characteristics of criminal homicides and their perpetrators which call for the death penalty, and to express these characteristics in language which can be fairly understood and applied by the sentencing authority, appear to be tasks which are beyond present human ability” (McGautha, 1971, p. 204). The Court in Furman ignored this dire warning, and proceeded on the assumption that legislatures could somehow square the circle. Psychological research, unfortunately, seems to have confirmed Justice Harlan’s pessimistic assessment. It has certainly shown that efforts to date have not been successful.

The guidelines that jurors are supposed to apply in deciding between life and death are opaque. They include terms such as “aggravating circumstances” and “mitigating circumstances” that are unfamiliar or, worse yet, have different meanings in everyday discourse than they do in law (Haney & Lynch, 1994, 1997). In addition, many juries have difficulty understanding the rules for balancing aggravating and mitigating factors to reach a decision – what they should do if they find that aggravating factors outweigh mitigating factors, that mitigating factors outweigh aggravating, or that they are equally strong; and whether they must agree unanimously that an aggravating
or mitigating factor is present. Haney and Lynch (1997) found that 41% of jurors incorrectly believed that if the mitigating factors equaled the aggravating factors, they could still vote for death, and 27% believed that they could vote for death even if the mitigating factors outweighed the aggravating factors. More alarming, 41% of the jurors they studied thought that when the aggravating factors outweighed the mitigating factors the law required them to vote for death. This is not true. The law never requires a jury to vote for death, but apparently jurors believe that it does.

The Supreme Court has been remarkably oblivious to the research on this issue, basically subscribing to the legal fiction that if the jury is given the legally correct instructions, the jury understands the law. The case of *Weeks v. Angelone* (2000) is a vivid example. In that case the prosecutor argued that there were two applicable aggregating circumstances: that Weeks would "constitute a continuing serious threat to society," and that his crime was "outrageously or wantonly vile, horrible, or inhumane . . ." The jury was instructed: "If you find from the evidence that the Commonwealth has proved beyond a reasonable doubt, either of the two alternatives, and as to that alternative you are unanimous, then you may fix the punishment of the defendant at death; or, if you believe from all the evidence that the death penalty is not justified, then you shall fix the punishment of the defendant at life imprisonment" (*Weeks*, 2000, p. 229). The jury had a written copy of the instructions, but apparently had trouble figuring out what they meant. After deliberating for 4½ hours, they asked the judge: "If we believe that Lonnie Weeks Jr., is guilty of at least one of the alternatives, then is it our duty as a jury to issue the death penalty? Or must we decide (even though he is guilty of one of the alternatives) whether or not to issue the death penalty, or one of the life sentences? What is the Rule? Please clarify" (p. 241).

The judge could have answered the question by saying: "You are not required to give the death penalty even if you find one of the aggravating circumstances; it is up to you to decide between life and death." Instead he simply referred them to the paragraph in the instructions describing their choices, the very paragraph they said they did not understand. The jury, in tears, sentenced Weeks to death. Weeks appealed on the grounds that the jury had not understood that a life sentence was permissible. In this case, as in others (e.g., *Buchanan v. Angelone*, 1998) the consensus of the social science research was shared by the jurors themselves. They did not understand the instructions; they knew it; they said so; they asked for help – and they did not get any.
The Supreme Court was not impressed. The majority held that "A jury is presumed to follow the instructions [citations omitted]. Similarly, a jury is presumed to understand a judge's answers to its questions" (p. 234). No social science research was mentioned. Did the Court really believe that the jury understood the judge's answer? Or did it maintain the legal fiction of understanding, because the problem of juror misunderstanding was surely not unique to Weeks's case but had troublesome implications for many other capital and noncapital cases? Either way, it is clear that unintelligible instructions are no safeguard against arbitrary decisions.

An empirical study of people's understanding of capital jury instructions was presented to a federal trial court in the case of United States ex rel Free v. McGinnis (1993). The judge held that the instructions failed to provide jurors with a "clear understanding of how they are to go about deciding whether the defendant lives or dies" (p. 1129). The study, conducted by Hans Zeisel (described in Diamond & Levi, 1996), tested prospective jurors in a Chicago courthouse with 16 questions about the legal instructions for weighing aggravating and mitigating factors, and found that a majority of the respondents gave the correct answer on only three of the questions. The Seventh Circuit Court of Appeals reversed and reinstated Free's death sentence. However, unlike most courts, which simply affirm the legal presumption that jurors understand and follow the instructions, Judge Posner, writing for the majority in Free, said that empirical evidence that jurors misunderstand the instructions on capital sentencing could be grounds for a constitutional reversal. The data from Zeisel's study, however, were not persuasive, primarily because Zeisel had not shown that any alternative, rewritten instructions would have been more comprehensible to the jurors. A concurring judge also faulted the study for not including jury deliberations. In response, Diamond and Levi (1996) conducted a follow-up study in which they found that: (1) rewritten instructions were understood significantly better than the Illinois pattern instructions; and (2) deliberation generally did not increase the jurors' ability to understand the instructions. Diamond and Levi's data were not available to the Court when they reinstated Free's death sentence, and Free was executed in 1995. But the court's opinion in Free is an unusual and encouraging departure from the common judicial practice of altogether ignoring empirical data on juries' ability to understand the legal instructions. No matter how well crafted a set of guidelines might be in theory, if the jury cannot understand them, then in practice life or death decisions will be arbitrary.
Predicting dangerousness

So far we have been discussing the instructions governing the process by which jurors are to weigh the aggravating and mitigating factors in deciding between life and death. There are also problems with the factors themselves. Most states give jurors a list of aggravating factors and instruct them that they must find that one or more were present in order for the defendant to be eligible for the death penalty. Some of the factors are straightforward and unproblematic: for example, that the victim was a police officer. Many states include at least one factor that seems to allow the same unbridled discretion that was held unconstitutional in *Furman* (1972), allowing the jury to consider whether the murder was heinous, outrageous, wantonly cruel, vile, or depraved. Social science has generally had little to say about the validity of these specific factors, except for one. In several states the jury is asked to consider whether the defendant is likely to be a continuing threat to society, to commit violent crimes in the future. If so, the death penalty is warranted. The constitutionality of this factor was considered in the case of *Barefoot v. Estelle* (1983), in which the American Psychiatric Association filed a brief as a Friend of the Court. The APA pointed to a strong consensus among social scientists that not even experts can predict future dangerousness with any accuracy, and argued that future dangerousness should not be permissible as an aggravating factor because it was impossible to predict, and that expert psychological testimony of future dangerousness should be prohibited, because the data indicated that such predictions were usually wrong. In other words, the APA argued that its own members could not be trusted on this issue.

Nonetheless the Supreme Court held that the factor was constitutionally acceptable, and that expert psychologists and psychiatrists should be allowed to testify and to predict whether or not the defendant would pose a continuing threat to society. Astonishingly, the majority found the psychiatric evidence satisfactory because "Neither petitioner nor the [American Psychiatric Association] suggest that psychiatrists are always wrong with respect to the future dangerousness, only most of the time" (*Barefoot v. Estelle*, 1983, p. 901; emphasis added). In *Barefoot* the court accepted a form of social science "evidence" that was strongly repudiated by the scientific community. The

1 Of course, if predictions of dangerousness were "always wrong" they would be a perfect (negative) predictor of future behavior – which would be a strong argument for admitting them.
“common sense” of the Justices seems to be that predicting the violent potential of capital defendants is sufficiently important for capital sentencing that if a legislature wants juries to do it, they must be allowed to, even when the acknowledged experts on the issue say it cannot be done.

A more cynical view is that drawing jurors’ attention to future dangerousness is one of the most effective methods of getting them to return a verdict of death. Research suggests that one of the strongest justifications for the death penalty in the minds of the public in general and of capital jurors in particular is that it guarantees that the defendant will never kill again (Bowers & Steiner, 1999; Eisenberg & Wells, 1993; Gross, 1998). The “common knowledge” of prosecutors, courts, and legislatures is consistent with this research, and because they favor death sentences, they may want to reinforce the image of the dangerous killer out on the streets and eager to kill again. Unfortunately, this strategy is fundamentally misleading. In almost every state a capital defendant who is sentenced to life imprisonment is ineligible for parole, will die in prison, and is no danger to private citizens. But sometimes this information is kept from the jury, leaving them to believe that if he is not executed the murderer will be released—perhaps in 20 years, perhaps in 10, maybe sooner (Bowers & Steiner, 1999; Eisenberg & Wells, 1993). In Simmons v. South Carolina (1994) the Supreme Court held that when a capital defendant’s “future dangerousness is at issue, and state law prohibits the defendant’s release on parole, due process requires that the sentencing jury be informed that the defendant is parole ineligible” (p. 154). South Carolina attempted to evade this rule in a subsequent case (Kelly v. South Carolina, 2002), arguing that their evidence that Kelly was a “butcher” was relevant to character, not to dangerousness, but lost again in the Supreme Court. Judges seem to understand that evidence of future dangerousness is powerfully persuasive to juries, but although the Supreme Court has attempted to limit the influence of future dangerousness in cases in which it is fundamentally misleading, it has not reconsidered its position on the overall appropriateness of such evidence.

Racial discrimination

Broad, vague legal standards and juror misunderstanding of the law not only lead to decisions that are arbitrary, but also provide an opportunity for racial prejudice. As early as the mid-1960s, Marvin Wolfgang conducted a study that showed that across the South death sentences
for rape were about 18 times more likely when the defendant was Black and the victim was White than in other racial combinations (Maxwell v. Bishop, 1970; Wolfgang & Reidel 1973). Uneasiness about racial discrimination has haunted the death penalty ever since. For the most part it has lurked in the background, appearing in concurring and dissenting opinions, but never squarely addressed – until the case of McCleskey v. Kemp in 1987. The issue in McCleskey was discrimination in the entire process that led to death sentences – not just jury discrimination, but discrimination in the decisions to arrest, to go forward with the prosecution, and to try the case as a capital case.

By 1987 several studies had examined whether the new guided discretion statutes enacted after Gregg had managed to eliminate racial discrimination in capital sentencing. All showed that defendants who killed White victims were much more likely to receive the death penalty than those who killed Black victims. There was some evidence that Black defendants were more likely to be sentenced to death than White defendants, but it is less powerful and less consistent, and the effect often disappears when the studies control for legitimate aggravating factors such as killing a stranger or committing the murder in the course of another felony. The race of the victim, however, showed up as an independent factor even in the best-controlled studies. Regardless of other factors, killing a White person is more likely to result in a death sentence than killing a Black person.

McCleskey was a Georgia case, and the centerpiece of the case was a comprehensive landmark study of Georgia’s homicide prosecutions (Baldus, Pulaski, and Woodworth, 1983). In this study David Baldus and his colleagues analyzed data on more than 400 variables in over 1,000 Georgia homicide cases and found that nothing but race could explain the discrimination against defendants who killed White victims. The study also found that Blacks who killed Whites were more likely to be sentenced to death than Whites who killed Whites in some subsets of the cases. Similar patterns were found in less comprehensive studies in Florida (Radelet, 1981; Radelet & Pierce, 1985), South Carolina (Jacoby & Paternoster, 1982), and eight other states (Gross & Mauro, 1984), and in many studies across the country in the years since McCleskey.

The Supreme Court rejected the claim of discrimination in McCleskey by 5 to 4, with Justice Powell writing for the majority. Powell begins by assuming the accuracy of Baldus’s findings, but, all the same, goes on to say that: “At most the Baldus study indicates a disparity that appears to correlate with race” (McCleskey, 1987, p. 312), and “we decline to assume that what is unexplained is invidious” (p. 313). In
fact, Baldus and his colleagues had found a powerful correlation between capital sentencing and race of victim that could not be explained by any combination of other variables that they considered, or that was offered by any other researcher or interested party. Powell’s central argument, however, is that – at least in this context\(^2\) – statistical evidence is inherently “insufficient to support an inference that any of the [individual] decisionmakers in McCleskey’s case acted with discriminatory purpose” (p. 297). That holding created an essentially insuperable barrier to proof of racial discrimination in capital sentencing.

Years later, after he had retired, Justice Powell told his biographer that his decision in *McCleskey* was the only one in his career that he truly regretted, that his understanding of statistics “ranges from limited to zero,” and that he had come to believe that the death penalty ought to be abolished (Jeffries, 1994, p. 451). Given the importance of the issue of racial discrimination, abolition would have been a likely consequence if the Court had sided with McCleskey. In 1987 this was not an acceptable outcome to Justice Powell or his colleagues in the majority on the Supreme Court.

Death qualification

Death qualification is the practice of excluding all citizens who adamantly oppose the death penalty from serving as jurors in capital cases. Thus in capital cases, unlike all other cases, the jury is made up exclusively of people who would be willing to sentence a person to death. The empirical question is whether such a “death-qualified” jury is more than usually likely to decide that the defendant is guilty.

As with race bias, the common knowledge of prosecutors and judges is that the answer is yes. Jurors who favor the death penalty tend to favor the arguments of the prosecution, and prosecutors have admitted to using death qualification in order to increase the likelihood of conviction (Oberer, 1961). As with race bias, the empirical data coincide perfectly with legal intuition: Death-qualified jurors are more favorable to the prosecution and more likely to vote for guilt than the citizens who are excluded from capital juries. Conceptually, this creates a problem of discrimination of a different sort than the

\(^2\) In other contexts, such as discrimination in employment in and in the composition of juries, the Court has accepted statistical evidence as proof of discrimination (Gross & Mauro, 1989, pp. 173–191.)
one we have just considered. The issue is not discrimination between different categories of capital defendants, but against capital defendants by comparison to other criminal defendants.

In 1968, when the issue first came before the Supreme Court (*Witherspoon v. Illinois*), there was little research, and the Court held that the evidence was insufficient to demonstrate that death-qualified juries were biased towards conviction. However, the Court acknowledged that the question was an empirical one, and that they might be persuaded in the future by a stronger empirical record. By the time the issue came back to the Supreme Court in 1986, in the case *Lockhart v McCree*, the evidence consisted of 15 empirical studies, conducted over a 30-year period, using samples from different regions and demographic groups, and multiple methods: attitude, surveys, simulations, and interviews with actual jurors. All of the research converged on the conclusion that death-qualified jurors were more favorable to the prosecution and more likely to vote guilty than the citizens who were excluded, and that jury deliberation did not erase the bias (Ellsworth, 1988). As in *McCleskey*, the Court was faced with a substantial, consistent, and highly persuasive body of research that pointed to a conclusion opposite to the one the majority wanted to reach.

Justice Rehnquist, writing for the majority, first criticized the research, and then declared that even if the studies proved what they claimed, death qualification would still be constitutional. The critique was fairly extensive. Rehnquist examined the 15 studies one by one, finding a flaw in each and discarding it from the set, until only one study was left (Cowan, Thompson, & Ellsworth, 1984) – and “Surely,” he concluded, “a constitutional decision should not be based on a lone study” (*Lockhart*, 1986, p. 173). Although the APA amicus brief, the lower court cases, and Justice Marshall’s dissent clearly described the concept of convergent validity, the majority managed to overlook it. As in *McCleskey*, the majority opinion also rejected McCree’s claim because he had failed to demonstrate that the particular jury that had tried him included jurors who were biased towards guilt. Rehnquist’s general position is that if a jury contains 12 impartial individuals, it is an impartial jury, even if large segments of the population are excluded from serving. In effect, the Court decided that the constitutionality of death-qualification was not an empirical question not all, despite the holding in *Witherspoon*: “We do not ultimately base our decision today on the invalidity of the lower courts’ ‘factual’ findings” (*Lockhart*, 1986, p. 1762, n.3). In capital cases it is constitutional to decide guilt or innocence with juries that are biased toward guilt.
Conclusion: Retribution, Innocence, and Public Opinion

In *Lockhart v McCree* (1986) the Court held that empirical research was not relevant to the constitutionality of death qualification. In *McCleskey v. Kemp* (1987) the Court held that statistical data could not prove race discrimination in capital sentencing. Many people – members of the public and members of the judiciary – question whether empirical research is relevant to the constitutionality of capital punishment at all. For many supporters and opponents of the death penalty the issue is a moral one, not an empirical one. Supporters believe that death is appropriate as a form of retribution: just punishment for heinous murderers, a life for a life (Ellsworth & Gross, 1994). Opponents believe just as strongly that it is morally wrong for the State to kill. The empirical research we have reviewed in this chapter has been evaluated in the contexts of these moral principles. Over the past 35 years – the period when the constitutionality of various aspects of the death penalty has been at issue – moral beliefs about the death penalty have been intense and passionate. The Supreme Court of course was aware of the public’s fervent support for the death penalty during most of this period, support that was compatible with the Court majority’s own ideological commitment to the constitutionality of capital punishment. In this context, dealing with the empirical data on public opinion and evolving standards of decency was smooth sailing: The public favored the death penalty, the Court favored it, and the Court’s decisions matched the data.

On the other issues, the empirical research revealed serious problems in the actual practice of capital punishment in America, in particular the continuing existence of the arbitrariness and discrimination that were supposedly eliminated by the guided discretion standards endorsed in *Gregg v. Georgia* (1976). Challenges to the death penalty have relied heavily on social science data, and the empirical arguments have been exceptionally sophisticated and thorough, particularly on the issues of racial discrimination and death qualification. The Court’s attempts to evade research with implications that they did not like has contributed to what Gross and Mauro (1989) called “an expanding swamp of uncertain rules and confusing opinions” (p. 7).

This is not a matter of judicial “common knowledge”, but of values and ideology. Most lawyers and judges know that death-qualification produces juries that favor the prosecution, and that the criminal justice system is not color-blind. Nor is it a matter of a general reluctance to
consider social science data. In other contexts involving race, such as employment, education, and the use of peremptory challenges, the Court has found discrimination based on weaker data than those presented in *McCleskey*. In other cases involving the fairness of the jury, such as the cases on jury size, they have relied on weaker evidence than that presented in *Lockhart*. They have generally avoided considering the research that shows that juries do not understand the legal standards for sentencing a person to death, but have allowed the jury to hear inaccurate and misleading expert testimony on future dangerousness. The Court's use of social science evidence is related less to its compatibility with scientific standards than to its compatibility with a decision already reached. The death penalty is morally justified, and data that threaten this premise are circumvented.

There is a growing uneasiness, however. The moral principle of just retribution only makes sense if the people we are executing are heinous murderers – sane, intellectually competent people who choose evil over good. Doubts that our system of capital punishment has been effective in singling out these morally depraved killers are reflected in the recent Supreme Court decisions outlawing the execution of mentally retarded defendants and juveniles.

In the past decade, however, a much more important challenge has emerged. Until recently, most people believed that the safeguards built into our system of capital punishment – trial by jury, a seemingly endless series of appeals, executive clemency – were so complete that the possibility of executing someone by mistake was negligible. If anything, people believed that these protections, particularly the appellate process, were greater than necessary and should be curtailed. The near infallibility of the system was common knowledge. That view has fundamentally changed.

Between 1973 and January 2007, 122 convicted capital defendants have been exonerated and released from death row in the United States (Death Penalty Information Center). In 2006 new investigations produced strong evidence that several executed defendants were innocent (Shaw, 2006). In addition, hundreds of innocent noncapital defendants have been released in the past two decades, mostly rape defendants who were cleared by DNA evidence (Gross, Jacoby, Matheson, Montgomery, & Patil, 2005). The problem of false convictions in general – and false capital convictions in particular – has received widespread attention in the news media, and has been the subject of many popular books, television shows, and movies. It has generated a growing movement to reform police investigative procedures. It led to a moratorium on executions in Illinois in 1999, and a more recent one in
New Jersey in 2006. It is the most important development in the American criminal justice system in the past 25 years.

The Supreme Court might someday consider a direct challenge to the death penalty on the grounds that it leads to executions of innocent people. One lower court did hold the death unconstitutional on that basis, but it was reversed on appeal (United States v. Quinones, 2002). More likely the issue will affect the Court indirectly.

Concerns about the execution of innocent people have already weakened public support for capital punishment. This is the engine that drove down support in the late 1990s. Most Americans already believe, abstractly, that innocent defendants have been put to death in recent years (Gross & Ellsworth, 2003). If specific false executions are proven by DNA, or other incontrovertible evidence, support will drop further. If the public comes to oppose the death penalty, legislatures and courts will follow suit, as they did in the case of the execution of the mentally retarded. The "evolving standards of decency" standard puts public acceptance at the center of the constitutional debate over capital punishment. For years the public enthusiastically accepted capital punishment, and the Supreme Court’s decisions were consistently pro-death penalty; recently, public support has weakened, and the Court has ruled against the death penalty for particular kinds of defendants. In the future, if a majority of the American public comes to oppose the death penalty altogether, it will be history.

REFERENCES


NORC General Social Survey, April, 1976.


