Behavioral Sciences and the Law
Behav. Sci. Law (2013)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/bsl.2064

Probability of Criminal Acts of Violence: A Test of Jury Predictive Accuracy

Thomas J. Reidy*,†, Jon R. Sorensen‡ and Mark D. Cunningham§

The ability of capital juries to accurately predict future prison violence at the sentencing phase of aggravated murder trials was examined through retrospective review of the disciplinary records of 115 male inmates sentenced to either life ($n = 65$) or death ($n = 50$) in Oregon from 1985 through 2008, with a mean post-conviction time at risk of 15.3 years. Violent prison behavior was completely unrelated to predictions made by capital jurors, with bidirectional accuracy simply reflecting the base rate of assaultive misconduct in the group. Rejection of the special issue predicting future violence enjoyed 90% accuracy. Conversely, predictions that future violence was probable had 90% error rates. More than 90% of the assaultive rule violations committed by these offenders resulted in no harm or only minor injuries.

The U.S. Supreme Court in *Jurek v. Texas* (1976) expressed confidence that capital jurors could reliably identify capital defendants at risk for committing future violence. This confidence has subsequently been widely adopted in capital sentencing statutes and case law as an available metric for determining the deathworthiness of a capital offender. At odds with this confidence in jury predictive capabilities, however, retrospective reviews of the post-conviction prison behavior of capital offenders have consistently demonstrated alarming error rates in jury accuracy in predictions of future violence (Cunningham, Sorensen, & Reidy, 2009; Cunningham, Sorensen, Vigen, & Woods, 2010; Cunningham, Sorensen, Vigen, & Woods, 2011; Edens, Buffington-Vollum, Keilin, Roskamp, & Anthony, 2005; Marquart, Ekland-Olson, & Sorensen, 1989; Marquart, Ekland-Olson, & Sorensen, 1994). Despite these data, considerations of future violence persist in death penalty schemes throughout the United States. This suggests that public policy makers are unaware of these findings or view them as insufficiently replicated to abandon future violence predictions as a fixture in modern American death penalty jurisprudence.

THE ROLE OF FUTURE VIOLENCE AS A DEATH-SENTENCING CONSIDERATION

Statutes

The potential that a capital offender will continue to perpetrate serious violence has varied applications in modern death penalty jurisprudence (see Krauss, McCabe, &

*Correspondence to: Thomas J. Reidy, 154 Central Avenue, Salinas, CA 93901, U.S.A. E-mail: tomreidy@comcast.net
†Private Practice, 154 Central Avenue, Salinas, CA 93901, U.S.A
‡Department of Criminal Justice, East Carolina University, NC
§Private Practice, DallasTX

McFadden, 2009). In two jurisdictions, Oregon and Texas, a jury, in order to sentence an offender to death, must give an affirmative answer to a "special issue:" "whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society" (for a discussion of the origin of this issue and related assumptions, see Cunningham, 2006). In four states (Idaho, Oklahoma, Virginia, and Wyoming), an assertion that the offender will commit future violence is available as a statutory aggravating factor.[1] Prosecutors in many other death penalty states and in the federal system regularly introduce the probability of future violence as a nonstatutory aggravating factor for jurors to consider (Shapiro, 2009). The alternative, that an offender is likely to have a positive or non-violent adjustment to prison, is available as a mitigating factor in all jurisdictions (see *Skipper v. South Carolina*, 1986).

Pragmatic and Case Law Contexts of the Predicted Violence

In all American death penalty jurisdictions, life without parole (LWOP) is a sentencing alternative to the death penalty. In some jurisdictions, there is a third option of life with parole (LWP). As a practical matter, when death is being considered as a preventative intervention, prison is the only context where future violence could occur/originate. Consistent with this pragmatic reality that an LWOP inmate will never again be in the community, federal district court decisions have repeatedly limited the Government to evidence that is relevant to a context of life in prison (see *United States v. Cooper*, 2000; *United States v. Gilbert*, 2000; *United States v. Glover*, 1999; *United States v. Llera Plaza*, 2001; *United States v. Peoples*, 1999; *United States v. Rodriguez*, 2006; *United States v. Sablan*, 2007). In rare instances (e.g., terrorist, organized crime figure), capital offenders may have external resources available to them that make ordered-violence in the community *from* prison a risk assessment consideration.

State appellate courts in some states (e.g., Oregon, Texas, Virginia), however, citing the silence of *Jurek* in defining "society," have declined to limit a jury's consideration of the future violence of a capital offender to a prison context. For example, the Oregon Supreme Court held in *State v. Douglas* (1990) that when the jury considers the threat that the defendant might pose because of future violent crimes, the jury could consider both threat to society-at-large and threat to prison society. Similarly, the Texas Court of Criminal Appeals has asserted that the special issue should be construed as asking, "whether a defendant would constitute a continuing threat 'whether in or out of prison' without regard to how long the defendant would actually spend in prison if sentenced to life" (*Estrada v. State*, 2010; see also *Coble v. State*, 2010). The rationale expressed in the Texas Court of Criminal Appeals for a context-free determination is similar to that articulated by the Virginia Supreme Court in 2001 in *Burns v. Commonwealth* (see also *Porter v. Commonwealth*, 2008; *Morva v. Commonwealth*, 2009): "[T]he relevant inquiry is not whether Burns *could* commit criminal acts of violence in the future but whether he *would*" (i.e., if unfettered by context).

From a scientific perspective, there are fundamental problems with a context-free assessment of "theoretical" violent acts that "would" occur if the stars aligned. First,

[1] Aggravating factors fall into two categories: statutory and nonstatutory. The jury must unanimously find at least one aggravating factor defined by statute for the defendant to be death-*eligible*. Nonstatutory aggravating factors, in those jurisdictions that allow them, inform the ultimate sentencing decision, but do not establish eligibility for the death penalty.

DOI: 10.1002/bsl

a maxim of violence risk assessment is that violence is always a function of context (see Monahan, 1981). As Hall (1987) succinctly described: "Individual persons are never dangerous in toto" (p. 10). Second, the character- and offense-driven risk assessment methodology prescribed by both the Texas and Virginia high courts in reaching a context-free risk determination asserts factors that are not predictive of serious violence in prison (see Cunningham, 2006, 2010; Cunningham et al., 2011; Edens et al., 2005). Third, a context-free assessment of theoretical acts that "would" occur fails to narrow the class of deathworthy defendants. Every capital offender presents a disproportionate risk under such criteria, when compared with the "average" citizen. Fourth, such a context-free and theoretical assessment is fundamentally arbitrary, as no systematic or scientific metric can illuminate this consideration or its error rate. Such arbitrariness appears antithetical to the Supreme Court's prescription (*Kansas v. March*, 2006; as cited by *Coble v. State*, 2010, at footnote 23) that capital statutes *rationally* narrow the class of death-eligible defendants; and permit a jury to render a *reasoned*, individualized sentencing determination.

Severity of Violence

The *Jurek* court did not specify the severity or persistence of violence that constitutes "a continuing threat to society." Some jurisdictions have adopted more restrictive language, such as "...propensity to commit murder which will probably constitute a continuing threat to society" (Idaho Code §19-2515). Similarly, the Governor's Commission on Capital Punishment (Illinois, 2002) specified that justifications that would minimize arbitrary application of the death penalty included: "Incapacitating persons with a clearly demonstrated propensity to murder again" (p. 69). A number of scholars (see Clausen-Schulz, Pearce, & Schopp, 2004; Cunningham & Reidy, 1998; Cunningham et al., 2009; Edens et al., 2005; Slobogin, 2009) have asserted that a preventative intervention of death should bear some reasonable proportionality to the act(s) that are presumed to occur in its absence. Such a proportionality analysis would arguably require acts that would result in major injuries or death.

To the extent that more severe types of violence are contemplated, as would be suggested by proportionality considerations, the problem of unreported violence as a source of error is much reduced. The most severe forms of prison violence are quite likely to be both reported and properly attributed. Studies have shown that all inmate homicides are almost always attributed to the inmate offender(s), as are serious acts of violence against other inmates and staff members (see Cunningham et al., 2010; Cunningham et al., 2011; Sorensen, Cunningham, Vigen, & Woods, 2011).

Juror Concern with Future Violence

Even when future violence is not expressly introduced in capital trials, Capital Jury Project studies funded by the National Science Foundation have found that "future danger" is always a consideration on the minds of capital jurors and a compelling motivation for invoking the death penalty (Blume, Garvey, & Johnson, 2001; Costanzo & Costanzo, 1994; Sandys, Pruss, & Walsh, 2009). Post-trial interviews of actual capital jurors revealed that fear of future homicide, and not just serious violence, is an influential factor in a death penalty verdict. Interestingly, while an anticipation of future violence serves as a strong aggravating force driving juror decision-making

DOI: 10.1002/bsl

toward a preference for death, the lack of future violence only has a modest mitigating effect on capital juror determinations (Garvey, 1998; Sandys et al., 2009).

The use of scientific data in defining future violence risk is fundamental to promoting reasoned decisions, rather than allowing irrational and unfounded fears of capital jurors to influence decision-making (Cunningham et al., 2009). Unfortunately, jurors often fail to grasp the import of statistical/actuarial data, finding intuitive or "clinical" approaches more persuasive, however erroneous the asserted correlations (Krauss & Sales, 2001; Sandys et al., 2009). Juror and potentially judicial susceptibility to intuitive risk methodologies and illusory correlations (see also Dawes, Faust, & Meehl, 1989; Shah, 1978) does not bode well for predictive accuracy or effective evidentiary gate-keeping regarding this issue (see *Daubert v. Merrill Dow Pharmaceuticals*, 1993).

PREDICTIVE ACCURACY OF CAPITAL JURIES

To the extent that capital juries assessing the violence risk of capital offenders were contemplating the prison context, the accuracy of their predictions is illuminated by inspection of the post-conviction prison files of these capital inmates. The studies providing these retrospective reviews of inmate files are of two types (see Table 1). In the unidirectional studies, the capital juries endorsed future violence as a special issue and the subsequent prison behavior of these violence-predicted offenders was examined. These studies demonstrate an extraordinarily high error rate for jury predictions of future violence, assuming predictions contemplating prison assaults resulting in more than first aid injury.

Bidirectional studies of jury predictive performance have performed retrospective reviews of the prison records of convicted capital offenders following trials where juries asserted or rejected the probability of future violence. In two bidirectional studies of jury performance to date (Cunningham et al., 2009; Marquart et al., 1989), the predictions of these jurors had high accuracy rates when rejecting a prediction of future violence, and high error rates in predictions of violence. However, no discriminating abilities were demonstrated. The predictions of these juries were no better than random guesses, i.e., they showed no improvement over the very low base rates of serious violence among the samples.

In both unidirectional and bidirectional studies, as the severity of criterion violence increases, the accuracy rates become fully polarized: virtually 100% error rates for predictions of life-threatening violence and 100% accuracy rates for rejecting such predictions. The findings from these retrospective reviews provide consistent evidence that capital juries are not able to reliably (in the legal sense, not the statistical sense) identify capital offenders who will go on to commit the rare act of serious violence.

CURRENT STUDY

It could be argued that the studies to date have been limited to two jurisdictions (i.e., Texas and the Federal Bureau of Prisons), thus limiting whether these data generalize more widely. Further, only two studies (Cunningham et al., 2009; Cunningham et al., 2010) provide data on capital jury predictive performance for inmates confined in the modern (post-1990) corrections era. Although there is no apparent rationale as to why

DOI: 10.1002/bsl

Table 1. Error rates in capital jury predictions of future violence by severity of violence using prison as the criterion context

Study	Type	Jurisdiction	Sample N/type	At risk	Error rate of jury prediction (%ER)	
					Endorse PCAV[a]	Reject PCAV[a]
Marquart et al. (1989)	BiD	Texas	90 FDR	$M = 6.3$ years[b]	Serious assault[c] – 94.5% ER Homicide – 98.9% ER	
			107 LS	$M = 7.2$ years		Serious assault[c] – 7.5% ER
Marquart et al. (1994)	UniD	Texas	421 DR	1974–1988 M = n.a.	Assault – 89.3% ER Homicide – 99.5% ER	Homicide – 0% ER
Edens et al. (2005)	UniD	Texas	65 EDR	$M = 12$ years	Serious assault[d] – 95.4% ER Homicide – 100% ER	
			42 CDR	$M = 8$ years	Serious assault[d] – 92.9% ER Homicide – 100% ER	
			48 FDR	$M = 22$ years	Serious assault[d] – 95.8% ER Homicide – 100% ER	
Cunningham et al., 2009	BiD	Federal	72 Cap $n = 34$	$M = 5.7$ years $M = 6.1$ years	Serious assault[e] – 94.1% ER Mod. injury[f] – 100% ER Major injury[g] – 100% ER Fatal injury – 97.1% ER	
			$n = 38$	$M = 5.4$ years		Serious assault[e] – 5.3% ER Mod. Injury[f] – 0% ER Major injury[g] – 0% ER Fatal injury – 0% ER
			33 DR $n = 27$		Serious assault[e] – 100% ER	
			$n = 6$			Serious assault[e] – 0% ER
			39 LS $n = 7$		Serious assault[e] – 71.4% ER	
			$n = 32$			Serious assault[e] – 6.2% ER
Cunningham et al. (2011)	UniD	Texas	111 FDR	$M = 18.3$ years 9.9 years DR 8.4 years BP[b]	Serious assault[d] – 91.9% ER Major injury[g] – 100% ER Homicide – 100% ER	

Note. BiD, bidirectional; BP, broader prison population; Cap, capital offenders sentenced by juries; CDR, offenders who continued to be on death row; DR, death row; EDR, offenders who had been executed; ER, error rate of jury prediction; FDR, former death row, i.e., offenders initially sentenced to death who obtained relief from this sentence and were then followed in the broader prison population; LS, life-sentenced by jury; PCAV, whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society; UniD, unidirectional.

[a]As a "special issue" for Texas juries, or a non-statutory aggravating factor for federal juries.

[b]Time at risk in the broader prison population following relief from the death sentence.

[c]Assault on inmate with a weapon, sex by force, assault on staff.

[d]Resulting in injury requiring more than first-aid treatment.

[e]Includes BOP codes 100, 100A, 101, 101A, 107.

[f]Resulting in hospital admission for serious but not life-threatening injuries.

[g]Resulting in hospital admission for life-threatening injuries.

DOI: 10.1002/bsl

Texas or federal capital juries would be particularly impaired in their predictive abilities, further research with modern era inmates in other jurisdictions is indicated to further test capital jury predictive performance.

Adopting the Texas capital sentencing scheme, the Oregon death penalty statute requires jurors, in order to sentence an offender to death, to unanimously and affirmatively answer a special issue: "Whether there is a probability that defendant would commit criminal acts of violence that would constitute a continuing threat to society." Alternatively, Oregon juries may sentence a convicted capital offender to either a true life option (life without the possibility of parole) or a 30-year minimum term before parole consideration. The availability of Oregon jury verdicts on this special issue and longitudinal inmate data from the Oregon Department of Corrections (ODOC) present an opportunity to directly study the accuracy of predictions of future violent conduct, under the sentencing scheme affirmed in *Jurek*, over an extended period of confinement for a group of aggravated murderers (i.e., death-eligible offenders).

Results from the current longitudinal study will permit a bidirectional comparison of jury decisions under the Oregon "special issue," taking into account acts of prison misconduct and violence disaggregated by type and severity, prior criminal history, and characteristics of the aggravated murder defendant and offense. The lengthy "at risk" prison tenures of many of the offender participants in this study, permitting scrutiny of over two decades of inmate misconduct, provides an extended opportunity for violence, minimizing the impact of violence that has not yet occurred on jury accuracy rates.

Research on prison violence has been criticized (see Cunningham & Reidy, 1998; Cunningham et al., 2009; Edens et al., 2005) for failing to sufficiently operationalize specific types of institutional misconduct by aggregating into a single all-inclusive measure of aggression or violence behaviors as diverse as self-mutilation, belligerence, property damage, threats, obscene gestures, throwing urine/feces, fights, and different levels of assault. Meta-analyses have shown that the type of outcome measure influences the efficacy of prediction instruments (Guy, Edens, Anthony, & Douglas, 2005; Yang, Wong, & Coid, 2010). Consistent with recommended procedures, this study will disaggregate measures of institutional violence by type, level of severity, and nature of injuries to provide greater clarity and specificity for examining jury determinations of future violence.

Institutional review board (IRB) approval for this study was obtained through Prairie View A&M University. Because these data are based on archival data and reports only group findings, informed consent was not sought from Oregon inmates in this study.

METHODS

Participants

The pool of participants was drawn from a database supplied by the Oregon Capital Resource Center (OCRC). As part of their "Aggravated Murder Sentencing Phase Verdicts Project," the OCRC collected information on Oregon aggravated murder cases[2] that had proceeded to the sentencing phase of trial since 1985. In identifying the pool of eligible cases, the OCRC relied on a list of aggravated murder cases provided by the tracking

[2] The only crime subject to the death penalty in Oregon is aggravated murder. As used in ORS 163.105 aggravated murder refers to murder as defined in ORS 163.115.

DOI: 10.1002/bsl

system of the Office of Public Defense Services (OPDS). Those files were cross-referenced with records available online through the Oregon Judicial Information Network (OJIN) and, for older cases, through direct contact with county courts. After identifying the pool of eligible cases, the OCRC contacted county courts, requesting indictments, judgments, jury instructions and verdict forms. Relying on these forms, the OCRC developed a comprehensive database in Access that included information on defendants, co-perpetrators, victims, charges, verdicts, and jury responses to sentencing phase questions. By the fall of 2010, the database included 164 cases that had proceeded to the sentencing phase of aggravated murder trials.

Further restrictions were placed on the sample for the current study. Because of its primary focus, the study includes only those cases wherein jurors responded to the "PCAV" sentencing query: "whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society." This resulted in the exclusion of cases ($n = 18$) where guilty pleas or directed verdicts precluded the possibility of a death sentence, leaving juries to consider only the issue of whether the life sentence given to the defendant would allow parole eligibility. For defendants subjected to more than one penalty phase trial, only the initial jury determination was included in the present study, thus excluding subsequent cases ($n = 21$) for individual defendants.[3] Aggravated murder cases involving prison/institutional killings were excluded ($n = 3$) because that type of killing "begs the question" of violence propensity in a prison context. The current study also excluded cases involving female defendants ($n = 4$), for whom the conditions of confinement vary considerably from that of their male counterparts. Finally, a small number of cases ($n = 3$) were excluded because the defendants spent too little time incarcerated in the ODOC to assess behavioral outcome measures.[4]

The final sample consisted of 115 male inmates sentenced to either life ($n = 65$) or death ($n = 50$) during the sentencing phase of an aggravated murder trial in Oregon from 1985 through 2008. With the exception of those who died of natural causes ($n = 5$), were executed ($n = 2$), moved to another state for safekeeping ($n = 2$), or released due to an overturned conviction ($n = 1$), their behavior in prison was retrospectively coded from their entrance to prison through 2010.[5] The average time at risk ranged from 2.3 to 25.2 years, with a mean time served of 15.3 years (SD =6.4 years). The average time served was slightly longer for those sentenced to death ($M = 16.1$ years) than for those sentenced to life ($M = 14.7$ years). Although the bulk of death-sentenced inmates' time was served on death row ($M = 11.1$ years), more than one-third ($n = 18$) obtained relief from their death sentences and ended up serving a significant portion of time ($M = 13.9$ years) among the broader prison population.

[3] An exception to this rule was made for a defendant who received a life sentence for one murder, followed by a death sentence for another during the following year. He has remained on death row since that time.

[4] None of the excluded inmates, those convicted of murder while incarcerated or the female inmates, killed in prison during the observation period. The remaining three excluded cases did not serve time in ODOC during the observation period. One of these situations arose as a result of a verdict being overturned and the case dismissed. Another exclusion resulted from a defendant being transported to another state to face a capital murder charge for killing a police officer, where he was ultimately convicted and sentenced to death. In the final case, the defendant did not enter the prison system until after the 12/31/08 cutoff date.

[5] The observation period was cut short for one additional inmate who, while serving a life sentence for aggravated murder, murdered a fellow inmate. Physical inmates' folders could not be retrieved for some of the cases sentenced during the earliest years of the series. Hence, 12 cases were missing disciplinary data for the time period prior to the advent of computerization, representing 6 years of their carceral term on average.

DOI: 10.1002/bsl

Measures

This study focused on the ability of juries to accurately predict continuing acts of violence among defendants whose lives hang in the balance. A jury's response to the PCAV sentencing question is the "prognosis," while a defendant's subsequent behavior serves as the criterion for determining the accuracy of the prognosis. The operationalization of a jury's prognosis is a straightforward dichotomous determination – yes or no – retrieved from the OCRC database. The criterion is a bit more complicated. What constitutes "continuing acts of violence" is not as easily defined. Although the outcome could be hypothesized as theoretical future acts of violence in the community (see earlier discussion), the pragmatic reality is that convicted aggravated murderers for whom future community risk is of any concern will spend their lives in prison whether sentenced to death or LWOP. Further, only six ODOC inmates (of any offense/sentence) have escaped from high-security Oregon prisons in 30 years, and these were quickly recaptured. Although parole eligibility after 30 mandatory years of incarceration is a sentencing option, it has been issued in less than 10% (five of 52) of the cases in the current sample where the jury has been queried. Of this very limited pool of parole-eligible aggravated murderers in Oregon, the first does not reach parole eligibility until 2014 (H. Dinh, personal communication, 20 September 2010). Accordingly, the most appropriate and the only available criterion, then, are within the context of the prison environment.

The measurement of "violence" in the prison context requires further clarification. Though ODOC classifies certain disciplinary offenses as major acts of misconduct (level 1 offenses), the designation is based on a host of factors, such as risk to the security of the institution, in addition to the actual level of violence involved. In fact, most of the level 1 offenses (e.g., possession of a weapon, disturbances, racketeering, arson), though involving the potential for violence, seldom result in physical harm to other inmates or staff members. If an inmate commits a serious violation of the sort listed above while harming another inmate or staff member, they are also charged with an assault. For this reason, assaults are viewed as the best indicator of violence in the institution.[6]

The classification of infractions as assaults, as well as whether they are considered major or minor, depends on several factors aside from the level of resulting harm; and the operational definitions shifted during the course of the observation period. For these reasons, a standardized measure of the level of harm resulting from assaultive inmate misconduct was employed (see Sorensen et al., 2011). The extent of injury resulting from assaultive actions, coded from electronically scanned disciplinary reports, was measured on the following scale: *no injury* – no injuries noted; *outpatient minor injury* – cuts, bruises, abrasions, muscle strains, and injuries requiring tetanus shots; *outpatient moderate injury* – lacerations requiring sutures, broken or knocked out teeth, broken and dislocated bones, and concussions; *in-patient serious injury* – hospitalization for non-life threatening injuries and injuries requiring reconstructive surgery; *in-patient severe injury* – hospitalization with life-threatening injuries; and *fatal injury* – injuries resulting in death.

Measuring the level of harm in such detail offered several possible violent outcome measures. The initial plan was to operationalize "criminal acts of violence" as assaults resulting in moderate to fatal injuries. Upon examining the sample's disciplinary files,

[6] While other level 1 offenses might be considered violent, particularly sexual assault and hostage-taking, none of the inmates in the sample committed those serious acts of misconduct.

DOI: 10.1002/bsl

however, it was clear that this definition failed to capture some inmates with lengthy records of serious and assaultive misconduct. In reading the reports, it became obvious that some of these minor assaultive acts could easily have resulted in more harm if, for instance, the fight had not been broken up quickly, or if the victor had a more powerful left jab. Based on this review, it was decided that the definition of violence should be expanded to include inmates who committed three or more assaultive acts of misconduct, regardless of the level of harm resulting from those acts. The primary outcome measure, then, refers to inmates who have committed either injurious or repetitive assaults. Other outcome measures were also examined to insure the robustness of findings.

Additional control measures include characteristics of the defendant previously identified as potentially relevant correlates of violence (Cunningham, Reidy, & Sorensen, 2008), criminal history, and elements of the aggravated murder case. Gathered from electronic ODOC prisoner files, the defendant characteristics include race, age on entrance to prison, and indicators of drug/alcohol problems and mental health disturbances. Race is self-reported as Black, White, or Other. Drug/alcohol problems were determined by testing in prison either at intake or later during the course of their prison term. The ODOC drug/alcohol needs scale ranges from none to severe problems. Inmates were only tested if initial screening indicated a need for it. Those not tested and inmates shown to have none or only minor problems after testing were combined for the purpose of this research. Those with moderate to severe drug/alcohol problems were also grouped together herein. Mental health indicators were handled in a similar fashion, so that the lowest designated category included inmates with no disorder after assessment, personality disorders, and minor mental illness diagnoses [e.g., adjustment disorders, attention deficity hyperactivity disorder (ADHD), phobias] along with the inmates not targeted for testing. Major mental illness diagnoses [e.g., *Diagnostic* and *Statistical Manual* of *Mental Disorders* (DSM) Axis I psychotic and mood disorders], with or without medication, were combined. Those designated as mentally retarded were included as a separate category. Aside from the lowest designation, the latter two categories were not mutually exclusive, in that an inmate could have been diagnosed with a major mental illness and mental retardation.

Prior criminal history data were gather from the Oregon State Police (OSP) and matched to prison files by an identifier supplied by ODOC. These records include arrests in Oregon dating back to the mid-1970s. Numerous variables were culled from these records, including the raw number of prior "person" (violent) arrests. Inmates were later categorized as having none, one to two, or three or more prior violent arrests for the purpose of analysis. The presence or absence of a violent arrest in the previous 5 years was included as measure of recency. Dichotomous indicators of prior rape/sexual assault, robbery, and assault arrests were also extracted from the database. Two variables available from the ODOC also contained criminal history measures, including record of a prior ODOC incarceration and revocation from community supervision.

Features of the aggravated murder case and its disposition were culled from files provided by the OCRC. The first set of indicators noted various elements of the aggravated murder specified in charging and sentencing documents. These include whether the defendant had a prior homicide conviction, and whether the case involved multiple victims, maiming or torture, victims under the age of 14, victims performing their official duties as officers of the court, a contemporaneous felony, efforts to conceal the crime or perpetrator's identity, or an escape from a penal institution. The element concerning the killing of another inmate was left out for the reason discussed earlier, as were elements not relevant to any of the current cases. The presence of co-perpetrators and gender of the victims were also

DOI: 10.1002/bsl

extracted from the OCRC database, as was the cause of death (e.g. gunshot, stabbed/lacerated, other). The disposition, either death or life, along with any further modification of the sentence (e.g., death reduced to life), and other changes to the inmates' status (e.g., death from natural causes, release from prison, execution), and the dates of each, were also recorded.

FINDINGS

Table 2 provides a description of the characteristics of aggravated murder defendants broken down by jury determinations of future violence risk. Among the sample of cases, more than two-thirds of the initial penalty phase verdicts resulted in affirmative responses to the PCAV sentencing query. The available characteristics of the defendant, including age and race, were not related to the rate at which juries answered the PCAV query in the affirmative. The only exception was for defendants over the age of 40, for whom juries predicted a higher likelihood of future violence potential. While perhaps unknown to juries at the time, defendants later identified as having moderate to severe drug/alcohol problems drew more negative responses to the PCAV inquiry. Those later diagnosed with severe mental illness, particularly those requiring medication, drew more affirmative responses.

Table 3 presents information related to the criminal history of aggravated murder defendants by jury finding of likelihood of future violence. The results show that jurors were less likely to respond affirmatively in the absence of a prior record of violent crimes,

Table 2. Characteristics of aggravated murder defendants by jury finding of future violence

	Jury finding that future violence likely			
	No		Yes	
	n	%	*n*	%
Total cases	37	32.2%	78	67.8%
Race				
White	25	30.5%	57	69.5%
Black	6	33.3%	12	66.7%
Other	6	40.0%	9	60.0%
Age at entrance to prison				
18–24 years	7	31.8%	15	68.2%
25–29 years	8	44.4%	10	55.6%
30–34 years	10	37.0%	17	63.0%
35–39 years	7	35.0%	13	65.0%
≥ 40 years	5	17.9%	23	82.1%*
Drug/alcohol problems				
None to minor/not identified	18	26.5%	50	73.5%
Moderate to severe	19	40.4%	28	59.6%†
Mental health indicators				
None to minor/not identified	32	38.1%	52	61.9%
Major mental illness (MI) diagnosis	2	13.3%	13	86.7%†
Major MI diagnosis requiring medication	0	0.0%	11	100.0%*
Mental retardation	4	33.3%	8	66.7%

†$p < 0.10$; *$p < 0.05$; **$p < 0.01$.

DOI: 10.1002/bsl

Table 3. Criminal history of aggravated murder defendants by jury finding of future violence

	Jury finding that future violence likely			
	No		Yes	
	n	%	*n*	%
Total cases	37	32.2%	78	67.8%
Prior person crime arrests				
None	24	42.9%	32	57.1%*
1 to 2	8	22.9%	27	77.1%†
≥ 3	2	14.3%	12	85.7%†
Rape/sexual assault	2	22.2%	7	77.8%
Robbery	1	7.1%	13	92.9%*
Assault	6	20.7%	23	79.3%†
Any during 5 years prior to current arrest	7	21.2%	26	78.8%†
Prior prison incarceration				
No	31	36.5%	54	63.5%
Yes	3	12.0%	22	88.0%*
Revocation				
No	30	33.3%	60	66.7%
Yes	4	20.0%	16	80.0%

†$p < 0.10$; *$p < 0.05$; **$p < 0.01$.

and more likely to respond affirmatively when the defendant had prior violent crime arrests, especially for robbery or assault. Those with a prior record of incarceration were also more likely to receive an affirmative jury response to the PCAV inquiry. While it appears that those receiving a community service revocation were more likely to garner an affirmative response, the results were not statistically significant.

As shown in Table 4, certain aspects of the killings were related to juries' responses to the PCAV query. In comparing specific categories, the overall likelihood of an affirmative PCAV response (67.8%) should be considered for reference. Murders for hire, like crimes involving co-perpetrators more generally, elicited fewer affirmative responses. Killings involving maiming or torture and contemporaneous felonies were more likely to draw affirmative responses. Responses indicate that defendants who killed with a gun were seen as having less of a propensity toward future violence than defendants who killed with a knife, and particularly those who relied on other methods. Of the 78 affirmative responses, 50 (64.1%) resulted in death sentences, whereas only 28 out of 65 (43.1%) life sentences resulted after an affirmative response to the PCAV query. Further, the rate of affirmative answers to the PCAV query was not affected by the 1989 change to the statute, which added a sentencing query concerning mitigation, and also a "natural life" (i.e., LWOP) sentencing option.

Table 5 presents a comparison of the rate and prevalence of disciplinary misconduct by jury finding related to future violence risk. The first row shows that when all types of rule violations are considered, rule misconduct is quite common among the inmates, reaching a rate of nearly 50 per 100 inmates per year, and involving nearly 90% of all aggravated murderers. The next two rows show that serious acts (level 1 violations) of misconduct and assaultive infractions are much less common, with a gross rate of around six per 100 per year, yet still involve about one-third of inmates across their terms of

DOI: 10.1002/bsl

imprisonment. The remaining rows show that as the level of injury resulting from an assaultive infraction increases, the occurrence of such offenses becomes far less frequent. In fact, greater than 90% (97 out of 105) of the assaultive rule violations resulted in either no reported harm or only minor injuries.[7] These offenses typically involved mutual fights among inmates, or a push or kick to a guard while resisting restraints, which were coded as involving "minor injury" if abrasions or bruises were reported, or if a trip to the infirmary was called for. Spitting on guards or dousing them or fellow inmates with bodily fluids, including urine, semen, and "shit bombs," was also common among the "no injury" assaults. Attempted assaults of staff were prominent among this category.

When considering only those assaultive acts resulting in moderate injury or greater, the rate of occurrence decreased to less than five per 1,000 inmates annually. None of these injurious assaults involved prison staff members. Among the assaults resulting in moderate injuries, two involved fist fights requiring a trip to the hospital and sutures. Four others involved the same treatment, but were more severe because weapons were used to either stab or slash victims, the most serious of which resulted in a 9-inch cut that penetrated the victim's fatty tissue. The one assault categorized as involving "serious injury" was a fist fight, which resulted in a broken jaw, requiring surgery and a brief hospital stay; this same aggressor was also responsible for one of the attacks on an inmate with a weapon. The most serious incident resulted in a homicide by an inmate serving a life sentence for aggravated murder after his jury had answered the PCAV inquiry negatively; this aggressor was also involved in one of the attacks on an inmate with a weapon.

While the details of these violent acts are quite disturbing, the figures presented in Table 5 point to some general conclusions concerning violence among incarcerated aggravated murderers. The first, already mentioned, is that assaultive acts, particularly those involving significant injuries, were quite rare. The second is that jury prognoses concerning defendants' propensity to commit further acts of violence appear to be completely unrelated to the actual commission of such acts. The rate of serious and assaultive behaviors was almost identical for the two groups; none of the differences were statistically significant. As noted earlier, the only fatality resulted at the hands of an inmate for whom the jury predicted that future violence was not likely.

Table 6 presents a contingency table with jury predictions related to the violence potential of aggravated murder defendants and the defendants' actual behavioral outcome while incarcerated. The main outcome measure included injurious/repetitive assaults. Positive outcomes included the group of violent inmates described above who committed assaults ranging from moderately injurious to fatal, and an additional six inmates who committed three or more assaults. The contingency table shows that although juries predicted that more than two-thirds of the aggravated murder defendants were likely to "commit criminal acts of violence that would constitute a continuing threat to society," only one-tenth (12 out of 115) actually committed qualifying violent acts. The findings on the prevalence of violence among sampled inmates from Table 6 agree with the figures from Table 5, both showing that violent outcomes were completely unrelated to

[7] The number of assaults was adjusted for three individual "outlying inmates" to avoid skewing the results. Rather than removing the outliers resulting in a loss of data for those cases, the maximum value was set at 50 for total violations and 10 each for serious and assaultive violations. The most extreme outlier had racked up more than 300 disciplinary violations and 50 level 1 assaults on guards, which nearly always involved spitting or attempted assaults, only one of which resulted in a "minor injury," an abrasion to the forearm. One of the other inmates had numerous level 1 disturbances, and another had more than 50 garden-variety disciplinary offenses.

DOI: 10.1002/bsl

Table 4. Aggravated murder case characteristics by jury finding of future violence

	Jury finding that future violence likely			
	No		Yes	
	n	%	*n*	%
Total cases	37	32.2%	78	67.8%
Aggravating murder elements				
Solicited – for hire	9	75.0%	3	25.0%**
Prior homicide conviction	1	50.0%	1	50.0%
Multiple victims	6	23.1%	20	76.9%
Maiming or torture	0	0.0%	9	100.0%*
Victim under the age of 14 years	0	0.0%	2	100.0%
Victim performing official duties	1	16.7%	5	83.3%
During commission of a felony	17	25.4%	50	74.6%†
Effort to conceal crime or identity	14	30.4%	32	69.6%
While on escape from penal facility	1	33.3%	2	66.7%
Co-perpetrators involved	14	43.8%	18	56.3%†
Female victim	22	29.3%	53	70.7%
Cause of death				
Gunshot	23	40.4%	34	59.6%†
Knife/cut	11	32.4%	23	67.6%
Other	8	24.2%	25	75.8%
Sentence				
Life	37	56.9%	28	43.1%
Death	n/a		50	100.0%
Statute in effect				
Before July 19, 1989	12	33.3%	24	66.7%
After July 19, 1989	25	31.6%	54	68.4%

†$p < 0.10$; *$p < 0.05$; **$p < 0.01$.

Table 5. A comparison of the yearly rate per 100 inmates and prevalence of disciplinary misconduct by jury finding of future violence

	Jury finding that future violence likely				Aggravated	
	No		Yes		murderers	
	(*n* = 37)		(*n* = 78)		(*N* = 115)	
Disciplinary misconduct	Rate[a]	(prevalence)	Rate[a]	(prevalence)	Rate	(prevalence)
Any disciplinary violations	47.7	(81.1%)	46.9	(89.7%)	47.2	(87.0%)
Serious (level 1) misconduct	5.9	(37.8%)	6.8	(37.2%)	6.5	(37.4%)
Assaultive infractions	5.9	(43.2%)	5.9	(37.2%)	6.0	(39.1%)
No injury	4.3	(37.8%)	3.1	(24.4%)	3.5	(28.7%)
Outpatient minor injury	1.2	(13.5%)	2.4	(20.5%)	2.0	(18.3%)
Outpatient moderate injury	0.4	(5.4%)	0.3	(6.4%)	0.3	(6.1%)
In-patient serious injury	0.0	(0.0%)	0.1	(1.3%)	0.1	(0.9%)
In-patient severe injury	0.0	(0.0%)	0.0	(0.0%)	0.0	(0.0%)
Fatal injury	0.2	(2.7%)	0.0	(0.0%)	0.1	(0.9%)

[a]*t*-tests non-significant for comparisons between group rates for each type of infraction.

DOI: 10.1002/bsl

Table 6. Jury predictions of future violence and prison behavioral outcomes among aggravated murderers

Jury prediction	Actual behavioral outcome	
	Assaultive	Not assaultive
Future violence likely	True positive	False positive
($n = 78$)	8 (10.3%)[a]	70 (89.7%)
	1 (1.3%)[b]	77 (98.7%)
Future violence unlikely	False negative	True negative
($n = 37$)	4 (10.8%)[a]	33 (89.2%)
	1 (2.7%)[b]	36 (97.3%)

[a]Indicates assaults resulting in moderate to fatal injury or three or more assaults at any level; MCC(Φ) = –0.01 ($p = 0.581$).
[b]Indicates assaults resulting in in-patient serious to fatal injury: MCC(Φ) = –0.05 ($p = 0.542$).

predictions made by jurors.[8] The observation that naturally flows from such a finding is that juries were right 90% of the time when predicting that future violence was not likely, and wrong 90% of the time when they predicted that future violence was likely. When the outcome is restricted to only inmates committing assaults resulting in serious to fatal injury, then the findings indicate that juries were nearly always right (97% of the time) when predicting that future violence was not likely, and nearly always wrong (99% of the time) when predicting that future violence was likely.[9]

Additional analyses presented in Table 7 were completed to test the robustness of these findings. First, the contingency table was replicated using less stringent measures of "violence." When the behavioral outcome threshold was lowered to include either the commission of any assaultive act of misconduct or any serious (level 1) violation, the overall prevalence of "violence" in the sample increased to 39.1% and 37.2%, respectively. Regardless, swapping the behavioral outcome in the contingency table failed to provide any improvement in juries' predictive performance. The prevalence rates of assaultive or serious acts were nearly identical for the violence-predicted and nonviolence-predicted groups, mimicking the rate comparisons for the more broadly defined outcomes in Table 5.

Second, differences between the groups in their opportunity to commit violent acts had to be considered. Each of the subjects for whom a jury's response was negative to the PCAV inquiry entered the general prison population, while two-third of the subjects for whom a jury's response was positive to the PCAV inquiry were sent to death row. It is plausible that the more restrictive conditions of confinement on death

[8] To insure that the missing data did not influence results, figures were re-computed excluding cases missing disciplinary data for a portion of their incarceration. The findings changed very little when the 12 cases were dropped from the analysis. For example, the rates of disciplinary infraction for the first three categories in Table 5 by jury finding future violence became: any disciplinary (48.3, 48.9, 48.7); serious (level 1) misconduct (6.5, 6.9, 6.8); and assaultive infractions (6.5, 6.4, 6.5). The results in Table 6 also remained substantively the same: true positive, eight (11.1%); false positive, 64 (88.9%); false negative, four (12.9%); and true negative, 27 (87.1%).
[9] Matthews Correlation Coefficient (MCC) (Baldi, Brunak, Chauvin, Anderson, & Nielsen, 2000) was utilized to test the strength of association between the prediction and the outcome (i.e., predictive performance) for the respective 2×2 matrices in Table 6. The resulting correlation coefficient is always between –1 and +1, with a value of –1 indicating total disagreement and +1 indicating total agreement. The correlation coefficient is 0 for completely random predictions. The MCCs for the contingency table were near zero (–0.01 and –0.05). Their performance shows that juror predictions were no better than random guesses.

DOI: 10.1002/bsl

row had the effect of reducing the opportunity to commit assaultive acts among a large portion of the violence-predicted inmates, thus serving to confound the relationship between the PCAV response and behavioral outcomes. The figures bear out this possibility to an extent. While 17.9% of the violence-predicted, life-sentenced inmates behaved violently, only 8.0% of the death-sentenced inmates, all of whom were violence-predicted, behaved violently. The lower prevalence of violence among death-sentenced inmates, however, could not be solely attributed to a reduced opportunity. One of the moderately injurious assaults described above occurred on death row when an aggressor entered another inmate's cell and slashed him multiple times with a bladed weapon. This event shows that freedom of movement was not completely curtailed on death row, nor was access to materials that could be used to fashion a weapon. Further, violence among formerly death-sentenced inmates did not increase after being transferred to the general prison population as a result of overturned sentences. Only one of the 18 (5.6%) former death row inmates behaved violently in the broader prison population.

A final attempt to rule out the more restrictive death row setting as a potential confound separated violence-predicted inmates serving time in the general prison population (GP) – life-sentenced and former death row inmates ($n = 46$) – from those serving time only on death row (DR; $n = 32$). The prevalence rates of violence were very similar between the group of violence-predicted inmates serving time in the GP (10.9%) and those serving time only on DR (9.4%). Jury prognosis and housing assignment were then included as interaction terms, PCAV-YES*GP and PCAV-YES*DR, in a logistic regression model. Compared with the exclude category PCAV-NO*GP, neither of the interaction terms was significant. This indicates that jury prediction of future violence in combination with housing assignment was not related to behavioral outcome in prison. Several control variables that were most strongly related to violent outcome were included in the model. A forward selection procedure was employed for the remaining control variables, but none added to the predictive power of the model. While the overall model was statistically significant ($r^2 = 0.255$; $x^2 = 14.596$; $p < 0.05$), most of the individual coefficients were not, primarily due to the size of the sample. Their influence, however, was in the expected direction, with age of 40 years and over being associated with a lower

Table 7. Logistic regression predicting prison violence among aggravated murderers

Variables	B	S.E.	Wald	Sig.	Exp (B)	95% C.I. for EXP(B)	
						Lower	Upper
Time at risk	0.035	0.059	0.343	0.558	1.035	0.922	1.163
PCAV-YES*DR	0.055	0.931	0.003	0.953	1.056	0.170	6.545
PCAV-YES*GP	–0.639	0.871	0.539	0.463	0.528	0.096	2.908
Age 40 plus	–0.661	1.212	0.298	0.585	0.516	0.048	5.553
MHR	2.285	0.859	7.071	0.008	9.831	1.824	52.986
Prior person ≥ 3	1.296	0.836	2.402	0.121	3.654	0.710	18.808
Contemp. felony	1.309	0.832	2.472	0.116	3.701	0.724	18.921
Constant	-3.884	1.297	8.967	0.003	0.021		

Time at risk, years spent in prison; PCAV-YES*DR, interaction term indicating an affirmative response to PCAV and time served only on death row; PCAV-YES*GP, interaction term indicating an affirmative response to PCAV and time served in the general prison population (the omitted reference category is PCAV-NO-GP); age 40 plus, inmates age 40 and over at entrance to prison; MHR, diagnosis of mental retardation; prior person ≥ 3, three or more violent arrests prior to the current offense; contemp. felony, murder occurred during commission of a felony; Sig., significance level.

DOI: 10.1002/bsl

likelihood of violent behavior, while three or more prior person offenses, a mental retardation (MHR) diagnosis, and contemporaneous felony were associated with a higher likelihood of violent prison misconduct.

DISCUSSION

The ability of capital juries at the sentencing phase of aggravated murder trials to accurately predict future prison violence was examined through a retrospective review of the disciplinary records of 115 male inmates sentenced to either life ($n = 65$) or death ($n = 50$) in Oregon from 1985 through 2008. The study had a number of strengths: (1) the cases subjected to follow-up were both large in number and represented the universe of Oregon cases for a 23-year period meeting study criteria where capital juries made future violence predictions regarding male inmates who had committed capital murders in the community; (2) by following inmates whose juries had endorsed or rejected a prediction of future violence, a bidirectional test of jury accuracy resulted; (3) the time at risk in prison averaged 15.3 years, reflecting ample opportunity for violence proclivity to be behaviorally expressed; (4) the dependent variable, assault in prison, was disaggregated by severity and level of injury to inform proportionality considerations; (5) the emphasis on assaults with injuries reduced error that might have been associated with unreported violence; (6) both offender and capital offense characteristics were subjected to analysis to illuminate the bases of the jury predictions; and (7) the analyses controlled for conditions of confinement.

Violence Prediction Accuracy

Oregon capital juries exhibited chance-level predictive performance, mirroring that of Texas capital juries recently reported by Cunningham et al. (2010). In both jurisdictions utilizing the *Jurek*-affirmed special issue, the predictions of the juries represented no improvement over base rates. In other words, these offenders exhibited base rates of a 10% incidence of serious violence and a 2% incidence of violence resulting in life-threatening injury. Consistent with random guesses, predictions that offenders would be violent were in error 90–98% of the time, depending on the severity of violence specified. Similarly, through no special predictive talent, rejection of the special issue was accurate 90–98% of the time. The abysmal performances of Oregon and Texas juries in predicting future violence are remarkably similar to "alarming error rates" recently reported for federal capital jurors (Cunningham et al., 2009). The violence predictions by jurors in all of these studies were no better than random guesses.

The inability of Oregon juries to demonstrate greater than chance-level accuracy was a consistent finding, regardless of the severity of violence specified. Thus, while the error rate of the violence predictions increased with the severity of assault specified, at no level of assault did the juries demonstrate a better than chance ability to specify the particular defendants who would go on to commit that violence. The poor violence prediction performance of these Oregon juries cannot be attributed to the heightened security of death row. Controlling for individual characteristics, elements of the capital offense, and prior community criminality, conditions of confinement were not found to impinge on the accuracy of jury assessments.

DOI: 10.1002/bsl

Consistent with prior studies demonstrating a very low base rate of serious prison violence among capital offenders in other jurisdictions (Cunningham 2010; Cunningham et al., 2008; Cunningham et al., 2009; Cunningham et al., 2011; Edens et al., 2005; Reidy, Sorensen, & Cunningham, 2012; Sorensen & Cunningham, 2010), the current findings demonstrate that assaultive acts among aggravated murderers in Oregon, particularly those involving significant injuries, are quite rare, with the prevalence decreasing with the severity of injury. The majority of assaultive infractions (90%) resulted in no injury or only minor injuries. Only 6.1% of the aggravated murderers committed violence resulting in moderate injuries, such as lacerations requiring sutures, fractured bones, or concussions, and none of these moderate injuries necessitated in-patient hospitalization. Two percent committed assaults resulting in in-patient serious or fatal injuries. These low base rates are a key factor in the poor predictive performance of Oregon juries, i.e., it is very difficult to predict a low base rate behavior.

Public Policy Implications

Generalizing from contexts such as civil commitment, sentencing, and parole – where a risk of violence is assumed to be contained and public safety assured by confinement (i.e., therapeutic, correctional) – the U.S. Supreme Court in *Jurek v. Texas* (1976) affirmed that jurors could make determinations of future violence in all contexts (in prison, as well as in the community), with sufficient reliability (in the legal sense) to make this prediction a part of life and death sentencing. In the decades following *Jurek*, the availability of LWOP sentencing as an alternative to the death penalty in Oregon and every other death penalty jurisdiction has made prison the only realistic context for any future violence prediction.

Acknowledging the elegance of the intuitive logic suggesting that jurors should be able to make these *Jurek* judgments regarding future violence in prison, the critical issue would seem not to be whether jurors should be able to accurately make these predictions, but rather whether they can do so. The continuing presence of the special issue as an essential element in the death penalty statutes in Oregon and Texas has allowed a direct test of the *Jurek* Court's confidence in the predictive capabilities of capital juries. Follow-up of jury predictive performance in these two jurisdictions in the modern prison era, with offender prison tenures averaging 15.3 years (Oregon) and 18.3 years (Texas, see Cunningham et al., 2010), demonstrates that they fail miserably. Jurors in both Texas and Oregon, over 20-year periods and involving an aggregate follow-up of 226 case predictions ($n = 115$ Oregon; $n = 111$ Texas), were unable to more than guess at future conduct.

The inability of capital juries to accurately forecast future prison violence is not limited to the "special issue" jurisdictions of Oregon and Texas. Rather, this predictive incapacity extends to capital contexts where *Jurek* is a precedent for the availability of future violence as an aggravating factor at death penalty sentencing. As referenced earlier, 72 federal capital juries nationwide also made chance-level predictions where future dangerousness had been alleged as a non-statutory aggravating factor (see Cunningham et al., 2009). Thus, jurors perform no better when considering future violent prison conduct as an aggravating factor than they do as a special issue.

The consistency of the data in several large modern-era studies, with extended follow-up, in multiple jurisdictions establishes as a robust and well-generalized finding the inability of capital juries to predict prison violence. The confidence of the *Jurek* Court in the future violence judgments of capital juries has been dispelled.

DOI: 10.1002/bsl

The inability of capital juries to predict future prison violence has implications beyond a high rate of error. In other words, it is not simply that capital juries are wrong 90% of the time in predicting serious assaults and virtually always wrong in anticipating life-threatening assaults; rather, these juries show no improvement over random guesses. In other words, no authentic "individualization" is occurring regarding the only setting where the offender could exhibit violence. Such arbitrary guesses are the antithesis of a reasoned, individualized death penalty determination.

Correcting a mythology of jury predictive capability, at least regarding future prison violence, that has become reified in death penalty jurisprudence is a complex undertaking. This is particularly the case in jurisdictions where the meaning of the special issue has been expansively interpreted to variously reference violence in prison, violence in the community under hypothetical circumstance, the inclinations of "bad" character, and/or a catch-all integration of any and all sentencing considerations. The most obvious remedy is for *Jurek* to be revisited by the Court. Withdrawing "future dangerousness" as a death penalty consideration would be a pragmatic response to jury predictive failure, the realities of LWOP sentencing, and the incapacitation capabilities of modern corrections departments. It would also free the integration of mitigating and aggravating factors from a contaminating fiction of future conduct, arguably enhancing the reliability of these determinations.

Should *Jurek* not be superseded, in the face of these now well-replicated findings, a number of difficult public policy questions remain. Is it permissible or even mandatory that capital jurors be informed that, however the special issue is construed, they are incapable of predicting serious prison violence and its occurrence is quite rare? Should the prosecution be allowed to assert in argument "predictors" or "probabilities" for future prison violence that are refuted by scientific consensus? If statutes or case law invoke illusory correlations of offense features or character factors for predicting prison violence that are controverted by available science, do they remain constitutional? What will be the standard for review of death sentences that followed testimony or argument that misinformed the jury regarding the probability that the defendant would perpetrate serious prison violence? How wide must the gulf grow between science and jurisprudence before a rapprochement occurs?

DISCLOSURE

The authors derive income from evaluations and testimony at capital sentencing, specifying varying levels of improbability of future prison violence and/or describing prison classification and security. The withdrawal of future violence as a capital sentencing consideration would have an adverse economic impact on the authors. The Oregon Indigent Defender Fund at the request of defense counsel provided funding for data analysis by the second author.

ACKNOWLEDGEMENTS

The authors acknowledge and thank the Oregon Department of Corrections and Oregon Capital Defense Resource Counsel for providing assistance in data collection. We wish to especially acknowledge the contributions of Paul Bellatty, Tom O'Connor, Jeff Duncan, and Huong Dinh. The views expressed in this paper are not necessarily those of the Oregon Department of Corrections.

DOI: 10.1002/bsl

REFERENCES

Baldi, P., Brunak, S., Chauvin, Y., Anderson, C. A. F., & Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: An overview. *Bioinformatics Review*, *16*, 412–424.

Blume, J. H., Garvey, S. P., & Johnson, S. L. (2001). Future dangerousness in capital cases: Always "at issue". *Cornell Law Review*, *86*, 387–399.

Burns v. Commonwealth, 261 Va. 307, 541 S.E.2d 872, cert. denied, 534 U.S. 1043, 122 S.Ct. 621, 151 L. Ed.2d 542 (2001).

Claussen, A. M., Pearce, M. W., & Schopp, R. F. (2004). Dangerousness, risk assessment, and capital sentencing. *Psychology, Public Policy and Law*, *10*, 471–491.

Coble v. State, 330 S.W. 3d 253 (Tex. Crim. App. 2010).

Costanzo, S., & Costanzo, M. (1994). Life or death decisions: An analysis of capital jury decision making under the special issues sentencing framework. *Law and Human Behavior*, *18*, 151–168.

Cunningham, M. D. (2006). Dangerousness and death: A nexus in search of science and reason. *American Psychologist*, *61*, 828–839. doi: 10.1037/0003-066x.61.8.827

Cunningham, M. D. (2010). Evaluation for capital sentencing. New York, NY: Oxford University Press.

Cunningham, M. D., & Reidy, T. J. (1998). Integrating base rate data in violence risk assessments at capital sentencing. *Behavioral Sciences & the Law*, *16*, 71–95.

Cunningham, M. D., Reidy, T. J., & Sorensen, J. R. (2008). Assertions of "future dangerousness" at federal capital sentencing: Rates and correlates of subsequent prison misconduct and violence. *Law and Human Behavior*, *32*, 46–63. doi:10.1007/s10979-007-9107-7

Cunningham, M. D., Sorensen, J. R., & Reidy, T. J. (2009). Capital jury decision-making: The limitations of predictions of future violence. *Psychology, Public Policy, and Law*, *15*, 223–256. doi:10.1177/1073191104272815

Cunningham, M. D., Sorensen J. R., Vigen, M. P., & Woods, S. O. (2010). Life and death in the lone star state: Three decades of violence predictions by capital juries. *Behavioral Sciences & the Law*, *29*, 1–22. doi:10:1002/bsl.963

Cunningham, M. D., Sorensen J. R., Vigen, M. P., & Woods, S. O. (2011). Correlates and actuarial models of assaultive prison misconduct among violence-predicted capital offenders. *Criminal Justice and Behavior*, *38*, 5–25. doi:10.1177/0093854810384 830

Daubert v. Merrill Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668–1674.

Edens, J., Buffington-Vollum, J., Keilin, A., Roskamp, P., & Anthony, C. (2005). Predictions of future dangerousness in capital murder trials: Is it time to "disinvent the wheel?" *Law and Human Behavior*, *26*, 59–87. doi: 10.1007/s10979-005-x

Estrada v. State, Adrian ESTRADA, Appellant v. The State of Texas. No. AP-75634 (Tex. Crim. App. 2010)

Garvey, S. P. (1998). Aggravation and mitigation in capital cases: What do jurors think? *Columbia Law Review*, *98*, 1538–1576.

Guy, L. S., Edens, J. F., Anthony, C., & Douglas, K. S. (2005). Does psychopathy predict institutional misconduct among adults? A meta-analytic investigation. *Journal of Consulting and Clinical Pyschology*, *73*, 1056–1064. doi: 10.1037/0022-006X.73.6.1056

Hall, H. V. (1987). Violence prediction: Guidelines for the forensic practitioner. Springfield. Il: Charles C. Thomas.

Idaho Code. (§19-2515). Title 19 - Criminal Procedure; Chapter 25 - Judgment, Sentence in capital cases

Illinois (2002). Governor's Commission on Capital Sentencing. State of Illinois, Chicago.

Jurek v. Texas, 428 U.S. 153 (1976).

Kansas v. March 548 U.S. 163 (2006).

Krauss, D. A., & Sales, B. D. (2001). The effects of clinical and scientific expert testimony on juror decision making in capital sentencing. *Psychology, Public Policy, and Law*, *7*, 267–310.

Krauss, D. A., McCabe, J. G., & McFadden, S. (2009). Limited expertise and experts: Problems with the continued use of future dangerousness in capital sentencing. In R. F. Schopp, R. L. Weiner, B. H. Bornstein, & S.L. Willborn (Eds.), Mental disorder and criminal law: Responsibility, punishment, and competence (pp. 135–157). New York: Springer.

Marquart, J.W., Ekland-Olson, S., & Sorensen, J. R. (1989). Gazing Into the crystal ball: Can jurors accurately predict dangerousness in capital cases? *Law & Society Review*, *23*, 449–468.

Marquart, J. W., Ekland-Olson, S., & Sorensen, J. R. (1994). The Rope, the chair, and the needle: Capital punishment in Texas, 1923–1990. University of Texas Press: Austin, TX.

Monahan, J. (1981). Predicting violent behavior: An assessment of clinical techniques. Beverly Hills: Sage.

Morva v. Commonwealth, 278 Va. 329, 683 S.E.2d 553 (2009).

Porter v. Commonwealth, 276 Va. 203, 661 S.E.2d 415 (2008).

Reidy, T. J., Cunningham, M. D., & Sorensen, J. R. (2001). From death to life: Prison behavior of former death row inmates in Indiana. *Criminal Justice and Behavior*, *28*, 67–82.

Sandys, M., Pruss, H. C., & Walsh, S. M. (2009). Aggravation and mitigation: Findings and implications. *Journal of Psychiatry and Law*, *37*, 189–236.

DOI: 10.1002/bsl

Shah, S. (1978). Dangerousness: A paradigm for exploring some issues in law and psychology. *American Psychologist*, *33*, 224–238. doi: 00030066X/78/3303-0224
Shapiro, M. (2009). An overdose of dangerousness: How "future dangerousness" catches the least culpable capital defendants and undermines the rationale for the executions it supports. *American Journal of Criminal Law*, *35*, 101–156.
Skipper v. South Carolina, 461 U.S. 1 (1986).
Slobogin, C. (2009). Capital punishment and dangerousness. In R. Schopp (Ed.), Mental disorder and criminal law: Responsibility, punishment, and competence. New York: Springer Press.
Sorensen, J. R., & Cunningham, M. D. (2009). Once a killer, always a killer? Prison misconduct of former death sentenced inmates in Arizona, *Journal of Psychiatry & Law*, *37*, 237–267.
Sorensen, J. R., Cunningham, M. D., Vigen, M. P., & Woods, R. O. (2011). Serious assaults on prison staff: A descriptive analysis. *Journal of Criminal Justice*, *39*, 143–150. doi:10.1016/j.jcrimjus.2011.01.002
State v. Douglas, 310 Or 438, 450, 800 P2d 288 (1990).
United States v. Cooper, 91 F. Supp. 2d 90, 112 (D. D. C. 2000).
United States v. Gilbert, 120 F. Supp. 2d 147, 154 (D. Mass. 2000).
United States v. Glover, 43 F. Supp. 2d 1217, 1227 n. 6 (D. Kan. 1999).
United States v. Llera Plaza, 179 F. Supp. 2d 464, 487–88 (E. D. Pa. 2001).
United States v. Peoples, 74 F. Supp. 2d 930, 931 (W.D.Mo. 1999).
United States v. Rodriguez, No. CRIM. 2:04-CR-55, 2006 WL 487117, at _ 5 (D. N. D. 2006).
United States v. Sablan, 555 F. Supp. 2d 1177, 1185–86 (D. Colo. 2007).
Yang, M., Wong, S. C. P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, *136*, 740–767. doi: 10.1037/a0020473.

DOI: 10.1002/bsl

POLITICS

A Deadly Question

Have juries sentenced hundreds of people to death by trying to predict the unpredictable?

ABBIE VANSICKLE NOV 19, 2016

An undated photo of the execution bed at the Texas State Penitentiary at Huntsville (STR NEW / REUTERS)

Bill Meier says he can't remember exactly how he arrived at the deadly question, back in 1973.

"I frankly don't have the kind of memory that would allow me to remember just what was said," he said.

Meier is a Texas lawyer whose deep drawl makes it easy to imagine him wooing a jury. He now sits as a judge on the Texas Second Court

of Appeals, but back in the 1970s, he served as a state senator. He's a colorful character who's made *Texas Monthly* magazine's top 10 legislators list both as one of the state's best and as one of its worst. In 1973, the magazine called him open-minded, highly accessible, and never dogmatic. It later panned him as a legislator who "masqueraded as one who would advance the cause of conservatism; in fact, his cause was himself." Meier holds the record for the nation's longest filibuster after talking for 43 hours, wearing house slippers to ease his aching feet and an "astronaut bag" to prevent bathroom breaks to try and stop a bill that he believed would erode public records laws.

Back in 1973, Texas legislators were wrestling with how fix one of the state's most infamous institutions—its death penalty. The previous year, the U.S. Supreme Court had ruled the death penalty unconstitutional, finding it was too arbitrary, too riddled with discrimination and racism. The case put a stop to executions nationwide and sent state legislatures scrambling to write new laws to fix these flaws. Most states passed laws that allowed juries to consider past behavior and crimes, but Texas focused on predicting the future. Before jurors could sentence someone to death, they must first decide if the person will be a future danger.

The precise wording of the question is convoluted, asking jurors "whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society." At its core, it contains an incredible idea: Can we predict whether or not a killer will kill again?

Meier was one of the key legislators who crafted the law. In the years since, he's gotten "probably a 100 calls about this" from people

asking how he and his colleagues thought it up, he said. But he said he can't remember the discussion that led to it.

"Imagine trying to remember what you said and did back then," he said with a chuckle.

Texas holds a unique place in the annals of the American death penalty. It is, by far, the country's most prolific executioner. Since 1976, when the Supreme Court reinstated the death penalty in a case called *Gregg v. Georgia*, Texas has executed 538 people. Oklahoma is in second place with a relatively paltry 112 executions. And, more than 30 years later, Meier's question remains at the heart of this deadly system.

The Supreme Court recently heard arguments in a Texas death penalty case called *Buck v. Davis*. The defendant, Duane Buck, fatally shot his ex-girlfriend and her friend in July 1995. Two years later, a jury found him guilty and sentenced him to death. The case arrived at the Supreme Court because of a bizarre twist—at the sentencing, Buck's own lawyer introduced testimony by an expert who said that Buck's race made him more likely to be a future danger. Buck is black.

Should someone be sentenced to death, in part, because of his race? Should expert witnesses use race to predict whether someone will be violent? Those questions sound shocking, and rightfully so. But the circumstances of Buck's case—where an expert testified explicitly that race should be a factor for the jury to consider in sentencing—make it an outlier. The framing of the question before the court in Buck's case is narrow, and doesn't address whether future dangerousness is the right question to ask.

What if hundreds of people have been sentenced to death using a question aimed at predicting the unpredictable?

It's a tempting notion that we can predict who will live peaceful, productive lives and who will erupt in violence. If we could, it would certainly take some of the arbitrariness out of the death penalty. But how good are we at making those predictions?

The modern age of the death penalty began in 1972. That's when the Supreme Court heard a Georgia murder case called *Furman v. Georgia*. The defendant, William Henry Furman, had broken into a home in Savannah and was rummaging around when the homeowner awoke. Furman ran off, but on his way out, he dropped his loaded gun, which fired, killing the homeowner. A jury convicted Furman, who was black, in a one-day trial. In its decision, the Supreme Court held the death penalty, as then applied, was unconstitutional, that it was too haphazardly applied and violated the Eighth Amendment's prohibition against cruel and unusual punishment.

In his concurring opinion, Justice Thurgood Marshall wrote, "Capital punishment is imposed discriminatorily against certain identifiable classes of people; there is evidence that innocent people have been executed before their innocence can be proved; and the death penalty wreaks havoc with our entire criminal justice system."

The decision created a de facto moratorium on the death penalty nationwide as states tried to craft laws that would satisfy the Eighth Amendment. In Texas, Meier and a few other legislators hurried to draft such a bill before the end of the legislative session.

Meier sponsored a Senate bill to revamp the death penalty. A similar

bill moved through the Texas House of Representatives called House Bill 200. That bill passed both houses and then was reviewed by what's called a conference committee, a small group of legislators from both houses. Meier told me that's where he and another legislator, Terry Doyle, came up with the deadly question, known as the "future dangerousness" question.

At the time, Craig A. Washington Sr., was a state representative. Although he was opposed to the death penalty, he attended the conference committee discussions about the bill because he wanted to hear the debate. But, unlike Meier, Washington told me that he remembers the discussion around future dangerousness, or, rather, the lack of discussion.

"It was made up out of thin air," said Washington Sr., a Houston lawyer. "Thin air."

Washington said that the question seemed to be a way for the legislature to create some standard that seemed scientific, while still giving jurors a way to come to arbitrary decisions about life or death.

"If we could figure out who will be a danger, then we wouldn't have crime," Washington told me. He scoffed at the idea that even experts could make that prediction. "Experts," he said."I call them a guess in an evening gown, a g-u-e-s-s."

No one has embodied the troubling questions embedded in the "future dangerousness" question quite like psychiatrist Dr. James Grigson, who earned the sobriquet Dr. Death because of his willingness to testify in favor of executing the convicted.

After Texas passed its revamped death penalty law with the "future

dangerousness" question, prosecutors used expert witnesses to testify about a defendant's risk for violence. Grigson testified in 167 capital cases. Known for his genial, folksy mannerisms, Grigson frequently claimed certainty in predicting a defendant's risk for future violence, even if he'd never talked to the person. In 1983, *Barefoot v. Estelle*, a case involving Grigson's predictive powers made it to the U.S. Supreme Court, which ruled that experts could testify about a person's future dangerousness based on hypothetical questions, even if the expert had never talked with the person.

Grigson's confident predictions led to fame—*Vanity Fair* chronicled his colorful career—but also made him a pariah to many in his profession. The American Psychiatric Association expelled him from the professional organization in 1995 and sharply criticized Grigson's methods as junk science. Nowadays, Texas juries rarely hear from such experts. Grigson passed away in 2004, and that type of expert witness, one who relies on hypotheticals, has fallen out of favor.

But the "future dangerousness" question remains.

Dr. Mark Cunningham, a Seattle-based psychologist, and Dr. John Edens, a psychologist at Texas A&M University, have devoted their professional lives to the question of whether we can predict the future dangerousness of those convicted of crimes. Both have published extensively on the topic. And both have reached much the same conclusion.

"Juries show absolutely no predictive ability whatsoever," Cunningham said. "And, in fact, experts are similar."

The American Psychiatric Association—the same one that expelled Grigson—has taken a similar position and implored the Supreme

Court to ban the future dangerousness question in capital cases, saying in an amicus brief that "[t]he unreliability of psychiatric predictions of long-term future dangerousness is by now an established fact within the profession." The APA concluded that the "future dangerousness" question relies on junk science, and found that experts are wrong in two out of three predictions of "future dangerousness."

Part of the problem, Cunningham explained, is the sample size. Death penalty defendants represent a small sample of people, all of whom have been convicted of a violent act, so there's some risk for violence for all of them. Trying to predict who will be violent again doesn't work.

The other part of the problem is the actual circumstance versus the jury's fears, he said. In Texas, the alternative to the death penalty is life without parole, so the defendant will spend the rest of his life in prison. But jurors can make their decision as if the person would be released back into society.

"It renders this issue much more ambiguous and hard to define and maybe kind of fantastical," Cunningham said.

Edens agreed. Unlike in Texas, in the federal system and in many other states, a person's likelihood of future violence is only one of a number of factors that a jury can take into consideration. But the person's life doesn't hinge on that single question. And the jury isn't allowed to talk about hypotheticals, they must focus on the actual circumstances that await a convicted defendant—a life sentence in prison or execution.

"People are supposed to be constrained to what's really going to

happen," Edens said.

Cunningham and Edens are both skeptical about our ability to predict future danger in society, and they're even more skeptical about it once someone is confined to prison. "If your goal is 'let's put somebody to death before they do something bad in the future,' then the good news is that our prisons work relatively well at suppressing violence, but trying to pick those needles out of a haystack, we're not good at doing that," Edens said.

Their view—that it's very difficult to predict who will be a future danger, especially in prison—isn't universally shared. One man who is sharply critical of their findings is Dr. Richard Coons, a psychiatrist based in Austin, Texas, who has testified in dozens of capital cases. Coons himself has faced criticism for his evaluation methods. In 2010, the Texas Court of Criminal Appeals ruled that his methods for predicting future dangerousness were not reliable, but Coons stands by his techniques. When we spoke, he claimed that Cunningham underestimated the likelihood that a defendant would be dangerous in prison.

"[Cunningham and his researchers] talk about how little violence there is in the criminal justice system. Well, hogwash—that's a good old Texas term," he said.

Coons, who is now retired, told me that there are a number of factors that he'd use to predict a person's future danger. He'd look at their history of violence, their attitude toward violence, the facts of the case itself, whether or not the person has a conscience. Then, he'd make his prediction. Unlike the expert in the Buck case, Coons said, he'd never use race as a predictor, although he told me that it was

clear to him why blacks were overrepresented in jails and prisons.

"It's in the papers and the magazines about how prevalent black-on-black violence is," he told me. "It's not because they are black; it's because they're raised funny."

Major studies have concluded that both the race of the defendant and the race of the victim play a pivotal role in whether a person is sentenced to death. A study by criminologists that was included in Buck's appeal examined racial disparity in Harris County, Texas, the county where Buck was sentenced to death. The study found that, from 1992 to 1999, Harris County prosecutors were three and a half times more likely to seek the death penalty against black defendants than white ones. Jurors were more than twice as likely to sentence blacks to death.

Another study of capital cases in Harris County found that juries were more likely to impose a death sentence if the victim was white than if the victim was black, a pattern found in studies of other jurisdictions. Another study led by a Stanford University psychologist found that, in cases that involved a white victim, the more stereotypically black a defendant was perceived to be, the more likely the defendant was to be sentenced to death.

As philosophically at-odds as researchers like Cunningham and Edens are with Coons, surprisingly, they all agreed on one thing: We're asking jurors the wrong question. Cunningham and Edens, who often testify for the defense, have been open about their belief that the question is flawed, an open invitation to the jury to make a moral judgment on the defendant. But when I asked Coons directly if we're asking the right question, he came to the same conclusion:

"Maybe not." He added, "There is no credible way of evaluating the situation to a high degree of likelihood."

Coons, who peppered our conversation with colorful anecdotes and stories, used this analogy when I asked him to explain what he meant about the question's predictive power. "That's kind of like saying, "Which of these hummingbirds are going to go to Panama for the winter and which are going to Costa Rica?'" he said. "You can't know."

So if experts don't believe we can accurately predict whether someone convicted of a capital crime will be a future danger, is the deadly question on its deathbed?

The answer is likely no.

The reasons, according to mental health experts and lawyers, are two-fold. First, the question is written into the law, and Texas juries have answered it in every death penalty case since 1976. The U.S. Supreme Court upheld a direct challenge to future dangerousness in a 1976 case called *Jurek v. Texas*, finding that even if an expert is wrong about a person's dangerousness, the adversarial process of a trial will sort it out. The court added that judges routinely make predictions about people's future behavior in other contexts, like setting bail or in hospitalizing someone for a mental illness. Experts told me that it's unlikely the Texas Legislature would change the law. It might be viewed as an admission the question was flawed, which could result in all 244 people on death row challenging their sentences.

"Once a public policy is in place, it's almost like it takes on a life of its own. It becomes a cherished notion," Cunningham said.

The second reason is that the way courts deal with the question has morphed. It used to be that prosecutors would bring Grigson, Coons, or some other expert, and the defense would counter with their own expert. Now, though, those types of experts have largely fallen out of favor, said Kase, the defense lawyer. Instead, more and more, it's left to juries. Prosecutors often present a list of bad acts by defendants—prior convictions and disciplinary problems in prison—and the defense offers mitigating evidence, such as evidence that their client is adjusting well to prison.

Experts and lawyers say that trend makes "future dangerousness" even less of a science and more of a guess or moral judgment. It's certainly not what Meier envisioned when he wrote the question all those years ago. "The question is not something that's subject to being answered by a non-expert because it's predicting future conduct or the chance or danger or likelihood of future action, and that's something that requires an expert," Meier told me.

And, for defense lawyers like Kase, it's seems proof that the Texas death penalty is not about science, but about something much more abstract, a troubling moral judgment about who is too evil, too dangerous to live.

"It is akin to giving jurors two cotton swabs, asking them to look at them and saying, 'Does the DNA match?'" she said. "If an expert can't figure it out, then how can jurors do that? It is no accident that African Americans are overrepresented on death row."

We want to hear what you think about this article. Submit a letter to the editor or write to letters@theatlantic.com.