

Module 2: The (Un)reliability of Clinical and Actuarial Predictions of (Dangerous) Behavior

I would not say that the future is necessarily less predictable than the past. I think the past was not predictable when it started.

– Donald Rumsfeld

Abstract: The prediction of dangerous and/or violent behavior is important to the conduct of the United States justice system in making decisions about restrictions of personal freedom such as preventive detention, forensic commitment, or parole. This module discusses behavioral prediction both through clinical judgement as well as actuarial assessment. The general conclusion drawn is that for both clinical and actuarial prediction of dangerous behavior, we are far from a level of accuracy that could justify routine use. To support this later negative assessment, two topic areas are discussed at some length: 1) the MacArthur Study of Mental Disorder and Violence, including the actuarial instrument developed as part of this project (the Classification of Violence Risk (COVR)), along with all the data collected that helped develop the instrument; 2) the Supreme Court case of *Barefoot v. Estelle* (1983) and the American Psychiatric Association “friend of the court” brief on the (in)accuracy of clinical prediction for the commission of future violence. An elegant Justice Blackmun dissent is given in its entirety that contradicts the majority decision that held: There is no merit to petitioner’s argument that psychiatrists, individually and as a group, are incompetent to predict with an acceptable degree of reliability that a particular criminal will

commit other crimes in the future, and so represent a danger to the community.

Contents

1	Introduction	2
2	Clinical Prediction	4
3	Actuarial Prediction	6
4	Barefoot v. Estelle (1983)	12
5	Appendix: Continuation of the American Psychiatric Association, <i>Amicus Curiae</i> Brief: Barefoot v. Estelle	29
6	Appendix: Opinion and Dissent in the U.S. Supreme Court, Barefoot v. Estelle (Decided, July 6, 1983)	31

1 Introduction

An ability to predict and treat dangerous or violent behavior in criminal offenders is important to the administration of the criminal justice system in the United States. This prediction might be in the context of preventive detentions, parole decisions, forensic commitments, or other legal forms of restriction on personal liberty. Behavioral prediction might rely on clinical judgement (usually through trained psychologists or other medically versed individuals) or by actuarial

(statistical) assessments. In any case, concern should be on the reliability of such predictions, and more pointedly, on the state of clinical and actuarial prediction. So, the question: are we at such a level of predictive accuracy that as a society we can justify the necessary false positives that would inappropriately restrict the personal liberty of those who would prove to be neither dangerous or violent. Unfortunately, the conclusion reached in this module is that for both clinical or actuarial prediction of dangerous behavior, we are quite far from a level that could sanction routine use.

Evidence on prediction accuracy can typically be presented in the form of a 2×2 contingency table defined by a cross-classification of individuals according to the events A and \bar{A} (whether the person proved dangerous (A) or not (\bar{A})); and B and \bar{B} (whether the person was predicted to be dangerous (B) or not (\bar{B})):

Prediction:

B (dangerous)

\bar{B} (not dangerous)

Outcome (Post-Prediction):

A (dangerous)

\bar{A} (not dangerous)

A generic 2×2 table presenting the available evidence on prediction accuracy might then be given in the following form (arbitrary cell frequencies are indicated using the appropriate subscript combinations of A and \bar{A} and B and \bar{B}):

		Outcome		row sums
		A (dangerous)	\bar{A} (not dangerous)	
Prediction	B (dangerous)	n_{BA}	$n_{B\bar{A}}$	n_B
	\bar{B} (not dangerous)	$n_{\bar{B}A}$	$n_{\bar{B}\bar{A}}$	$n_{\bar{B}}$
column sums		n_A	$n_{\bar{A}}$	n

2 Clinical Prediction

The 2×2 contingency table given immediately below illustrates the poor prediction of dangerous behavior when based on clinical assessment. These data are from Kozol, Boucher, and Garofalo (1972), “The Diagnosis and Treatment of Dangerousness”:

		Outcome		row sums
		A (dangerous)	\bar{A} (not dangerous)	
Prediction	B (dangerous)	17	32	49
	\bar{B} (not dangerous)	31	355	386
column sums		48	387	435

For these data, 2 out of 3 predictions of “dangerous” are wrong ($.65 = 32/49$ to be precise). This is the source of the “error rate” reported in the Supreme Court opinion in *Barefoot v. Estelle* (1983), discussed at great length later. Also, 1 out of 12 predictions of “not dangerous” are wrong ($.08 = 31/386$).

In his dissent opinion in the *Barefoot v. Estelle* case, Justice Blackmun quotes the American Psychiatric Association *amicus curiae* brief as follows: “[the] most that can be said about any individual is that a history of past violence increases the probability that future violence will occur.” In other words, the best we can say is that “past violence” (B) is facilitative of “future violence” (A) but the error in that prediction can be very large as it is here for the Kozol

et al. data: $P(A|B) = \frac{17}{49} = .35$ is greater than $P(A) = \frac{48}{435} = .11$. But this implies that 2 out of 3 such predictions of “dangerous” are wrong (or, 1 out of 3 are correct). To us, the accuracy of these behavioral predictions is insufficient to justify any incarceration policy based on them; the same conclusion will hold for the type of actuarial prediction of “dangerous” discussed in the section to follow.

In Module 4 on diagnostic testing, the Meehl and Rosen (1955) notion of “clinical efficiency” is formally discussed, or when a diagnostic test is more accurate than just predicting using base rates. For these data, prediction by base rates would be to say everyone will be “not dangerous” because the number of people who are “not dangerous” (387) is larger than the number of people who are “dangerous” (48). Here, we would be correct in our predictions 89% of the time ($.89 = 387/435$). Based on clinical prediction, we would be correct a *smaller* 86% percentage of the time ($.86 = (17 + 355)/435$). So, according to the Meehl and Rosen characterization, clinical prediction is *not* “clinically efficient” because one can do better by just predicting according to base rates.

In commenting on the Kozol, et al. study, Monahan (1973) takes issue with the article’s principal conclusion that “dangerousness can be reliably diagnosed and effectively treated” and notes that it “is, at best, misleading and is largely refuted by their own data.” Monahan concludes his critique with the following quotation from Wenk, Robison, and Smith (1972):

Confidence in the ability to predict violence serves to legitimate intrusive types of social control. Our demonstration of the *futility* of such prediction should have consequences as great for the protection of individual liberty as a demonstration of the utility of violence prediction would have for the

protection of society. (p. 402)

3 Actuarial Prediction

Paul Meehl in his iconic 1954 monograph, *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*, created quite a stir with his convincing demonstration that mechanical methods of data combination, such as multiple regression, outperform (expert) clinical prediction. The enormous amount of literature produced since the appearance of this seminal contribution has uniformly supported this general observation; similarly, so have the extensions suggested for combining data in ways other than by multiple regression, for example, by much simpler unit weighting schemes, or those using other prior weights. It appears that individuals who are conversant in a field are better at selecting and coding information than they are at actually integrating it. Combining such selected information in a more mechanical manner will generally do better than the person choosing such information in the first place.¹

¹A 2005 article by Robyn Dawes in the *Journal of Clinical Psychology* (61, 1245–1255) has the intriguing title “The Ethical Implications of Paul Meehl’s Work on Comparing Clinical Versus Actuarial Prediction Methods.” Dawes’ main point is that given the overwhelming evidence we now have, it is unethical to use clinical judgment in preference to the use of statistical prediction rules. We quote from the abstract:

Whenever statistical prediction rules . . . are available for making a relevant prediction, they should be used in preference to intuition. . . . Providing service that assumes that clinicians “can do better” simply based on self-confidence or plausibility in the absence of evidence that they can actually do so is simply unethical. (p. 1245)

The MacArthur Study of Mental Disorder and Violence

The MacArthur Research Network on Mental Health and the Law was created in 1988 by a major grant to the University of Virginia from the John D. and Catherine T. MacArthur Foundation. The avowed aim of the Network was to construct an empirical foundation for the next generation of mental health laws, assuring the rights and safety of individuals and society. New knowledge was to be developed about the relation between the law and mental health; new assessment tools were to be developed along with criteria for evaluating individuals and making decisions affecting their lives. The major product of the Network was the MacArthur Violence Risk Assessment Study; its principal findings were published in the very well-received 2001 book, *Rethinking Risk Assessment: The MacArthur Study of Mental Disorder and Violence* (John Monahan, et al., Oxford University Press). More importantly for us (and as a source of illustrations used throughout), the complete data set (on 939 individuals over 134 risk factors) is available on the web.

The major analyses reported in *Rethinking Risk Assessment* are based on constructed classification trees; in effect, these are branching decision maps for using risk factors to assess the likelihood that a particular person will commit violence in the future. All analyses were carried out with an SPSS classification-tree program, called CHAID, now a rather antiquated algorithm (the use of this method without a modern means of cross-validation most likely led to the overfitting difficulties to be discussed shortly). Moreover, these same classification tree analyses have been incorporated into a proprietary software product called the Classification of Violence Risk (COVR);

it is available from the Florida-based company PAR (Psychological Assessment Resources). The program is to be used in law enforcement/mental health contexts to assess “dangerousness to others,” a principal standard for inpatient or outpatient commitment or commitment to a forensic hospital.

One of the authors of the current module taught a class entitled Advanced Multivariate Methods for the first time in the Fall of 2011, with a focus on recent classification and regression tree methods developed over the last several decades and implemented in the newer environments of Matlab and R (but not in SPSS). These advances involve “random forests,” “bootstrap aggregation (bagging),” “boosting algorithms,” “ensemble methods,” and a number of techniques that avoid the dangers of overfitting and allow several different strategies of internal cross-validation. To provide interesting projects for the class to present, a documented data set was obtained from the statistician on the original MacArthur study; this was a more transparent packaging of the data already available on the web. This “cleaned-up” data set could provide a direct replication of the earlier SPSS analyses (with CHAID); but in addition and more importantly, all the “cutting-edge” methods could now be applied that were unavailable when the original MacArthur study was completed in the late 1990s. At the end of the semester, five subgroups of the graduate students in the class reported on analyses they did on the MacArthur data set (each also had a different psychological test battery to focus on, for example, Brief Psychiatric Rating Scale, Novaco Anger Scale, Novaco Provocation Inventory, Big Five Personality Inventory, Psychopathy Checklist (Screening Version)). Every one of the talks essentially reported a “wash-out” when cross-validation and predic-

tion was the emphasis as opposed to just fitting the classification structures. We could not do better than just predicting with base rates. This was a first indication that the prediction of “dangerousness” was possibly not as advanced as the MacArthur Network might have us believe.

The second major indication of a difficulty with prediction even with the newer MacArthur assessment tools was given by a close read of the first small cross-validation study done to justify this actuarial software COVR (mentioned earlier): “An Actuarial Model of Violence Risk Assessment for Persons with Mental Disorders” (John Monahan, et al., *Psychiatric Services*, 2005, 56, 810–815). The abstract of this article is given below:

Objectives: An actuarial model was developed in the MacArthur Violence Risk Assessment Study to predict violence in the community among patients who have recently been discharged from psychiatric facilities. This model, called the multiple iterative classification tree (ICT) model, showed considerable accuracy in predicting violence in the construction sample. The purpose of the study reported here was to determine the validity of the multiple ICT model in distinguishing between patients with high and low risk of violence in the community when applied to a new sample of individuals.

Methods: Software incorporating the multiple ICT model was administered with independent samples of acutely hospitalized civil patients. Patients who were classified as having a high or a low risk of violence were followed in the community for 20 weeks after discharge. Violence included any battery with physical injury, use of a weapon, threats made with a weapon in hand, and sexual assault.

Results: Expected rates of violence in the low- and high-risk groups were 1 percent and 64 percent, respectively. Observed rates of violence in the low- and high-risk groups were 9 percent and 35 percent, respectively, ... These findings may reflect the “shrinkage” expected in moving from construction to validation samples.

Conclusions: The multiple ICT model may be helpful to clinicians who are faced with making decisions about discharge planning for acutely hospitalized civil patients.

John Monahan in his influential NIMH monograph, *The Clinical Prediction of Violent Behavior* (1977), observed that, even allowing for possible distortions in the research data, “it would be fair to conclude that the best clinical research currently in existence indicates that psychiatrists and psychologists are accurate in no more than one out of three predictions of violent behavior over a several year period among institutionalized populations that had both committed violence in the past (and thus had high base rates for it) and who were diagnosed as mentally ill.” In other words, predictions that someone will be violent (and therefore subject to detention) will be wrong two out of three times. With such a dismal record of clinical prediction, there were high expectations that the MacArthur Network could produce a much better (actuarial) instrument in COVR. Unfortunately, that does not appear to be the case. The figures of 64% and 35% given in the abstract suggest two conclusions: in the training sample, the error in predicting dangerousness is 1 out of 3; whether this shows “considerable accuracy in predicting violence in the construction sample” is highly questionable, even assuming the inflated value is correct. The cross-classified proportion of 35% gives the error of being wrong in prediction of dangerousness as 2 out of 3. It is quite an understatement to then say: “These findings may reflect the “shrinkage” expected in moving from construction to validation samples.” What it reflects is that actuarial prediction of violence is exactly as bad as clinical prediction. This may be one of the only (if not the only) examples from the behavioral science literature in

which actuarial prediction doesn't do better than clinical prediction.

The complete 2×2 table from the COVR validation study follows:

		Outcome		row sums
		A (dangerous)	\bar{A} (not dangerous)	
Prediction	B (dangerous)	19	36	55
	\bar{B} (not dangerous)	9	93	102
column sums		28	129	157

As noted above, a high prediction of “dangerous” is wrong 65% (= 36/55) of the time. A prediction of “not dangerous” is incorrect 9% (= 9/102) of the time (again, this is close to the 1 out of 12 incorrect predictions of “not dangerous” typically seen for purely clinical predictions). The accuracy or “hit-rate” is $(10 + 93)/157 = .71$. If everyone were predicted to be nondangerous, we would be correct 129 out of 157 times, the base rate for \bar{A} : $P(\bar{A}) = 129/157 = .82$. Obviously, the accuracy of prediction using base rates (82%) is better than for the COVR (71%), making the COVR not “clinically efficient” according to the Meehl and Rosen terminology.

We give two more examples from the MacArthur data set mentioned earlier that involve the variables of “prior arrest” or “prior violence” as diagnostic “tests” in their own right. The first adopts prior arrest as a diagnostic “test”: dangerous—one or more prior arrests (B); not dangerous—no prior arrests (\bar{B}).

		Outcome		row sums
		A (dangerous)	\bar{A} (not dangerous)	
Prediction	B (dangerous)	103	294	397
	\bar{B} (not dangerous)	39	354	393
column sums		142	648	790

Here, 3 out of 4 predictions of “dangerous” are wrong ($.74 = 294/397$); 1 out of 10 predictions of “not dangerous” are wrong ($.10 = 39/393$).

The accuracy of the test is $(103 + 354)/790 = .50$, and the correctness of prediction by base rates is $648/790 = .82$; thus, “prior arrest” is not a clinically efficient “test.”

The second example uses prior violence as a diagnostic “test”: dangerous—prior violence (B); not dangerous—no prior violence (\bar{B}).

		Outcome		row sums
		A (dangerous)	\bar{A} (not dangerous)	
Prediction	B (dangerous)	48	106	154
	\bar{B} (not dangerous)	128	657	785
column sums		176	763	939

In this case, 7 out of 10 predictions of “dangerous” are wrong ($.69 = 106/154$); 1 out of 6 predictions of “not dangerous” are wrong ($.16 = 128/785$). The accuracy of the test, $(48 + 657)/939 = .75$, is less than the the correctness of prediction by base rates: $763/939 = .81$; thus, “prior violence” is not a clinically efficient “test.”

4 Barefoot v. Estelle (1983)

The present discussion on probabilistic reasoning concerns the unreliability of clinical and actuarial behavioral prediction, particularly for violence; the particular section to follow includes two extensive redactions in appendices: one is the majority opinion in the Supreme Court case of *Barefoot v. Estelle* (1983) and an eloquent Justice Blackmun dissent; the second is an *amicus curiae* brief in this same case from the American Psychiatric Association on the accuracy of clinical prediction of future violence. Both of these documents are detailed, self-explanatory, and highly informative about our current

lack of ability to make clinical assessments that lead to accurate and reliable predictions of future behavior. To set the background for the *Barefoot v. Estelle* case, the beginning part of the *amicus curiae* brief follows; a redaction of the remainder of the brief, as already noted, is given in an appendix at the end of the chapter.

Brief for American Psychiatric Association as *Amicus Curiae* Supporting Petitioner, *Barefoot v. Estelle*

Petitioner Thomas A. Barefoot stands convicted by a Texas state court of the August 7, 1978 murder of a police officer—one of five categories of homicides for which Texas law authorizes the imposition of the death penalty. Under capital sentencing procedures established after this Court’s decision in *Furman v. Georgia*, the “guilt” phase of petitioner’s trial was followed by a separate sentencing proceeding in which the jury was directed to answer three statutorily prescribed questions. One of these questions—and the only question of relevance here—directed the jury to determine: whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society. The jury’s affirmative response to this question resulted in petitioner being sentenced to death.

The principle evidence presented to the jury on the question of petitioner’s “future dangerousness” was the expert testimony of two psychiatrists, Dr. John T. Holbrook and Dr. James Grigson, both of whom testified for the prosecution. Petitioner elected not to testify in his own defense. Nor did he present any evidence or testimony, psychiatric or otherwise, in an attempt to rebut the state’s claim that he would commit future criminal acts of violence.

Over defense counsel’s objection, the prosecution psychiatrists were permitted to offer clinical opinions regarding petitioner, including their opinions on the ultimate issue of future dangerousness, even though they had not performed a psychiatric examination or evaluation of him. Instead, the critical psychiatric testimony was elicited through an extended hypothetical question propounded by the prosecutor. On the basis of the assumed facts stated in the hypothetical, both Dr. Holbrook and Dr. Grigson gave essentially the same testimony.

First, petitioner was diagnosed as a severe criminal sociopath, a label variously defined as describing persons who “lack a conscience,” and who “do things which serve their own purposes without regard for any consequences or outcomes to other people.” Second, both psychiatrists testified that petitioner would commit criminal acts of violence in the future. Dr. Holbrook stated that he could predict petitioner’s future behavior in this regard “within reasonable psychiatric certainty.” Dr. Grigson was more confident, claiming predictive accuracy of “one hundred percent and absolute.”

The prosecutor’s hypothetical question consisted mainly of a cataloguing of petitioner’s past antisocial behavior, including a description of his criminal record. In addition, the hypothetical question contained a highly detailed summary of the prosecution’s evidence introduced during the guilt phase of the trial, as well as a brief statement concerning petitioner’s behavior and demeanor during the period from his commission of the murder to his later apprehension by police.

In relevant part, the prosecutor’s hypothetical asked the psychiatrists to assume as true the following facts: First, that petitioner had been convicted of five criminal offenses—all of them nonviolent, as far as the record reveals—and that he had also been arrested and charged on several counts of sexual offenses involving children. Second, that petitioner had led a peripatetic existence and “had a bad reputation for peaceful and law abiding citizenship” in each of eight communities that he had resided in during the previous ten years. Third, that in the two-month period preceding the murder, petitioner was unemployed, spending much of his time using drugs, boasting of his plans to commit numerous crimes, and in various ways deceiving certain acquaintances with whom he was living temporarily. Fourth, that petitioner had murdered the police officer as charged, and that he had done so with “no provocation whatsoever” by shooting the officer in the head “from a distance of no more than six inches.” And fifth, that subsequent to the murder, petitioner was observed by one witness, “a homosexual,” who stated that petitioner “was not in any way acting unusual or that anything was bothering him or upsetting him . . .”

Testimony of Dr. Holbrook:

Dr. Holbrook was the first to testify on the basis of the hypothetical ques-

tion. He stated that the person described in the question exhibited “probably six or seven major criterias (sic) for the sociopath in the criminal area within reasonable medical certainty.” Symptomatic of petitioner’s sociopathic personality, according to Dr. Holbrook, was his consistent “antisocial behavior” from “early life into adulthood,” his willingness to take any action which “serves [his] own purposes” without any regard for the “consequences to other people,” and his demonstrated failure to establish any “loyalties to the normal institutions such as family, friends, politics, law or religion.”

Dr. Holbrook explained that his diagnosis of sociopathy was also supported by petitioner’s past clinical violence and “serious threats of violence,” as well as an apparent history of “escaping or running away from authority” rather than “accepting a confrontation in the legal way in a court of law.” And finally, Dr. Holbrook testified that petitioner had shown a propensity to “use other people through lying and manipulation . . . ” According to Dr. Holbrook, by use of such manipulation the sociopath succeeds in “enhancing [his] own ego image . . . It makes [him] feel good.”

After stating his diagnosis of sociopathy, Dr. Holbrook was asked whether he had an “opinion within reasonable psychiatric certainty as to whether or not there is a probability that the Thomas A. Barefoot in that hypothetical will commit criminal acts of violence in the future that would constitute a continuing threat to society?” Without attempting to explain the implied clinical link between his diagnosis of petitioner and his prediction of future dangerousness, Dr. Holbrook answered simply: “In my opinion he will.”

Testimony of Dr. Grigson:

On the basis of the prosecutor’s hypothetical question, Dr. Grigson diagnosed petitioner as “a fairly classical, typical, sociopathic personality disorder” of the “most severe category.” The most “outstanding characteristic” of persons fitting this diagnosis, according to Dr. Grigson, is the complete “lack of a conscience.” Dr. Grigson stated that such persons “repeatedly break the rules, they con, manipulate and use people, [and] are only interested in their own self pleasure [and] gratification.”

Although Dr. Grigson testified that some sociopathic individuals do not pose a continuing threat to society, he characterized petitioner as “your most severe sociopath.” Dr. Grigson stated that persons falling into this special

category are “the ones that . . . have complete disregard for another human being’s life.” Dr. Grigson further testified that “there is not anything in medicine or psychiatry or any other field that will in any way at all modify or change the severe sociopath.”

The prosecutor then asked Dr. Grigson to state his opinion on the ultimate issue—“whether or not there is a probability that the defendant . . . will commit criminal acts of violence that would constitute a continuing threat to society?” Again, without explaining the basis for his prediction or its relationship to the diagnosis of sociopathy, Dr. Grigson testified that he was “one hundred percent” sure that petitioner “most certainly would” commit future criminal acts of violence. Dr. Grigson also stated that his diagnosis and prediction would be the same whether petitioner “was in the penitentiary or whether he was free.”

The psychiatrist featured so prominently in the opinions for *Barefoot v. Estelle* and the corresponding American Psychiatric Association *amicus* brief, James Grigson, played the same role repeatedly in the Texas legal system. For over three decades before his retirement in 2003, he testified when requested at death sentence hearings to a high certainty as to “whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society.” An affirmative answer by the sentencing jury imposed the death penalty automatically, as it was on Thomas Barefoot; he was executed on October 30, 1984. When asked if he had a last statement to make, he replied:

Yes, I do. I hope that one day we can look back on the evil that we’re doing right now like the witches we burned at the stake. I want everybody to know that I hold nothing against them. I forgive them all. I hope everybody I’ve done anything to will forgive me. I’ve been praying all day for Carl Levin’s wife to drive the bitterness from her heart because that bitterness that’s in her heart will send her to Hell just as surely as any other sin. I’m sorry for everything I’ve ever done to anybody. I hope they’ll forgive me.

James Grigson was expelled in 1995 from the American Psychiatric Association and the Texas Association of Psychiatric Physicians for two chronic ethics violations: making statements in testimony on defendants he had not actually examined, and for predicting violence with 100% certainty. The press gave him the nickname of “Dr. Death.”

Barefoot v. Estelle has another connection to the distinction between actuarial and clinical prediction, and where the former is commonly better than the latter. There is evidence mentioned in the APA brief that actuarial predictions of violence carried out by statistically informed laymen might be better than those of a clinician. This may be due to a bias that psychiatrists might (unsuspectingly) have in overpredicting violence because of the clients they see or for other reasons related to their practice. There is a pertinent passage from the court opinion (not given in our redactions):

That psychiatrists actually may be less accurate predictors of future violence than laymen, may be due to personal biases in favor of predicting violence arising from the fear of being responsible for the erroneous release of a violent individual. . . . It also may be due to a tendency to generalize from experiences with past offenders on bases that have no empirical relationship to future violence, a tendency that may be present in Grigson’s and Holbrook’s testimony. Statistical prediction is clearly more reliable than clinical prediction . . . and prediction based on statistics alone may be done by anyone.

The two psychiatrists mentioned in *Barefoot v. Estelle*, James Grigson and John Holbrook, appeared together repeatedly in various capital sentencing hearings in Texas during the later part of the 20th century. Although Grigson was generally the more outrageous of the two with predictions of absolute certitude based on a sociopath diagnosis, Holbrook was similarly at fault ethically. This pair of psy-

chiatrists of Texas death penalty fame might well be nicknamed “Dr. Death” and “Dr. Doom.” They were both culpable in the famous exoneration documented in the award winning film by Errol Morris, *The Thin Blue Line*. To tell this story, we give the summary of the Randall Dale Adams exoneration from the Northwestern Law School’s Center on Wrongful Convictions (written by Robert Warden with Michael L. Radelet):

Sentenced to death in 1977 for the murder of a police officer in Dallas, Texas, Randall Dale Adams was exonerated as a result of information uncovered by film-maker Errol Morris and presented in an acclaimed 1988 documentary, *The Thin Blue Line*.

Patrolman Robert Wood was shot to death during a traffic stop on November 28, 1976, by sixteen-year-old David Ray Harris, who framed Adams to avoid prosecution himself. Another factor in the wrongful conviction was the surprise—and partly perjured—testimony of three eyewitnesses whose existence had been concealed from the defense until the witnesses appeared in the courtroom. A third factor was a statement Adams signed during interrogation that the prosecution construed as an admission that he had been at the scene of the crime.

The day before the murder, Adams was walking along a Dallas street after his car had run out of gasoline. Harris happened by, driving a stolen car. He offered Adams a ride and the two wound up spending the afternoon and evening together, drinking beer, smoking marijuana, pawning various items Harris had stolen, and going to a drive-in movie theater to watch porn movies. Adams then returned to a motel where he was staying.

Shortly after midnight, Wood and his partner, Teresa Turko, spotted Harris driving a blue car with no headlights. The officers stopped the car and, as Wood approached the driver’s side, Harris shot him five times. Wood died on the spot. As the car sped off, Turko fired several shots, but missed. She did not get a license number. She seemed certain that there was only one person in the car—the driver.

Harris drove directly to his home in Vidor, 300 miles southeast of Dallas. Over the next several days, he bragged to friends that he had “offed a pig”

in Dallas. When police in Vidor learned of the statements, they took Harris in for questioning. He denied having had anything to do with the murder, claiming he had said otherwise only to impress his friends. But when police told him that a ballistics test established that a pistol he had stolen from his father was the murder weapon, Harris changed his story. He now claimed that he had been present at the shooting, but that it had been committed by a hitchhiker he had picked up—Adams.

Adams, an Ohio native working in Dallas, was taken in for questioning. He denied any knowledge of the crime, but he did give a detailed statement describing his activities the day before the murder. Police told him he had failed a polygraph test and that Harris had passed one, but Adams remained resolute in asserting his innocence.

Although polygraph results are not admissible in Texas courts, the results provided some rationale for questioning Harris's story. However, when a police officer is murdered, authorities usually demand the most severe possible punishment, which in Texas, and most other United States jurisdictions, is death. Harris was only sixteen and ineligible for the death penalty; Adams was twenty-seven and thus could be executed.

At trial before Dallas County District Court Judge Don Metcalfe and a jury, Turko testified that she had not seen the killer clearly, but that his hair was the color of Adams's. She also said that the killer wore a coat with a fur collar. Harris had such a coat, but Adams did not.

Adams took the stand and emphatically denied having any knowledge of the crime. But then the prosecution sprang two surprises. The first was the introduction of Adams's purported signed statement, which police and prosecutors claimed was a confession, although it said only—falsely, according to Adams—that when he was in the car with Harris, they had at one point been near the crime scene. The second was the testimony of three purported eyewitnesses whose existence had until then been unknown to the defense. One of these witnesses, Michael Randell, testified that he had driven by the scene shortly before the murder and, in the car that had been stopped by the officers, had seen two persons, one of whom he claimed was Adams. The other two witnesses, Robert and Emily Miller, had happened by at about the same time, but claimed to have seen only one person in the car—Adams.

Because the eyewitnesses were called only to rebut Adams's testimony, prosecutors claimed that Texas law did not require them to inform the defense of their existence before they testified. The weekend after their surprise testimony, however, the defense learned that Emily Miller had initially told police that the man she had seen appeared to be Mexican or a light-skinned African American. When the defense asked to recall the Millers to testify, the prosecution claimed that the couple had left town. In fact, the Millers had only moved from one part of Dallas to another. When the defense asked to introduce Emily Miller's statement, Judge Metcalfe would not allow it. He said it would be unfair to impeach her credibility when she was not available for further examination.

The jury quickly returned a verdict of guilty and turned to sentencing. Under Texas law, in order for Adams to be sentenced to death, the jury was required to determine, among other things, whether there was "beyond a reasonable doubt [a] probability" that he or she would commit future acts of violence. To establish that Adams met that oxymoronic criterion, the prosecution called Dr. James Grigson, a Dallas psychiatrist known as "Dr. Death," and Dr. John Holbrook, former chief of psychiatry for the Texas Department of Corrections.

Although the American Psychiatric Association has said on several occasions that future dangerousness was impossible to predict, Grigson and Holbrook testified that Adams would be dangerous unless executed. Grigson testified similarly in more than 100 other Texas cases that ended in death sentences. After hearing the psychiatrists, Adams's jury voted to sentence him to death. Twenty one months later, at the end of January 1979, the Texas Court of Criminal Appeals affirmed the conviction and death sentence. Judge Metcalfe scheduled the execution for May 8, 1979.

Adams was three days away from execution when United States Supreme Court Justice Lewis F. Powell Jr. ordered a stay. Powell was troubled that prospective jurors with moral qualms about the death penalty had been excluded from service, even though they had clearly stated that they would follow the Texas law.

To most observers—including, initially, Dallas District Attorney Henry Wade (of *Roe v. Wade* fame) the Supreme Court's language meant that

Adams was entitled to a new trial. But a few days later Wade announced that a new trial would be a waste of money. Thus, he said, he was asking Governor Bill Clements to commute Adams's sentence to life in prison. When the governor promptly complied, Wade proclaimed that there now would be no need for a new trial. Adams, of course, thought otherwise, but the Texas Court of Criminal Appeals agreed with Wade. As a result of the governor's action, said the court, "There is now no error in the case."

In March 1985, Errol Morris arrived in Dallas to work on a documentary about Grigson—"Dr. Death." Morris's intent had not been to question the guilt of defendants in whose cases Grigson had testified but only to question his psychiatric conclusions. When Morris met Adams, the focus of the project changed.

Morris learned from Randy Schaffer, a volunteer Houston lawyer who had been working on the case since 1982, that Harris had not led an exemplary life after helping convict Adams. Harris had joined the Army and been stationed in Germany, where he had been convicted in a military court of a series [of] burglaries and sent to prison in Leavenworth, Kansas. A few months after his release, Harris had been convicted in California of kidnapping, armed robbery, and related crimes.

After his release from prison in California, and five months after Morris arrived in Dallas, Harris tried to kidnap a young woman named Roxanne Lockard in Beaumont, Texas. In an effort to prevent the abduction, Lockard's boyfriend, Mark Mays, exchanged gunfire with Harris. Mays was shot to death and Harris was wounded. For the Mays murder—a crime that would not have occurred if Dallas authorities convicted the actual killer of Officer Wood eight years earlier—Harris was sentenced to death.

Meanwhile, Morris and Schaffer discovered that Officer Turko had been hypnotized during the investigation and initially had acknowledged that she had not seen the killer—facts that the prosecution had illegally withheld from the defense. Morris and Schaffer also found that robbery charges against the daughter of eyewitness Emily Miller had been dropped after Miller agreed to identify Adams as Wood's killer. The new information, coupled with the fact that Miller initially had described the killer as Mexican or African American, became the basis for a new trial motion.

In 1988, during a three-day hearing on the motion before Dallas District Court Judge Larry Baraka, Harris recanted. “Twelve years ago, I was a kid, you know, and I’m not a kid anymore, and I realize I’ve been responsible for a great injustice,” Harris told Baraka. “And I felt like it’s my responsibility to step forward, to be a man, to admit my part in it. And that’s why I’m trying to correct an injustice.”

On December 2, 1988, Judge Baraka recommended to the Texas Court of Criminal Appeals that Adams be granted a new trial, and two months later he wrote a letter to the Texas Board of Pardons and Paroles recommending that Adams be paroled immediately. The board refused, but on March 1 the Texas Court of Criminal Appeals unanimously concurred with Baraka that Adams was entitled to a new trial. Three weeks later, Adams was released on his own recognizance, and two days after that, Dallas District Attorney John Vance, who had succeeded Wade, dropped all charges.

Harris was never tried for the murder of Officer Woods. On June 30, 2004, he was executed for the Mays murder.

The *Federal Rules of Evidence* and the admissibility of expert witnesses and scientific data was influenced heavily by the case of *Daubert v. Merrell Dow Pharmaceuticals* (1993) that promulgates the *Daubert* standard for admitting expert testimony in federal courts. The majority opinion in *Daubert* was written by Justice Blackman, the same justice who wrote the dissent in *Barefoot v. Estelle*. The court stated that Rule 702 of the *Federal Rules of Evidence* was the governing standard for admitting scientific evidence in trials held in federal court (and now in most state courts as well). Rule 702, Testimony by Experts, states:

If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles

and methods reliably to the facts of the case.

We give a redaction of that part of the Wikipedia article on *Daubert v. Merrell Dow Pharmaceuticals* devoted to the discussion of the *Daubert* standard governing expert testimony. We doubt that clinical predictions of violence based on a sociopath diagnosis would be admissible under the *Daubert* standard.

The Standard Governing Expert Testimony: Three key provisions of the Rules governed admission of expert testimony in court. The first was *scientific knowledge*. This means that the testimony must be scientific in nature, and that the testimony must be grounded in “knowledge.” Of course, science does not claim to know anything with absolute certainty; science “represents a *process* for proposing and refining theoretical explanations about the world that are subject to further testing and refinement.” The “scientific knowledge” contemplated by Rule 702 had to be arrived at by the scientific method.

Second, the scientific knowledge must *assist the trier of fact* in understanding the evidence or determining a fact in issue in the case. The trier of fact is often either a jury or a judge; but other fact-finders may exist within the contemplation of the federal rules of evidence. To be helpful to the trier of fact, there must be a “valid scientific connection to the pertinent inquiry as a prerequisite to admissibility.” Although it is within the purview of scientific knowledge, knowing whether the moon was full on a given night does not typically assist the trier of fact in knowing whether a person was sane when he or she committed a given act.

Third, the Rules expressly provided that the judge would make the threshold determination regarding whether certain scientific knowledge would indeed assist the trier of fact in the manner contemplated by Rule 702. “This entails a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid and of whether that reasoning or methodology properly can be applied to the facts in issue.” This preliminary assessment can turn on whether something has been tested, whether an idea has been subjected to scientific peer review or published in scientific journals, the rate of error involved in the technique, and even general accep-

tance, among other things. It focuses on methodology and principles, not the ultimate conclusions generated.

The Court stressed that the new standard under Rule 702 was rooted in the judicial process and intended to be distinct and separate from the search for scientific truth. “Scientific conclusions are subject to perpetual revision. Law, on the other hand, must resolve disputes finally and quickly. The scientific project is advanced by broad and wide-ranging consideration of a multitude of hypotheses, for those that are incorrect will eventually be shown to be so, and that in itself is an advance.” Rule 702 was intended to resolve legal disputes, and thus had to be interpreted in conjunction with other rules of evidence and with other legal means of ending those disputes. Cross examination within the adversary process is adequate to help legal decision makers arrive at efficient ends to disputes. “We recognize that, in practice, a gate-keeping role for the judge, no matter how flexible, inevitably on occasion will prevent the jury from learning of authentic insights and innovations. That, nevertheless, is the balance that is struck by Rules of Evidence designed not for the exhaustive search for cosmic understanding but for the particularized resolution of legal disputes.”

As noted in the various opinions and *amicus* brief given in *Barefoot v. Estelle*, the jury in considering whether the death penalty should be imposed, has to answer affirmatively one question: whether there was a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society. The use of the word “probability” without specifying any further size seems odd to say the least, but Texas courts have steadfastly refused to delimit it any further. So, presumably a very small probability of future violence would be sufficient for execution if this small probability could be proved “beyond a reasonable doubt.”

The point of much of the current discussion has been to emphasize that actuarial evidence about future violence involving variables such as age, race, or sex, is all there really is in making such pre-

dictions. More pointedly, the assignment of a clinical label, such as “sociopath,” adds nothing to an ability to predict, and to suggest that it does is to use the worst “junk science,” even though it may be routinely assumed true in the larger society. All we have to rely on is the usual psychological adage that the best predictor of future behavior is past behavior. Thus, the best predictor of criminal recidivism is a history of such behavior, and past violence suggests future violence. The greater the amount of past criminal behavior or violence, the more likely that such future behavior or violence will occur (a behavioral form of a “dose-response” relationship). At its basis, this is statistical evidence of such a likely occurrence and no medical or psychological diagnosis is needed or useful.

Besides the specious application of a sociopath diagnosis to predict future violence, after the Supreme Court decision in *Estelle v. Smith* (1981) such a diagnosis had to be made on the basis of a hypothetical question and not on an actual psychological examination of the defendant. In addition to a 100% incontrovertible assurance of future violence, offering testimony without actually examining a defendant proved to be Grigson’s eventual downfall and one reason for the expulsion from his professional psychiatric societies. This prevention of an actual examination of a defendant by the Supreme Court case, *Estelle v. Smith* (1981), also involved James Grigson. Ernest Smith, indicted for murder, had been examined by Grigson in jail and who determined he was competent to stand trial. In the psychiatric report on Smith, Grigson termed him “a severe sociopath” but gave no other statements as to future dangerousness. Smith was sentenced to death based on the sociopath label given by Grigson. In *Estelle v. Smith* the Supreme Court held that because of the well-

known case of *Miranda v. Arizona* (1966), the state could not force a defendant to submit to a psychiatric examination for the purposes of sentencing because it violated a defendant's Fifth Amendment rights against self-incrimination and the Sixth Amendment right to counsel. Thus, the examination of Ernest Smith was inadmissible at sentencing. From that point on, predictions of violence were made solely on hypothetical questions and Grigson's belief that a labeling as a sociopath was sufficient to guarantee future violence on the part of a defendant, and therefore, the defendant should be put to death.

The offering of a professional psychiatric opinion about an individual without direct examination is an ethical violation of the Goldwater Rule, named for the Arizona Senator who ran for President in 1964 as a Republican. Promulgated by the American Psychiatric Association in 1971, it delineated a set of requirements for communication with the media about the state of mind of individuals. The Goldwater Rule was the result of a special September/October 1964 issue of *Fact:* magazine, published by the highly provocative Ralph Ginzburg. The issue title was "The Unconscious of a Conservative: Special Issue on the Mind of Barry Goldwater," and reported on a mail survey of 12,356 psychiatrists, of whom 2,417 responded: 24% said they did not know enough about Goldwater to answer the question; 27% said he was mentally fit; 49% said he was not. Much was made of Goldwater's "two nervous breakdowns," because such a person should obviously never be President because of a risk of recurrence under stress that might then lead to pressing the nuclear button.

Goldwater brought a \$2 million libel suit against *Fact:* and its publisher, Ginzburg. In 1970 the United States Supreme Court de-

cided in Goldwater’s favor giving him \$1 in compensatory damages and \$75,000 in punitive damages. More importantly, it set a legal precedent that changed medical ethics forever. For an updated discussion of the Goldwater Rule, this time because of the many psychiatrists commenting on the psychological makeup of the former chief of the International Monetary Fund, Dominique Strauss-Kahn, after his arrest on sexual assault charges in New York, see Richard A. Friedman’s article, “How a Telescopic Lens Muddles Psychiatric Insights” (*New York Times*, May 23, 2011).²

References

- [1] Kozol, H. L., Boucher, R. J., & Garofalo, R. (1972). The diagnosis and treatment of dangerousness. *Crime and Delinquency*, 18, 371–392.
- [2] Meehl, P., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, 52, 194–215.

²For a recent and thorough review of the literature on the prediction of dangerous or violent behavior as it relates to the death penalty, see Michael L. Perlin, *Mental Disability and the Death Penalty: The Shame of the States* (2013; Rowman & Littlefield); Chapter 3 is particularly relevant: “Future Dangerousness and the Death Penalty.” A good resource generally for material on the prediction of dangerous behavior and related forensic matters is the Texas Defender Service (www.texasdefender.org), and the publications it has freely available at its web site:

A State of Denial: Texas Justice and the Death Penalty (2000)

Deadly Speculation: Misleading Texas Capital Juries with False Predictions of Future Dangerousness (2004)

Minimizing Risk: A Blueprint for Death Penalty Reform in Texas (2005)

- [3] Monahan, J. (1973). Dangerous offenders: A critique of Kozol et al. *Crime and Delinquency*, 19, 418–420.
- [4] Monahan, J. (1977). *The Clinical Prediction of Violent Behavior*. Lanham, Maryland: Jason Aronson, Inc.
- [5] Monahan, J., Steadman, H., et al. (2001). *Rethinking Risk Assessment: The MacArthur Study of Mental Disorder and Violence*. New York: Oxford University Press.
- [6] Monahan, J., et al. (2005). An actuarial model of violence risk assessment for persons with mental disorders. *Psychiatric Services*, 56, 810–815.
- [7] Wenk, E. A., Robison, J. O., & Smith, G. W. (1972). Can violence be predicted? *Crime and Delinquency*, 18, 393–402.

5 Appendix: Continuation of the American Psychiatric Association, *Amicus Curiae* Brief: Barefoot v. Estelle

INTRODUCTION AND SUMMARY OF ARGUMENT

The questions presented in this case are the logical outgrowth of two prior decisions by this Court. In the first, *Jurek v. Texas*, the Court dealt with the same Texas capital sentencing procedure involved here. The Court there rejected a constitutional challenge to the “future dangerousness” question, ruling that the statutory standard was not impermissibly vague. Although recognizing the difficulty inherent in predicting future behavior, the Court held that “[t]he task that [the] jury must perform . . . is basically no different from the task performed countless times each day throughout the American system of criminal justice.” The *Jurek* Court thus upheld the use of the Texas statutory question, but did not consider the types of evidence that could be presented to the jury for purposes of this determination.

Subsequently in *Estelle v. Smith*, the Court again dealt with the Texas sentencing scheme—this time in the context of a psychiatric examination to determine the defendant’s competency to stand trial. The Court held that the Fifth Amendment’s privilege against self-incrimination applied to such psychiatric examinations, at least to the extent that a prosecution psychiatrist later testifies concerning the defendant’s future dangerousness. The Court reasoned that although a defendant has no generalized constitutional right to remain silent at a psychiatric examination properly limited to the issues of sanity or competency, full *Miranda* warnings must be given with respect to testimony concerning future dangerousness because of “the gravity of the decision to be made at the penalty phase . . . ” The *Smith* decision thus enables a capital defendant to bar a government psychiatric examination on the issue of future dangerousness.

The [present] case raises the two issues left unresolved in *Jurek* and *Smith*. These are, first, whether a psychiatrist, testifying as an expert medical witness, may ever be permitted to render a prediction as to a capital defendant’s long-term future dangerousness. The second issue is whether such testimony may be elicited on the basis of hypothetical questions, even if there exists

no general prohibition against the use of expert psychiatric testimony on the issue of long-term future dangerousness. *Amicus* believes that both of these questions should be answered in the negative.

I. Psychiatrists should not be permitted to offer a prediction concerning the long-term future dangerousness of a defendant in a capital case, at least in those circumstances where the psychiatrist purports to be testifying as a medical expert possessing predictive expertise in this area. Although psychiatric assessments may permit short-term predictions of violent or assaultive behavior, medical knowledge has simply not advanced to the point where long-term predictions—the type of testimony at issue in this case—may be made with even reasonable accuracy. The large body of research in this area indicates that, even under the best of conditions, psychiatric predictions of long-term future dangerousness are wrong in at least two out of every three cases.

The forecast of future violent conduct on the part of a defendant in a capital case is, at bottom, a lay determination, not an expert psychiatric determination. To the extent such predictions have any validity, they can only be made on the basis of essentially actuarial data to which psychiatrists, qua psychiatrists, can bring no special interpretative skills. On the other hand, the use of psychiatric testimony on this issue causes serious prejudice to the defendant. By dressing up the actuarial data with an “expert” opinion, the psychiatrist’s testimony is likely to receive undue weight. In addition, it permits the jury to avoid the difficult actuarial questions by seeking refuge in a medical diagnosis that provides a false aura of certainty. For these reasons, psychiatric testimony on future dangerousness impermissibly distorts the fact-finding process in capital cases.

II. Even if psychiatrists under some circumstances are allowed to render an expert medical opinion on the question of future dangerousness, *amicus* submits that they should never be permitted to do so unless they have conducted a psychiatric examination of the defendant. It is evident from the testimony in this case that the key clinical determination relied upon by both psychiatrists was their diagnosis of “sociopathy” or “antisocial personality disorder.” However, such a diagnosis simply cannot be made on the basis of a hypothetical question. Absent an in-depth psychiatric examination and evaluation,

the psychiatrist cannot exclude alternative diagnoses; nor can he assure that the necessary criteria for making the diagnosis in question are met. As a result, he is unable to render a medical opinion with a reasonable degree of certainty.

These deficiencies strip the psychiatric testimony of all value in the present context. Even assuming that the diagnosis of antisocial personality disorder is probative of future dangerousness—an assumption which we do not accept—it is nonetheless clear that the limited facts given in the hypothetical fail to disprove other illnesses that plainly do not indicate a general propensity to commit criminal acts. Moreover, these other illnesses may be more amenable to treatment—a factor that may further reduce the likelihood of future aggressive behavior by the defendant.

...

6 Appendix: Opinion and Dissent in the U.S. Supreme Court, *Barefoot v. Estelle* (Decided, July 6, 1983)

Summary of the majority opinion:

(a) There is no merit to petitioner’s argument that psychiatrists, individually and as a group, are incompetent to predict with an acceptable degree of reliability that a particular criminal will commit other crimes in the future, and so represent a danger to the community. To accept such an argument would call into question predictions of future behavior that are constantly made in other contexts. Moreover, under the generally applicable rules of evidence covering the admission and weight of unprivileged evidence, psychiatric testimony predicting dangerousness may be countered not only as erroneous in a particular case but also as generally so unreliable that it should be ignored. Nor, despite the view of the American Psychiatric Association supporting petitioner’s view, is there any convincing evidence that such testimony is almost entirely unreliable, and that the factfinder and the adversary system will not be competent to uncover, recognize, and take due account of its shortcomings.

(b) Psychiatric testimony need not be based on personal examination of

the defendant, but may properly be given in response to hypothetical questions. Expert testimony, whether in the form of an opinion based on hypothetical questions or otherwise, is commonly admitted as evidence where it might help the factfinder do its job. Although this case involves the death penalty, there is no constitutional barrier to applying the ordinary rules of evidence governing the use of expert testimony.

...

Justice Blackmun dissenting:

I agree with most of what Justice Marshall has said in his dissenting opinion. I, too, dissent, but I base my conclusion also on evidentiary factors that the Court rejects with some emphasis. The Court holds that psychiatric testimony about a defendant's future dangerousness is admissible, despite the fact that such testimony is wrong two times out of three. The Court reaches this result—even in a capital case—because, it is said, the testimony is subject to cross-examination and impeachment. In the present state of psychiatric knowledge, this is too much for me. One may accept this in a routine lawsuit for money damages, but when a person's life is at stake—no matter how heinous his offense—a requirement of greater reliability should prevail. In a capital case, the specious testimony of a psychiatrist, colored in the eyes of an impressionable jury by the inevitable untouchability of a medical specialist's words, equates with death itself.

To obtain a death sentence in Texas, the State is required to prove beyond a reasonable doubt that “there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society.” As a practical matter, this prediction of future dangerousness was the only issue to be decided by Barefoot's sentencing jury.

At the sentencing hearing, the State established that Barefoot had two prior convictions for drug offenses and two prior convictions for unlawful possession of firearms. None of these convictions involved acts of violence. At the guilt stage of the trial, for the limited purpose of establishing that the crime was committed in order to evade police custody, the State had presented evidence that Barefoot had escaped from jail in New Mexico where he was being held on charges of statutory rape and unlawful restraint of a minor child with intent to commit sexual penetration against the child's

will. The prosecution also called several character witnesses at the sentencing hearing, from towns in five States. Without mentioning particular examples of Barefoot's conduct, these witnesses testified that Barefoot's reputation for being a peaceable and law-abiding citizen was bad in their respective communities.

Last, the prosecution called Doctors Holbrook and Grigson, whose testimony extended over more than half the hearing. Neither had examined Barefoot or requested the opportunity to examine him. In the presence of the jury, and over defense counsel's objection, each was qualified as an expert psychiatrist witness. Doctor Holbrook detailed at length his training and experience as a psychiatrist, which included a position as chief of psychiatric services at the Department of Corrections. He explained that he had previously performed many "criminal evaluations," and that he subsequently took the post at the Department of Corrections to observe the subjects of these evaluations so that he could "be certain those opinions that [he] had were accurate at the time of trial and pretrial." He then informed the jury that it was "within [his] capacity as a doctor of psychiatry to predict the future dangerousness of an individual within a reasonable medical certainty," and that he could give

"an expert medical opinion that would be within reasonable psychiatric certainty as to whether or not that individual would be dangerous to the degree that there would be a probability that that person would commit criminal acts of violence in the future that would constitute a continuing threat to society."

Doctor Grigson also detailed his training and medical experience, which, he said, included examination of "between thirty and forty thousand individuals," including 8,000 charged with felonies, and at least 300 charged with murder. He testified that, with enough information, he would be able to "give a medical opinion within reasonable psychiatric certainty as to the psychological or psychiatric makeup of an individual," and that this skill was "particular to the field of psychiatry, and not to the average layman."

Each psychiatrist then was given an extended hypothetical question asking him to assume as true about Barefoot the four prior convictions for nonviolent offenses, the bad reputation for being law-abiding in various communities, the

New Mexico escape, the events surrounding the murder for which he was on trial and, in Doctor Grigson's case, the New Mexico arrest. On the basis of the hypothetical question, Doctor Holbrook diagnosed Barefoot "within a reasonable psychiatr[ic] certainty," as a "criminal sociopath." He testified that he knew of no treatment that could change this condition, and that the condition would not change for the better but "may become accelerated" in the next few years. Finally, Doctor Holbrook testified that, "within reasonable psychiatric certainty," there was "a probability that the Thomas A. Barefoot in that hypothetical will commit criminal acts of violence in the future that would constitute a continuing threat to society," and that his opinion would not change if the "society" at issue was that within Texas prisons, rather than society outside prison.

Doctor Grigson then testified that, on the basis of the hypothetical question, he could diagnose Barefoot "within reasonable psychiatric certainty" as an individual with "a fairly classical, typical, sociopathic personality disorder." He placed Barefoot in the "most severe category of sociopaths (on a scale of one to ten, Barefoot was "above ten"), and stated that there was no known cure for the condition. Finally, Doctor Grigson testified that whether Barefoot was in society at large or in a prison society there was a "one hundred percent and absolute" chance that Barefoot would commit future acts of criminal violence that would constitute a continuing threat to society.

On cross-examination, defense counsel questioned the psychiatrists about studies demonstrating that psychiatrists' predictions of future dangerousness are inherently unreliable. Doctor Holbrook indicated his familiarity with many of these studies, but stated that he disagreed with their conclusions. Doctor Grigson stated that he was not familiar with most of these studies, and that their conclusions were accepted by only a "small minority group" of psychiatrists—"[i]t's not the American Psychiatric Association that believes that.

After an hour of deliberation, the jury answered "yes" to the two statutory questions, and Thomas Barefoot was sentenced to death.

The American Psychiatric Association (APA), participating in this case as *amicus curiae*, informs us that "[t]he unreliability of psychiatric predictions of long-term future dangerousness is by now an established fact within the

profession.” The APA’s best estimate is that two out of three predictions of long-term future violence made by psychiatrists are wrong. The Court does not dispute this proposition, and indeed it could not do so; the evidence is overwhelming. For example, the APA’s Draft Report of the Task Force on the Role of Psychiatry in the Sentencing Process (1983) states that

“[c]onsiderable evidence has been accumulated by now to demonstrate that long-term prediction by psychiatrists of future violence is an extremely inaccurate process.”

John Monahan, recognized as “the leading thinker on this issue” even by the State’s expert witness at Barefoot’s federal habeas corpus hearing, concludes that

“the ‘best’ clinical research currently in existence indicates that psychiatrists and psychologists are accurate in no more than one out of three predictions of violent behavior,”

even among populations of individuals who are mentally ill and have committed violence in the past. Another study has found it impossible to identify any subclass of offenders “whose members have a greater-than-even chance of engaging again in an assaultive act.” Yet another commentator observes:

“In general, mental health professionals . . . are more likely to be wrong than right when they predict legally relevant behavior. When predicting violence, dangerousness, and suicide, they are far more likely to be wrong than right.”

Neither the Court nor the State of Texas has cited a single reputable scientific source contradicting the unanimous conclusion of professionals in this field that psychiatric predictions of long-term future violence are wrong more often than they are right.

The APA also concludes, as do researchers that have studied the issue, that psychiatrists simply have no expertise in predicting long-term future dangerousness. A layman with access to relevant statistics can do at least as well, and possibly better; psychiatric training is not relevant to the factors that validly can be employed to make such predictions, and psychiatrists consistently err on the side of overpredicting violence. Thus, while Doctors Grigson and Holbrook were presented by the State and by self-proclamation as experts at predicting future dangerousness, the scientific literature makes

crystal clear that they had no expertise whatever. Despite their claims that they were able to predict Barefoot's future behavior "within reasonable psychiatric certainty," or to a "one hundred percent and absolute" certainty, there was, in fact, no more than a one in three chance that they were correct.³

It is impossible to square admission of this purportedly scientific but actually baseless testimony with the Constitution's paramount concern for reliability in capital sentencing.⁴ Death is a permissible punishment in Texas

³Like the District Court . . . and the Court of Appeals, . . . the Court seeks to justify the admission of psychiatric testimony on the ground that

"[t]he majority of psychiatric experts agree that where there is a pattern of repetitive assaultive and violent conduct, the accuracy of psychiatric predictions of future dangerousness dramatically rises."

. . . The District Court correctly found that there is empirical evidence supporting the common sense correlation between repetitive past violence and future violence; the APA states that

"[t]he most that can be said about any individual is that a history of past violence increases the probability that future violence will occur."

But psychiatrists have no special insights to add to this actuarial fact, and a single violent crime cannot provide a basis for a reliable prediction of future violence. . . .

The lower courts and this Court have sought solace in this statistical correlation without acknowledging its obvious irrelevance to the facts of this case. The District Court did not find that the State demonstrated any pattern of repetitive assault and violent conduct by Barefoot. Recognizing the importance of giving some credibility to its experts' specious prognostications, the State now claims that the "reputation" testimony adduced at the sentencing hearing "can only evince repeated, widespread acts of criminal violence." . . . This is simply absurd. There was no testimony worthy of credence that Barefoot had committed acts of violence apart from the crime for which he was being tried; there was testimony only of a bad reputation for peaceable and law-abiding conduct. In light of the fact that each of Barefoot's prior convictions was for a nonviolent offense, such testimony obviously could have been based on antisocial but nonviolent behavior. Neither psychiatrist informed the jury that he considered this reputation testimony to show a history of repeated acts of violence. Moreover, if the psychiatrists or the jury were to rely on such vague hearsay testimony in order to show a "pattern of repetitive assault and violent conduct," Barefoot's death sentence would rest on information that might "bear no closer relation to fact than the average rumor or item of gossip," . . . and should be invalid for that reason alone. A death sentence cannot rest on highly dubious predictions secretly based on a factual foundation of hearsay and pure conjecture. . . .

⁴Although I believe that the misleading nature of any psychiatric prediction of future violence violates due process when introduced in a capital sentencing hearing, admitting

only if the jury finds beyond a reasonable doubt that there is a probability the defendant will commit future acts of criminal violence. The admission of unreliable psychiatric predictions of future violence, offered with unabashed claims of “reasonable medical certainty” or “absolute” professional reliability, creates an intolerable danger that death sentences will be imposed erroneously.

The plurality in *Woodson v. North Carolina*, stated:

“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 in a specific case.” The Court does not see fit to mention this principle today, yet it is as firmly established as any in our Eighth Amendment jurisprudence. Only two weeks ago, in *Zant v. Stephens*, the Court described the need for reliability in the application of the death penalty as one of the basic “themes . . . reiterated in our opinions discussing the procedures required by the Constitution in capital sentencing determinations.” (capital punishment must be “imposed fairly, and with reasonable consistency, or not at all”). State evidence rules notwithstanding, it is well established that, because the truth-seeking process may be unfairly skewed, due process may be violated even in a noncapital criminal case by the exclusion of evidence probative of innocence, or by the admission of certain categories of unreliable and prejudicial evidence (“[i]t is the reliability of identification evidence that primarily determines its admissi-

the predictions in this case—which were made without even examining the defendant—was particularly indefensible. In the APA’s words, if prediction following even an in-depth examination is inherently unreliable,

“there is all the more reason to shun the practice of testifying without having examined the defendant at all. . . . Needless to say, responding to hypotheticals is just as fraught with the possibility of error as testifying in any other way about an individual whom one has not personally examined. Although the courts have not yet rejected the practice, psychiatrists should.”

. . . Such testimony is offensive not only to legal standards; the APA has declared that “[i]t is unethical for a psychiatrist to offer a professional opinion unless he/she has conducted an examination.” . . . The Court today sanctions admission in a capital sentencing hearing of “expert” medical testimony so unreliable and unprofessional that it violates the canons of medical ethics.

bility”). The reliability and admissibility of evidence considered by a capital sentencing factfinder is obviously of still greater constitutional concern.

The danger of an unreliable death sentence created by this testimony cannot be brushed aside on the ground that the “jury [must] have before it all possible relevant information about the individual defendant whose fate it must determine.” Although committed to allowing a “wide scope of evidence” at presentence hearings, the Court has recognized that “consideration must be given to the quality, as well as the quantity, of the information on which the sentencing [authority] may rely.” Thus, very recently, this Court reaffirmed a crucial limitation on the permissible scope of evidence: “[s]o long as the evidence introduced . . . do[es] not prejudice a defendant, it is preferable not to impose restrictions.” The Court all but admits the obviously prejudicial impact of the testimony of Doctors Grigson and Holbrook; granting that their absolute claims were more likely to be wrong than right, the Court states that “[t]here is no doubt that the psychiatric testimony increased the likelihood that petitioner would be sentenced to death.” Indeed, unreliable scientific evidence is widely acknowledged to be prejudicial. The reasons for this are manifest. “The major danger of scientific evidence is its potential to mislead the jury; an aura of scientific infallibility may shroud the evidence, and thus lead the jury to accept it without critical scrutiny.”⁵

⁵There can be no dispute about this obvious proposition:

“Scientific evidence impresses lay jurors. They tend to assume it is more accurate and objective than lay testimony. A juror who thinks of scientific evidence visualizes instruments capable of amazingly precise measurement, of findings arrived at by dispassionate scientific tests. In short, in the mind of the typical lay juror, a scientific witness has a special aura of credibility.”

... “Scientific . . . evidence has great potential for misleading the jury. The low probative worth can often be concealed in the jargon of some expert . . . ” This danger created by use of scientific evidence frequently has been recognized by the courts. Speaking specifically of psychiatric predictions of future dangerousness similar to those at issue, one District Court has observed that, when such a prediction

“is proffered by a witness bearing the title of ‘Doctor,’ its impact on the jury is much greater than if it were not masquerading as something it is not.”

... In *United States v. Addison*, the court observed that scientific evidence may “assume a posture of mystic infallibility in the eyes of a jury of laymen.” Another court has noted that scientific evidence “is likely to be shrouded with an aura of near infallibility, akin to the ancient oracle of Delphi.” ...

Where the public holds an exaggerated opinion of the accuracy of scientific testimony, the prejudice is likely to be indelible. There is little question that psychiatrists are perceived by the public as having a special expertise to predict dangerousness, a perception based on psychiatrists' study of mental disease. It is this perception that the State in Barefoot's case sought to exploit. Yet mental disease is not correlated with violence, and the stark fact is that no such expertise exists. Moreover, psychiatrists, it is said, sometimes attempt to perpetuate this illusion of expertise, and Doctors Grigson and Holbrook—who purported to be able to predict future dangerousness “within reasonable psychiatric certainty,” or absolutely—present extremely disturbing examples of this tendency. The problem is not uncommon.

Furthermore, as is only reasonable, the Court's concern in encouraging the introduction of a wide scope of evidence has been to ensure that accurate information is provided to the sentencing authority without restriction. The joint opinion announcing the judgment in Gregg explained the jury's need for relevant evidence in these terms:

“If an experienced trial judge, who daily faces the difficult task of imposing sentences, has a vital need for accurate information . . . to be able to impose a rational sentence in the typical criminal case, then accurate sentencing information is an indispensable prerequisite to a reasoned determination of whether a defendant shall live or die by a jury of people who may never before have made a sentencing decision.”

So far as I am aware, the Court never has suggested that there is any interest in providing deceptive and inaccurate testimony to the jury. Psychiatric predictions of future dangerousness are not accurate; wrong two times out of three, their probative value, and therefore any possible contribution they might make to the ascertainment of truth, is virtually nonexistent (psychiatric testimony not sufficiently reliable to support finding that individual will be dangerous under any standard of proof). Indeed, given a psychiatrist's prediction that an individual will be dangerous, it is more likely than not that the defendant will not commit further violence. It is difficult to understand how the admission of such predictions can be justified as advancing the search for truth, particularly in light of their clearly prejudicial effect. Thus, the Court's remarkable observation that “[n]either petitioner nor the [APA]

suggests that psychiatrists are always wrong with respect to future dangerousness, only most of the time,” misses the point completely, and its claim that this testimony was no more problematic than “other relevant evidence against any defendant in a criminal case,” is simply incredible. Surely, this Court’s commitment to ensuring that death sentences are imposed reliably and reasonably requires that nonprobative and highly prejudicial testimony on the ultimate question of life or death be excluded from a capital sentencing hearing.

Despite its recognition that the testimony at issue was probably wrong and certainly prejudicial, the Court holds this testimony admissible because the Court is

“unconvinced . . . that the adversary process cannot be trusted to sort out the reliable from the unreliable evidence and opinion about future dangerousness.”

One can only wonder how juries are to separate valid from invalid expert opinions when the “experts” themselves are so obviously unable to do so. Indeed, the evidence suggests that juries are not effective at assessing the validity of scientific evidence.

There can be no question that psychiatric predictions of future violence will have an undue effect on the ultimate verdict. Even judges tend to accept psychiatrists’ recommendations about a defendant’s dangerousness with little regard for cross-examination or other testimony. The American Bar Association has warned repeatedly that sentencing juries are particularly incapable of dealing with information relating to “the likelihood that the defendant will commit other crimes,” and similar predictive judgments. Relying on the ABA’s conclusion, the joint opinion announcing the judgment in *Gregg v. Georgia*, recognized that,

“[s]ince the members of a jury will have had little, if any, previous experience in sentencing, they are unlikely to be skilled in dealing with the information they are given.”

But the Court in this case, in its haste to praise the jury’s ability to find the truth, apparently forgets this well-known and worrisome shortcoming.

As if to suggest that petitioner’s position that unreliable expert testimony should be excluded is unheard of in the law, the Court relies on the proposi-

tion that the rules of evidence generally

“anticipate that relevant, unprivileged evidence should be admitted and its weight left to the factfinder, who would have the benefit of cross-examination and contrary evidence by the opposing party.”

But the Court simply ignores hornbook law that, despite the availability of cross-examination and rebuttal witnesses,

“opinion evidence is not admissible if the court believes that the state of the pertinent art or scientific knowledge does not permit a reasonable opinion to be asserted.”

Because it is feared that the jury will overestimate its probative value, polygraph evidence, for example, almost invariably is excluded from trials despite the fact that, at a conservative estimate, an experienced polygraph examiner can detect truth or deception correctly about 80 to 90 percent of the time. In no area is purportedly “expert” testimony admitted for the jury’s consideration where it cannot be demonstrated that it is correct more often than not. “It is inconceivable that a judgment could be considered an expert’ judgment when it is less accurate than the flip of a coin.” The risk that a jury will be incapable of separating “scientific” myth from reality is deemed unacceptably high.⁶

The Constitution’s mandate of reliability, with the stakes at life or death, precludes reliance on cross-examination and the opportunity to present rebuttal witnesses as an antidote for this distortion of the truthfinding process. Cross-examination is unlikely to reveal the fatuousness of psychiatric predictions because such predictions often rest, as was the case here, on psychiatric categories and intuitive clinical judgments not susceptible to cross-examination and rebuttal. Psychiatric categories have little or no demon-

⁶The Court observes that this well-established rule is a matter of evidence law, not constitutional law. . . . But the principle requiring that capital sentencing procedures ensure reliable verdicts, which the Court ignores, and the principle that due process is violated by the introduction of certain types of seemingly conclusive, but actually unreliable, evidence, . . . which the Court also ignores, are constitutional doctrines of long standing. The teaching of the evidence doctrine is that unreliable scientific testimony creates a serious and unjustifiable risk of an erroneous verdict, and that the adversary process, at its best, does not remove this risk. We should not dismiss this lesson merely by labeling the doctrine nonconstitutional; its relevance to the constitutional question before the Court could not be more certain.

strated relationship to violence, and their use often obscures the unimpressive statistical or intuitive bases for prediction.⁷ The APA particularly condemns the use of the diagnosis employed by Doctors Grigson and Holbrook in this case, that of sociopathy:

“In this area confusion reigns. The psychiatrist who is not careful can mislead the judge or jury into believing that a person has a major mental disease simply on the basis of a description of prior criminal behavior. Or a psychiatrist can mislead the court into believing that an individual is devoid of conscience on the basis of a description of criminal acts alone. . . . The profession of psychiatry has a responsibility to avoid inflicting this confusion upon the courts, and to spare the defendant the harm that may result. . . . Given our uncertainty about the implications of the finding, the diagnosis of sociopathy . . . should not be used to justify or to support predictions of future conduct. There is no certainty in this area.”

It is extremely unlikely that the adversary process will cut through the facade of superior knowledge. The Chief Justice [Burger] long ago observed:

“The very nature of the adversary system . . . complicates the use of scientific opinion evidence, particularly in the field of psychiatry. This system of partisan contention, of attack and counterattack, at its best is not ideally suited to developing an accurate portrait or profile of the human personality, especially in the area of abnormal behavior. Although under ideal conditions the adversary system can develop for a jury most of the necessary fact material for an adequate decision, such conditions are rarely achieved in the courtrooms in this country. These ideal conditions would include a highly skilled and experienced trial judge and highly skilled lawyers on both sides of the case, all of whom, in addition to being well-trained in the law and in the techniques of advocacy, would be sophisticated in matters of medicine, psychiatry, and psychology. It is far too rare that all three of the legal actors in the cast meet these standards.”

⁷In one study, for example, the only factor statistically related to whether psychiatrists predicted that a subject would be violent in the future was the type of crime with which the subject was charged. Yet the defendant’s charge was mentioned by the psychiatrists to justify their predictions in only one-third of the cases. The criterion most frequently cited was “delusional or impaired thinking.” . . .

Another commentator has noted:

“Competent cross-examination and jury instructions may be partial antidotes . . . but they cannot be complete. Many of the cases are not truly adversarial; too few attorneys are skilled at cross-examining psychiatrists, laypersons overweigh the testimony of experts, and, in any case, unrestricted use of experts promotes the incorrect view that the questions are primarily scientific. There is, however, no antidote for the major difficulty with mental health ‘experts’—that they simply are not experts. . . . In realms beyond their true expertise, the law has little special to learn from them; too often, their testimony is . . . prejudicial.”

Nor is the presentation of psychiatric witnesses on behalf of the defense likely to remove the prejudicial taint of misleading testimony by prosecution psychiatrists. No reputable expert would be able to predict with confidence that the defendant will not be violent; at best, the witness will be able to give his opinion that all predictions of dangerousness are unreliable. Consequently, the jury will not be presented with the traditional battle of experts with opposing views on the ultimate question. Given a choice between an expert who says that he can predict with certainty that the defendant, whether confined in prison or free in society, will kill again, and an expert who says merely that no such prediction can be made, members of the jury, charged by law with making the prediction, surely will be tempted to opt for the expert who claims he can help them in performing their duty, and who predicts dire consequences if the defendant is not put to death.⁸

Moreover, even at best, the presentation of defense psychiatrists will convert the death sentence hearing into a battle of experts, with the Eighth

⁸“Although jurors may treat mitigating psychiatric evidence with skepticism, they may credit psychiatric evidence demonstrating aggravation. Especially when jurors’ sensibilities are offended by a crime, they may seize upon evidence of dangerousness to justify an enhanced sentence.” . . . Thus, the danger of jury deference to expert opinions is particularly acute in death penalty cases. Expert testimony of this sort may permit juries to avoid the difficult and emotionally draining personal decisions concerning rational and just punishment. . . . Doctor Grigson himself has noted both the superfluousness and the misleading effect of his testimony: “I think you could do away with the psychiatrist in these cases. Just take any man off the street, show him what the guy’s done, and most of these things are so clear-cut he would say the same things I do. But I think the jurors feel a little better when a psychiatrist says it—somebody that’s supposed to know more than they know.” . . .

Amendment's well-established requirement of individually focused sentencing a certain loser. The jury's attention inevitably will turn from an assessment of the propriety of sentencing to death the defendant before it to resolving a scientific dispute about the capabilities of psychiatrists to predict future violence. In such an atmosphere, there is every reason to believe that the jury may be distracted from its constitutional responsibility to consider "particularized mitigating factors," in passing on the defendant's future dangerousness.

One searches the Court's opinion in vain for a plausible justification for tolerating the State's creation of this risk of an erroneous death verdict. As one Court of Appeals has observed:

"A courtroom is not a research laboratory. The fate of a defendant . . . should not hang on his ability to successfully rebut scientific evidence which bears an 'aura of special reliability and trustworthiness,' although, in reality, the witness is testifying on the basis of an unproved hypothesis . . . which has yet to gain general acceptance in its field."

Ultimately, when the Court knows full well that psychiatrists' predictions of dangerousness are specious, there can be no excuse for imposing on the defendant, on pain of his life, the heavy burden of convincing a jury of laymen of the fraud.⁹

The Court is simply wrong in claiming that psychiatric testimony respect-

⁹The Court is far wide of the mark in asserting that excluding psychiatric predictions of future dangerousness from capital sentencing proceedings "would immediately call into question those other contexts in which predictions of future behavior are constantly made." . . . Short-term predictions of future violence, for the purpose of emergency commitment or treatment, are considerably more accurate than long-term predictions. In other contexts where psychiatric predictions of future dangerousness are made, moreover, the subject will not be criminally convicted, much less put to death, as a result of predictive error. The risk of error therefore may be shifted to the defendant to some extent. . . . The APA, discussing civil commitment proceedings based on determinations of dangerousness, states that, in light of the unreliability of psychiatric predictions, "[c]lose monitoring, frequent follow-up, and a willingness to change one's mind about treatment recommendations and dispositions for violent persons, whether within the legal system or without, is the only acceptable practice if the psychiatrist is to play a helpful role in these assessments of dangerousness." . . . In a capital case, there will be no chance for "follow-up" or "monitoring." A subsequent change of mind brings not justice delayed, but the despair of irreversible error. . . .

ing future dangerousness is necessarily admissible in light of *Jurek v. Texas*, or *Estelle v. Smith*. As the Court recognizes, *Jurek* involved “only lay testimony.” Thus, it is not surprising that “there was no suggestion by the Court that the testimony of doctors would be inadmissible,” and it is simply irrelevant that the *Jurek* Court did not “disapprov[e]” the use of such testimony. In *Smith*, the psychiatric testimony at issue was given by the same Doctor Grigson who confronts us in this case, and his conclusions were disturbingly similar to those he rendered here. The APA, appearing as *amicus curiae*, argued that all psychiatric predictions of future dangerousness should be excluded from capital sentencing proceedings. The Court did not reach this issue, because it found *Smith*’s death sentence invalid on narrower grounds: Doctor Grigson’s testimony had violated *Smith*’s Fifth and Sixth Amendment right. Contrary to the Court’s inexplicable assertion in this case, *Smith* certainly did not reject the APA’s position. Rather, the Court made clear that “the holding in *Jurek* was guided by recognition that the inquiry [into dangerousness] mandated by Texas law does not require resort to medical experts.” If *Jurek* and *Smith* held that psychiatric predictions of future dangerousness are admissible in a capital sentencing proceeding as the Court claims, this guiding recognition would have been irrelevant.

The Court also errs in suggesting that the exclusion of psychiatrists’ predictions of future dangerousness would be contrary to the logic of *Jurek*. *Jurek* merely upheld Texas’ substantive decision to condition the death sentence upon proof of a probability that the defendant will commit criminal acts of violence in the future. Whether the evidence offered by the prosecution to prove that probability is so unreliable as to violate a capital defendant’s rights to due process is an entirely different matter, one raising only questions of fair procedure.¹⁰ *Jurek*’s conclusion that Texas may impose the death penalty on capital defendants who probably will commit criminal acts of violence in no way establishes that the prosecution may convince a jury that this is so by

¹⁰The Court’s focus in the death penalty cases has been primarily on ensuring a fair procedure: “In ensuring that the death penalty is not meted out arbitrarily or capriciously, the Court’s principal concern has been more with the procedure by which the State imposes the death sentence than with the substantive factors the State lays before the jury as a basis for imposing death, once it has been determined that the defendant falls within the category of persons eligible for the death penalty.”

misleading or patently unreliable evidence.

Moreover, Jurek's holding that the Texas death statute is not impermissibly vague does not lead ineluctably to the conclusion that psychiatric testimony is admissible. It makes sense to exclude psychiatric predictions of future violence while admitting lay testimony, because psychiatric predictions appear to come from trained mental health professionals, who purport to have special expertise. In view of the total scientific groundlessness of these predictions, psychiatric testimony is fatally misleading. Lay testimony, frankly based on statistical factors with demonstrated correlations to violent behavior, would not raise this substantial threat of unreliable and capricious sentencing decisions, inimical to the constitutional standards established in our cases; and such predictions are as accurate as any a psychiatrist could make. Indeed, the very basis of Jurek, as I understood it, was that such judgments can be made by laymen on the basis of lay testimony.

Our constitutional duty is to ensure that the State proves future dangerousness, if at all, in a reliable manner, one that ensures that "any decision to impose the death sentence be, and appear to be, based on reason rather than caprice or emotion." Texas' choice of substantive factors does not justify loading the factfinding process against the defendant through the presentation of what is, at bottom, false testimony.