Mark D. Cunningham and Thomas J. Reidy
*Criminal Justice and Behavior* 2002; 29; 512
DOI: 10.1177/009385402236731

The online version of this article can be found at:
http://cjb.sagepub.com/cgi/content/abstract/29/5/512
VIOLENCE RISK ASSESSMENT AT FEDERAL CAPITAL SENTENCING

Individualization, Generalization, Relevance, and Scientific Standards

MARK D. CUNNINGHAM
THOMAS J. REIDY

The application of group statistical data to violence risk assessment enjoys strong empirical support. Accordingly, this method has dominated expert testimony regarding future dangerousness at federal capital sentencing trials across the past 5 years. Standards for admissibility of violence risk assessment testimony in federal capital sentencing are being considered but remain ambiguous. Challenges to violence risk assessment testimony in these cases have broadly centered on issues of relevancy and reliability. Corollary questions include when such assessments are sufficiently individualized, whether group data can be generalized across American correctional settings, what scientific evidence supports a given methodology, and how information regarding special conditions of confinement is relevant to risk. Conceptual perspectives and scientific evidence regarding each of these issues are discussed.

Group statistical methodology and data have been identified as fundamental to reliable violence risk assessments at capital sentencing (Cunningham & Reidy, 1998b, 1999; 2001; Reidy, Cunningham, & Sorensen, 2001; Sorensen & Pilgrim, 2000). Over the past 5 years, this methodology has been routinely presented at federal death penalty trials. Arguably, this application of group statistical methodology and data at federal capital sentencing moves these grav-
est of determinations toward the “greater degree of reliability” (p. 989) in death penalty litigation called for in Lockett v. Ohio (1978).

The use of group statistical information in violence risk assessments enjoys general scientific acceptance—both as a broad methodology and as data that substantially increase the reliability of these estimates (Cunningham & Reidy, 1998b, 1999, 2001; Gottfredson, 1987; Hall, 1987; Monahan, 1981, 1996; Morris & Miller, 1985; Poythress, 1992; Reidy et al., 2001; Serin & Amos, 1995; Showalter & Bonnie, 1984; Smith, 1993). Even those who would temper the enthusiasm for actuarial approaches to violence risk assessment because of methodological limitations in the underlying research (Litwack, 2001; Melton, Petrila, Poythress, & Slobogin, 1997) acknowledge the important role that broad perspectives regarding base rates of violence, as well as empirical correlates of violence risk derived from group statistical data, have in reliably performing these assessments.

Group statistical data in violence risk assessment are typically described in terms of the base rate, or frequency of violence, in a given sample or population under specified circumstances over a given period of time. Knowledge of the applicable base rate has been described as “the most important single piece of information” in making accurate violence risk assessments (Monahan, 1981, p. 60).

The fundamental reliance of empirically supported violence risk assessment models on base rates is not restricted to capital sentencing. Group statistical violence risk assessment applications also include noncapital sentencing determinations, prison classification, parole eligibility, and civil commitment and release. Furthermore, the use of group statistical techniques in evaluating risks of all types has had a longstanding role as the principal methodology in the insurance industry, where group casualty experience is used to set individual premiums. Group statistical data also form the fundamental scientific underpinnings of the practice of medicine and pharmacology. For example, the prescription of an antibiotic for a given infection relies on the clinical trials with this drug (i.e., the statistical response to this medication by a group of patients with a similar infection). Group statistical methodology then has demonstrated reliability and validity in a broad arena of commercial and clinical applications—and enjoys equally broad scientific acceptance.
A RECENT HISTORY OF VIOLENCE RISK ASSESSMENT
TESTIMONY AT FEDERAL CAPITAL SENTENCING

Mental health experts testifying at federal capital sentencing since 1994 have routinely employed group statistical methods to assess the likelihood of serious violence within a prison context over a capital life term if a defendant was sentenced to life without parole. According to the Federal Death Penalty Resource Counsel Project (Defender Services Division of the Administrative Office of the United States Courts), since 1994 50 federal death penalty trials, involving 68 defendants, have resulted in guilty verdicts at trial and proceeded to the penalty phase (K. McNally, personal communication, December 3, 2001). In 18 of these penalty phases, mental health experts have provided violence risk assessments. Group statistical methodology and data from psychologists or sociologists were heard in all 18 of the federal death penalty trials in which mental health experts testified regarding violence risk assessment. In most instances of this testimony, group statistical data regarding various correctional populations were utilized as risk assessment anchoring points regarding the likelihood that a federal capital defendant would commit acts of serious violence while serving a life-without-parole sentence in federal prison. Often this testimony was accompanied by conservative particularization to a given defendant based on individual demographic, dispositional, and past correctional response features. In all 18 of the cases hearing expert testimony regarding group statistical methodology and data, the associated mental health expert was called by the defense.

In 13 of these cases, the group violence base rates were conservatively particularized to the defendant based on interview, records review, and third-party interviews. In 5 of these federal capital cases, the defendants declined to be evaluated, and a psychologist subsequently testified as a teaching witness regarding violence risk assessment methodology, applicable base rates of violence in prison, and prison risk management options available in the federal Bureau of Prisons (BOP) including super-maximum confinement.

Mental health experts called by the government testified regarding violence risk assessment in 6 of the 18 cases, most often in rebuttal.
Significantly, in this rebuttal testimony neither the methodology of reliance on group statistical data nor the statistics offered were disputed. Rather, the risk assessment based on group statistical data was ignored, or some characteristic or characteristics of the defendant were asserted as more predictive of future violence. In three of the six instances of testimony by government-retained experts, psychologists based their violence risk assessments primarily on the Psychopathy Checklist–Revised (PCL-R) (Hare, 1991), with secondary reliance on clinical appraisal, personality assessment, and/or offense analysis. Two other federal death penalty cases heard violence risk assessment testimony from psychiatrists called by the government based on clinical appraisal of the defendant’s community misconduct history and implications of the instant offense to future violence in prison. In one case, a psychologist testifying for the government in rebuttal concluded that the defendant’s community violence history as well as allegations of soliciting violence in the community from prison superceded the group statistical data.

STANDARDS FOR ADMISSION OF VIOLENCE RISK ASSESSMENT TESTIMONY BY MENTAL HEALTH EXPERTS

Although violence risk assessment testimony based on group statistical methodology and data has been routinely admitted in federal capital sentencing, it has faced repeated challenges regarding admissibility. Response to these challenges is complicated by ambiguity in the standard for admitting expert testimony at federal capital sentencing. Both 21 United States Code (U.S.C.) Section 848(j) and the Federal Death Penalty Act (FDPA) 18 U.S.C. Section 3593(c) refer to the parties presenting “information” relevant to the aggravating and mitigating factors, and specifically state that

information is admissible regardless of its admissibility under the rules governing admission of evidence at criminal trials except that information may be excluded if its probative value is outweighed by the danger of creating unfair prejudice, confusing the issues, or misleading the jury.
The standards for violence risk assessment testimony at capital sentencing by mental health experts established by *Barefoot v. Estelle* (1983) were arguably low and provided no guidance for the Court in discriminating reliable, scientifically based testimony from unreliable speculation. Furthermore, statute and case law parameters for admissibility of scientific evidence at capital sentencing remain undecided (see *U.S. v. Barnette*, 2000). It could be argued under the Fifth and Eighth Amendments that the singular gravity and irreversible consequences of imposing a death penalty call for evidence that is at least as reliable and valid as the evidence admitted in civil litigation. In that civil context, Federal Rule of Evidence 702 governs admissibility of scientific evidence that qualified experts can provide:

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 by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise.

Rule 702 was interpreted by the U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993) (for a broad discussion of the implications of *Daubert* for mental health expert testimony, see Gatowski et al., 2001; Goodman-Delahunty, 1997). Briefly, in *Daubert* the Supreme Court held that proposed expert testimony must be demonstrated, by a preponderance of the evidence, to be scientifically grounded as well as relevant and reliable to the issues presented in the case. In subsequent elaboration (*Kumho Tire Co. v. Carmichael*, 1999), two separate determinations were specified: (a) Does the expert’s testimony consist of scientific, technical or other specialized knowledge; and (b) will the application of that knowledge to the particular facts and circumstances of the case aid the jury in understanding or deciding a fact that is in issue.

This standard placed more rigorous requirements on the scientific foundation for professional opinions than was the case under *Frye v. United States* (1923) that used a “general acceptance test” as the standard for admission of scientific evidence. Instead, *Daubert* held that the expert’s testimony must be based upon scientifically valid reasoning or methodology. The trial court is thus responsible for examining the body of science underlying the expert’s opinion and determining
whether or not that body of science is sufficiently developed and
cohort to justify permitting the expert to draw conclusions from it
and present them in a court of law. Four factors were identified in
Daubert that could provide the trial court guidance in making its
determination of the validity of the expert testimony: (a) whether the
knowledge or technique can be or has been tested by scientific meth-
ods; (b) whether the theory has been subjected to peer review and pub-
lished in scientific journals; (c) the rate of error associated with the
particular technique, and the existence and maintenance of standards
controlling the technique’s operation; and (d) the degree to which the
expert’s theory is recognized and accepted as valid within the relevant
scientific community.

After determining the scientific reliability and validity of the pro-
posed expert testimony, the trial court makes an additional determina-
tion regarding whether the testimony can be validly applied to the par-
ticular facts and issues in the case. Finally, the trial court must weigh
the costs of error if the testimony is excluded erroneously (Faigman,
1995).

ADMISSIBILITY OF VIOLENCE
RISK ASSESSMENT TESTIMONY

Four overlapping issues, stemming from the above-described stat-
utes and case law, currently frame the discussion of admissibility of
violence risk assessment testimony in federal capital cases:

1. Are some methods of violence risk assessment more individualized
and, therefore, more relevant to a particular defendant than others?

The distinction between individualized as opposed to group meth-
ods is a false dichotomy. Such a contrast reflects a fundamental mis-
derstanding of the nature of risk assessment and, more broadly, of
psychology as a science. Simply stated, there is no individualized
assessment of a particular person that does not rest on group data of
one sort or another. Psychology exists as a science because it provides
a database regarding the behavior of groups of individuals that is sys-

entifically derived expert knowledge in psychology (as distinct from personal experiential) consists of published observations and research on various groups of animals and individuals. Scientific expertise in psychology then consists solely of knowledge about this group-derived data. Thus, any scientifically informed evaluation, treatment, or behavioral prediction regarding a specific person relies on these collective observations and research data from a specified context.

Violence risk assessment, a particular area of research and knowledge within psychology, also fundamentally relies on the accumulation of group data that are then applied to a given individual. At the point that group data are applied to a particular person, the methodology becomes individualized. Regardless of whether the violence risk assessment method is clinical (interview and/or testing), anamnestic (past behavioral pattern), or explicitly group statistical, conclusions offered by an expert about a particular individual are purported to be reasonably and reliably inferred from group data.

How does interviewing rely on group data? Interviewing generates behavioral observations and history that have a demonstrated association with certain personality characteristics and behavior proclivities in groups of individuals studied. Any assertion of violence probabilities associated with these characteristics would necessarily rely on the group of persons displaying similar characteristics having had an increased frequency (base rate) of violence in a similar context.

How does psychological testing rely on group data? Psychological testing provides scores that allow a systematic comparison of the individual to a reference group. The scores and scale elevation patterns of the individual are interpreted in terms of performance and characteristics of a norm group or group-derived criterion.

How do violence risk assessment instruments rely on group data? Structured techniques for the appraisal of violence risk such as the PCL-R (Hare, 1991), Violence Risk Appraisal Guide (VRAG) (Quincey, Harris, Rice, & Cormier, 1998), or the Historial, Clinical, Risk Management-20 (HCR-20) (Webster, Douglas, Eaves, & Hart, 1997) each propose that various demographic, historical, personality
features, and/or diagnostic variables are predictive of future violence. Scoring based on the presence or absence of specified characteristics in a given individual is interpreted in terms of the group experience (i.e., the rate of violence among the standardization/experimental sample that were scored on the same set of characteristics).

_How does behavior pattern analysis rely on group data?_ The risk assessment implications of an individual’s past pattern of behavior rely on an assumption that people as a group maintain similar behaviors over time. It further assumes that the groups of individuals who have had a pattern of violent and/or criminal activity in the past have a much higher continuing incidence (base rate) of those misdeeds.

_How can a violence risk assessment rely on group data and still be individualized?_ All risk assessment methods are both individualized in terms of placing the individual in a reference group and group statistical in terms of using the group experience (base rate) to give meaning to the individual prediction. The various risk assessment methods simply group individuals in different ways: personality characteristics, diagnosis, historical variables, test scores, behavior patterns, and/or incarceration status. The question then is not whether individualized risk assessment will be based on group data—whether applied by an uninformed jury or expert testimony. Instead, the issue is whether the grouping provides empirically sound group data regarding the likelihood of prison violence.

_How are overt group statistical methods of violence risk assessment different?_ The difference between overt group statistical methods and interview/testing/risk assessment instruments is principally that in using a formal statistical approach the underlying group base rates of violence are explicitly stated rather than inferred. By allowing the capital jury the benefit of being able to directly scrutinize the group incidence (base rate) of violence, as well as its scientific basis, a formal group statistical methodology has the least likelihood of misleading or confusing the jury. This is contrasted with primary reliance in capital sentencing risk assessments on interview, personality testing, and/or risk assessment instruments. These give an appearance of
being more individualized while, in fact, simply obscuring the fit of the group and potentially misrepresenting and exaggerating that group’s actual or true rate of violence in prison.

Additional conservative particularization to a given capital defendant relative to the group base rates can be made based on various demographic factors such as age, the defendant’s past response to structure and confinement, the presence of untreated psychosis or other major mental illness, offense characteristics, and other variables. Whether more specific particularizing is made in expert testimony, the group base rates provide critically important statistical anchoring points for a federal capital sentencing jury’s risk assessment deliberations.

2. Can group statistical data regarding institutional misconduct and violence from state prison populations or from noncapital inmates in the federal BOP be reliably generalized to federal capital defendants?

In any application of group statistical data, there are ways in which the particular instance or individual person differs somewhat from the general research or respective group. The critical issue is not the presence of unique features in the instant appraisal but rather whether there is sufficient commonality in the characteristic of interest for the general research to reliably generalize to the specific case. All applications of scientific data involve this generalizing from broader research to the specific case. In the same way, the group base rate of violence can be a valid and reliable estimate of the individual risk, even when the match is not perfect. Determining the fit between prison correctional data and capital inmates in federal custody requires scrutiny of that data, detailed in the discussion that follows.

Applying violence rates of former death row inmates and incarcerated murderers. The studies in Table 1, representing national and state samples, track the frequency of serious violence in the general prison population by inmates convicted of murder, including former death row inmates (see also Cunningham & Reidy, 1998b; Reidy et al., 2001). A number of rationales support generalizing from these studies to capital offenders in other correctional settings, including the federal Bureau of Prisons:
### TABLE 1: Assaultive Rule Violations of Former Death Row Inmates and Comparison Inmates

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Follow-Up Interval</th>
<th>Assault Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Former Death Row Inmates</td>
<td>Comparison Inmates</td>
</tr>
<tr>
<td>Marquart, Ekland-Olson, &amp; Sorensen, 1994</td>
<td>N = 100, FDR, Texas</td>
<td>1924-1972 (avg. 12 yrs)</td>
<td>.20 cumulative</td>
</tr>
<tr>
<td>Marquart et al., 1994</td>
<td>N = 47, FDR, Texas</td>
<td>1973-1988 (avg. 10 yrs)</td>
<td>.07 cumulative</td>
</tr>
<tr>
<td></td>
<td>N = 156, LS, Texas</td>
<td>1973-1988 (avg. 11 yrs)</td>
<td>.10 cumulative</td>
</tr>
<tr>
<td></td>
<td>(128 murderers/28 rapists)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marquart, Ekland-Olson, &amp; Sorensen, 1989</td>
<td>N = 90, FDR, Texas</td>
<td>1974-1988 (avg. 6.3 yrs)</td>
<td>.016 annual</td>
</tr>
<tr>
<td></td>
<td>N = 107 CLS murderers, Texas</td>
<td>1974-1988 (avg. 7.2 yrs)</td>
<td>.026 annual</td>
</tr>
<tr>
<td></td>
<td>N = 38,246, TDC systemwide</td>
<td>1986</td>
<td>.12 annual</td>
</tr>
<tr>
<td></td>
<td>N = 1,712, TDC, high security unit</td>
<td>1986</td>
<td>.20 annual</td>
</tr>
<tr>
<td>Marquart &amp; Sorensen, 1989</td>
<td>N = 533, nationwide</td>
<td>1972-1987 (15 yrs)</td>
<td>.31 cumulative</td>
</tr>
<tr>
<td></td>
<td>(453 murderers/80 rapists)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akman, 1966</td>
<td>N = 67, FDR, Canada</td>
<td>1964-1965 (2 yrs)</td>
<td>0 cumulative</td>
</tr>
<tr>
<td>Bedau, 1964</td>
<td>N = 7,285, systemwide, Canada</td>
<td>1964-1965 (2 yrs)</td>
<td>.007 annual</td>
</tr>
<tr>
<td>Sorensen &amp; Wrinkle, 1996</td>
<td>N = 55, New Jersey</td>
<td>1907-1960 (53 yrs)</td>
<td>0 cumulative</td>
</tr>
<tr>
<td></td>
<td>N = 648 murderers, Missouri</td>
<td>1977-1992</td>
<td>.218 cumulative</td>
</tr>
<tr>
<td></td>
<td>(93 death row/323 LWOP/232 LWP)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sorensen &amp; Pilgrim, 2000</td>
<td>N = 3,390 murderers, Texas</td>
<td>1990-1999 (avg. 4.5 yrs)</td>
<td>.024 annual</td>
</tr>
</tbody>
</table>

**SOURCE:** Adapted from Reidy, Cunningham, & Sorensen (2001).

FDR = former death row inmates; LS = life sentence; CLS = capital life sentence; TDC = Texas Department of Corrections; LWOP = life without parole; LWP = life with parole.
1. The underlying capital offenses in past samples and current offenders were sufficiently aggravated (in comparison to other capital eligible offenses) that a death sentence was sought and returned.

2. In many cases, there is overlapping state and federal jurisdiction in cases that are ultimately prosecuted as capital cases in federal court. There is little plausible reason to believe that defendants committing capital murder in these or other federal cases are different in some essential way from murderers tried in state courts.

3. Fundamental aspects of state and federal prison incarceration have a high degree of similarity over time and across corrections departments—particularly during the past 30 years. Correction departments belong to common associations and accrediting organizations, disseminate statistical experience and research findings with each other, and are guided by common judicial determinations.

4. Despite the diversity of settings and time periods, there is broad similarity in the research results from study to study—reflecting a robust finding.

5. There is no research or other authority demonstrating that these base rate studies cannot be reliably generalized to federal capital defendants.

To summarize, these studies produce highly consistent data regarding the violence rate of these offenders regardless of the victim or weapon features of the murder, specific death penalty statute, decade of follow-up, particular state department of corrections, or sentence of the convicted murderer (former death row, life without parole, life with parole). In other words, none of these unique-case variables influenced the outcome rate of prison violence for the broad category of murderer sentenced to at least decades of confinement. Only with the extraordinary sample size of the Sorensen and Pilgrim (2000) study of Texas murderers were a limited number of demographic/offense characteristics demonstrated to raise or lower the projected risk of prison violence. There is neither evidence nor plausible rationale that such consistent and robust findings would not generalize to federal capital murderers in BOP.

Application of other violence rates. There are a number of relevant ways to group federal capital defendants and examine the applicable rates of prison violence in those reference groups (Cunningham & Reidy, 1998b, 1999, 2001; Reidy et al., 2001)—which then provide a transparent basis for individualizing to the specific federal capital
<table>
<thead>
<tr>
<th>Study</th>
<th>Follow-Up Interval</th>
<th>Assault Type</th>
<th>General Population</th>
<th>U.S. Penitentiaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harer (1992)</td>
<td>1988-1989 (18 months)</td>
<td>Total assaults on inmates</td>
<td>.011</td>
<td>.171</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total assaults on staff</td>
<td>.0116</td>
<td>.0280</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serious assaults on inmates</td>
<td>.0022</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serious assaults on staff</td>
<td>.0002</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assault on inmate without weapon</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assault on inmate with weapon</td>
<td>.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assault on staff without weapon</td>
<td>.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assault on staff with weapon</td>
<td>.0009</td>
<td></td>
</tr>
</tbody>
</table>
These include grouping federal capital inmates as (a) capital inmates sentenced to death but subsequently granted relief and followed in a general prison population, (b) murderers never sentenced to death and serving life terms, (c) federal capital inmates sentenced to life at trial, (d) long-term sentence inmates (see Flanagan, 1980), (e) violent and/or homicide offenders, (f) inmates in the BOP as a whole (see Tables 2 and 3) or inmates in U.S. penitentiaries (USP), (g) potential special federal inmate populations, if applicable, such as security threat groups, (h) potential inmates at Administrative Maximum (ADX) Florence (“super-max”), and (i) inmates of a particular age at entrance to BOP or across a capital term.

It is acknowledged that none of these groupings represent a perfect fit with a particular federal capital defendant. Such a perfect fit would involve a sample of federal capital inmates numbering in the thousands who had been followed in federal prison for decades. This sample does not exist, nor is it expected to in the foreseeable future. The above groupings would seem to represent a valid connection in most federal capital risk assessments.

3. Are clinical interview, offense analysis, psychological testing, and/or risk assessment instruments reliable predictors of serious prison violence?

How reliable are interview data and clinical analysis in assessing prison violence risk? There are no peer-reviewed studies demonstrating that interview data regarding personality characteristics, conscience, remorse, premeditation, psychodynamic features, instrumental versus affectively driven community offenses, or organized versus
disorganized community offense types are predictive of serious violence in prison. Similarly, there are no peer-reviewed findings demonstrating that “clinical analysis” of records, offense reports, or historical data offered in hypothetical are predictive of violence in prison. These conclusions are not surprising. Reviews of research have described the abysmal performance of violence risk assessments relying on unaided clinical judgment (Bonta, Law, & Hanson, 1998; Campbell, 2000; Grove & Meehl, 1996).

In the absence of reliable empirical data regarding the accuracy of clinical methods in the prediction of prison violence, it is not possible to determine an error rate for interviewing and clinical analysis. Peer-reviewed literature extending across 20 years has identified a number of errors that plague clinical methods of violence risk assessments, typically resulting in overestimation of risk (for a review see Cunningham & Reidy, 1999).

As the absence of empirical support and pervasive concern with error rates suggest, interviewing and clinical analysis are not generally accepted in the community of violence risk assessment scientists as reliable methods of predicting prison violence. When factors validated by group statistical data are utilized (i.e., guided clinical decision making), clinical methods of risk assessment become somewhat more reliable. Such factors for guided decision making regarding the probability of prison violence are quite limited and should only conservatively modify base rate data.

It is intuitively obvious why risk assessment formulation based on interview data would fare so poorly. Many of the risk characteristics identified in an interview (early childhood behavior problems, past aggression in the community, impulsivity, poor judgment, criminal history, alcohol/drug abuse, antisocial personality disorder, etc.) are so pervasive among an incarcerated population that they cease to discriminate who will be violent (Cunningham & Reidy, 1998a, 1999, 2001; see also Meloy, 1988; Widiger & Corbitt, 1995).

Furthermore, prison is a profoundly different context than the community. It cannot be assumed that factors associated with violence in the community will be predictive in a markedly different context. Accordingly, before such factors are employed in capital risk assessments, research would need to demonstrate that these characteristics are predictive of violence in a prison as well in a community context.
There is little, if any, group research demonstrating that these characteristics result in a high probability of violence in the context of prison.

*Does personality testing have a significant role in violence risk assessment at capital sentencing?* Traditional personality testing has been almost nonexistent in risk assessments at federal capital sentencing. This is not surprising, as personality testing has little empirical support for being predictive of prison violence. Specifically, no personality testing pattern, whether Minnesota Multiphasic Personality Inventory, second edition (MMPI-2), Millon Clinical Multiaxial Inventory (MCMI), Rorschach, or other, is reliably associated in the peer-reviewed literature with higher long-term rates of prison violence (Clements, 1996; Craig, 1996; Kennedy, 1986; Shaffer, Waters, & Adams, 1994; Van Voorhis, 1994; Zager, 1988; see also Cunningham & Reidy, 1998b). When psychological testing profiles cannot be reliably associated with long-term prison violence, an unacceptably high error rate is a given.

*Are the PCL-R, VRAG, or HCR-20 reliable bases for capital violence risk assessments?* The PCL-R has appeared occasionally in federal capital violence risk assessments. To date, there have been two published peer-reviewed studies assessing the utility of the PCL-R in predicting violence among a general inmate population in an American prison setting. Neither study supported the use of the instrument to predict prison violence. Kosson, Steuerwald, Forth, and Kirkhart (1997) reported nonsignificant correlations between institutional disciplinary offenses for violent and nonviolent behavior and PCL-R Factors 1 and 2 among a sample of male BOP inmates. Similarly, Edens, Poythress, and Lilienfield (1999) found a nonsignificant correlation between PCL-R total scores and institutional physical aggression among youthful offenders during the 1st year of incarceration and only a low correlation between PCL-R Factor 2 and institutional physical aggression \( r = .24, p < .05 \). There is additionally insufficient investigation of the predictive implications of the PCL-R for ethnic minorities (Cunningham & Reidy, 1998a; Salekin, Rogers, & Sewell, 1996). Edens, Petrila, and Buffington-Vollum (2001) in their analysis of the application of the PCL-R to violence risk assessments at death
penalty sentencing concluded, “It seems impossible to reconcile the (lack of) probative value of the PCL-R with the prejudicial impact that the label of ‘psychopath’ almost certainly will have on jurors” (p. 467).

In spite of the paucity of support for utilizing the PCL-R in violence risk assessments at capital sentencing, it has been purported by government-retained psychologists to be predictive of prison violence (for a discussion of associated ethical implications see Edens, 2001). In one of these instances a psychologist retained by the government in *U.S. v. Haynes* (2000) who had relied on these instruments for his risk assessment withdrew his report after being presented (via pretrial motion) with affidavits by several recognized violence risk assessment researchers detailing the above absence of empirical support for application of the PCL-R and HCR-20 to prison violence.

Both trial and appellate federal courts have also found concern with the prejudicial impact of a *psychopath* label stemming from the PCL-R. In *U.S. v. Lee* (1999/2001), the trial court reversed a death sentence (subsequently reinstated by the 8th Circuit) after determining that it had erred in allowing the government to go into PCL-R psychopathy in the cross-examination of a defense psychologist who had restricted his testimony on direct to mitigation. The 4th Circuit reversed the death sentence in *U.S. v. Barnette* (2000) after finding that the trial court had erred in not allowing expert surrebuttal to a government-called psychologist who identified the defendant as a psychopath (for a discussion of strategic role of the PCL-R in capital cases see Edens et al., 2001).

The VRAG and HCR-20 have been less frequently encountered in federal capital cases. These risk assessment instruments suffer from the same absence of empirical support for their utility in the prediction of serious prison violence. There are no published, peer-reviewed studies demonstrating the validity of the HCR-20 in predicting institutional violence in correctional offenders in the United States. Similarly, there are scant data (Kroner & Mills, 2001) on the relationship, if any, of VRAG scores to prison violence. More specifically, the VRAG has not been validated to predict the institutional adjustment of American correctional inmates.
Does analysis of a past pattern of behavior add to violence risk assessment at capital sentencing? Behavior pattern analysis in violence risk assessment at capital sentencing can be a very reliable method for estimating risk, assuming that there is sufficient behavior to form a pattern and that the context of prediction is sufficiently similar (Cunningham & Reidy, 1998b, 1999; Morris & Miller, 1985). This latter similarity of context is particularly important to attend to if risk assessment is to be accurate. Reviews sponsored by the U.S. Department of Justice concluded that a pattern of violence in the community has not been found to be reliably predictive of violence in prison (Alexander & Austin, 1992; National Institute of Corrections, 1992). This same discontinuity between community violence and prison violence has been confirmed in samples of former death row inmates (Reidy et al., 2001) and incarcerated murderers (Sorensen & Pilgrim, 2000). Confirming the predictive significance of past violence when in the same context, however, Sorensen and Pilgrim found that a history of past prison violence among incarcerated murderers markedly increased the likelihood of prison violence. Thus, in applying group pattern data to individual pattern prediction, it is critical to use group data from a similar context (prison conduct to prison conduct).

Does anecdotal evidence of prison violence increase the reliability of federal capital violence risk assessments? Somewhat more frequently observed in federal capital cases than poorly grounded testimony by mental health experts is the buttressing of future dangerousness allegations through anecdotal reports from BOP personnel of specific acts of violence, unrelated to the defendant, that have occurred in federal correctional facilities. These are ostensibly offered to convey the reality that inmate control is not absolute. This sort of anecdotal testimony poses a grave risk, though, to reliable and scientifically grounded risk determinations. Anecdotal descriptions of instances of prison violence present the jury with a memorable image that promotes a visceral emotional reaction of fear. This emotionally charged reaction tends to override more rational analysis and encourages the discarding of factual probability information (see Shah, 1978)—injecting a significant source of error into the jury’s determination of risk.
4. What is the relevance of risk management measures including the availability of super-maximum confinement and restricted communications to capital violence risk assessment?

What does prevention have to do with assessment? Assessment of violence risk is not simply a static enterprise. It also involves consideration of what preventive measures can be undertaken that would modify or reduce the level of violence risk posed by a particular individual (Heilbrun, 1997; Heilbrun, O’Neil, Strohman, Bowman, & Philipson, 2000; Rogers, 2000; Serin & Amos, 1995). Similarly, the risk of inmate violence is a function of what risk reduction interventions are brought to bear (Cunningham & Reidy, 1999, 2001). For example, BOP classification procedures result in capital life inmates being confined in a high-security facility (i.e., USP or higher level of custody) (U.S. Department of Justice, 2000). Additional prison-based risk reduction interventions in BOP include medication or treatment for psychological disorders, application of disciplinary contingencies, rehabilitation programming, isolation from codefendants or fellow gang members, special management provisions, or modified confinement. All of these are relevant to a capital jury’s consideration of the violence likelihood of a particular defendant.

What special management interventions are available? Even more secure confinement options than a high security USP are available in BOP. Under an administrative maximum level of confinement at ADX Florence or other special security units disbursed throughout BOP, an inmate is single-celled and locked down 23 hours daily, with individual or small-group exercise and shackled movement under escort. Under such conditions of security, opportunities for serious violence toward others are greatly reduced. Although special conditions of confinement are most typically in response to serious misconduct during federal prison confinement, these may be ordered preemptively—before a federal inmate exhibits serious institutional violence. For example, Timothy McVeigh, Ramsi Yousef, and Ted Kazinski were confined at ADX on the basis of concerns unrelated to any exhibited posttrial violent conduct.
Can “directed” violence be controlled? In some federal capital sentencing proceedings, there has been evidence that the defendant had the capability and/or history of using phone privileges, visitation contacts, letters, or inmate communication to order violence or direct criminal activity from prison. In considering the future likelihood of such directed violence, it is important for the jury to be aware that there are provisions for limiting the contact that a federal inmate may have with other inmates, the free community, or even their own attorney. Extraordinary limitations on defendant communications with other inmates and community members were ordered in U.S. v. Luis Felipe (1998) (head of the Latin Kings) and subsequently upheld by the 2nd Circuit Court of Appeals. Similar restrictions were ordered in U.S. v. Ramsi Yousef (1998) (World Trade Center bomber). Because an assertion of “future dangerousness” while confined in federal prison implicitly suggests that the BOP is unable to safely contain this defendant, the capabilities of BOP to bring higher levels of security and special conditions of confinement to bear is critically important to the jury’s risk analysis.

DISCUSSION

Violence risk assessment testimony by mental health expert witnesses at federal capital sentencing trials across the past 5 years has most often relied on group statistical methodology and explicitly described group data. This application of group statistical data to federal capital sentencing enjoys a sound scientific foundation that appears consistent with the evidentiary standards envisioned by Daubert and Kumo Tire.

To illustrate, there is extensive literature supporting the application of group statistical data to violence risk assessment. Moreover, the specific application of group statistical methodology and data to capital sentencing has passed peer review in independent scholarly journals (Cunningham & Reidy, 1998b, 1999, 2001; Marquart, Ekland-Olson, & Sorensen, 1989; Marquart & Sorensen, 1989; Reidy et al., 2001; Sorensen & Pilgrim, 2000). The underlying base rate data cited in these studies have additionally passed peer review or have been derived from data gathered and analyzed by BOP, Department of Jus-
tice, Bureau of Justice Statistics, and other reliable sources. The presentation of prison violence incidence for representative groups, at various levels of violence severity, an error rate through the range of probabilities expressed (i.e., the rate of serious prison violence among murderers and former death row inmates as a broad category ranges from 15%-30% across lifetime confinement). Particularly when accompanied by demonstrative exhibits that assist the jury’s understanding and retention, testimony detailing violence risk assessment methodology and empirical data is reliable, probative, and can be presented in a clear, organized, logical fashion that neither confuses nor misleads the jury (see 21 U.S.C. Section 848(j); FDPA 18 U.S.C. Section 3593(c)).

Despite the inroads that scientifically reliable violence risk assessment methodology and data are making in federal death penalty sentencing, unreliable methods are still being utilized by mental health experts in this arena of literal life and death. Psychiatric testimony at capital federal sentencing based on clinical intuition ungrounded by empirical research on prison violence has occurred as recently as June 2000 (U.S. v. Vialva). In some other federal capital cases, psychologists have utilized the PCL-R and/or violence risk assessment instruments even though these have not been validated to predict serious prison violence.

Although such lapses represent a minority of cases, the anticipation of faulty methodology in government-retained evaluations—and particularly the associated specter of testimony labeling a capital defendant as a psychopath—has had a chilling effect on the willingness of the defense to allow capital defendants to be evaluated. Additionally, the anticipation of government-sponsored rebuttal testimony involving unrelated anecdotal descriptions of prison violence, and/or the labeling of a defendant as a psychopath (or some variation of that conceptualization) on the basis of file data and conduct associated with the capital offense, has been a factor in defense decisions not to offer risk assessment testimony—even though needed, relevant, and empirically grounded.

Assertions of high violence probability based on methods without adequate empirical support raise grave ethical concerns. Numerous principles of the American Psychological Association Ethical Standards and Code of Conduct (ESCC) (American Psychological Associ-
(ESCC: 2.04; SGFP: II., III.A., VI.A., C., VII.D.; see Standards for Educational and Psychological Testing, American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999), and to candidly acknowledge the limitations in their data or methodology (ESCC: Principle B., 3.03, 7.04; SGFP: VI.C.; VII.A., D.). Adherence to the highest standards of ethics and practice become increasingly important as the magnitude of the issue at stake and potential for harm increases (see Cunningham & Reidy, 2001).

In a significant proportion of federal capital cases, the jury heard no mental health expert testimony addressing the likelihood of a defendant committing acts of serious violence in federal prison. Multiple factors suggest the potential for fundamental error in the jury’s violence risk determinations of a capital defendant in the absence of sound violence risk assessment methodology and data. Specifically,

1. Individuals undertaking violence risk assessment are likely to commit a number of fundamental errors unless guided by reliable scientific methodology and group data, more often resulting in an overestimation of violence risk source.

2. A capital jury’s familiarity with the heinousness of the capital offense and knowledge of other aggravating factors or acts encourages an expectation of high future violence risk in prison that is not justified by research (see Shah, 1978).

3. The combination of these factors is likely to result in a capital sentencing jury that is strongly biased toward overestimation of violence risk when scientifically grounded expert testimony is excluded.

4. A federal capital jury is at best almost certain to be ignorant of reliable violence risk assessment methodology, rates of prison violence, prison population demographics and characteristics, and federal prison confinement options and security procedures. At worst, federal capital jurors come to a capital sentencing risk assessment holding intuitive assumptions regarding future dangerousness that are false. It must be emphasized that much of the research data on the violence risk assessment of prison inmates are counterintuitive.
5. In the absence of expert testimony, a capital jury has no mechanism to discount memorable but infrequent events (anecdotal evidence), understand or incorporate base rate data, appreciate the importance of context, avoid illusory correlation, maintain skepticism of clinical methods, understand the minimal implications of antisocial personality disorder or related characterizations, incorporate the effects of aging, reliably evaluate patterns of behavior, factor in preventive measures, acquire the broad information relevant to violence risk considerations, appreciate the probabilistic nature of their task, or critically evaluate the arguments of the government or the defense.

Expert testimony regarding scientifically sound violence risk assessment methodology and associated group statistical empirical data has a fundamentally important role at federal capital sentencing where future dangerousness is at issue. With the assistance of informed expert testimony, a federal capital jury is more likely to avoid fundamental error (see Faigman, 1995) and make the most reliable assessment of a capital defendant.

NOTES

1. Despite the practical use of group data in multiple contexts, the extent to which group data can be applied to risk predictions in individual cases is a philosophy-of-science issue of ongoing debate. A full discussion of the underlying theoretical issues is beyond the scope of this article.

2. Federal capital trials are bifurcated. The jury first determines guilt. If the defendant is found guilty of the capital charge, the government and the defense may present additional evidence at a sentencing phase. The jury then makes a determination of a death or a life without parole sentence.


4. Fourteen of these representing testimony of Cunningham and two of Reidy.


6. Although the Psychopathy Checklist–Revised is a measure of the personality/behavioral construct of psychopathy, a substantial proportion of forensic interest and research with this test involves its utility as a predictor of future criminal and violent offenses.

7. These rates are not applicable to inmates charged with murdering inmates or staff, or risk estimates of ordering violence in the community from prison.
REFERENCES


Frye v. U.S., 293 F. 1013 (DC Cir. 1923).


U.S. v. Felipe, 148 F.3d 101 (2d Cir. 1998).
U.S. v. Coleman Johnson, Western District of Virginia, Charlottesville Division, Cause No. 3:00CR00026 (2001).
U.S. v Daniel Lewis Lee, 2001 WL 1590491 (8th Cir. Ark.).