

SELECTIVE INCAPACITATION: SOME QUERIES ABOUT RESEARCH DESIGN AND EQUITY

ANDREW VON HIRSCH*
DON M. GOTTFREDSON**

TABLE OF CONTENTS

I. Introduction	12
II. The Accuracy of the Prediction Method	13
A. Past Success and Failure in Forecasting Criminality	13
B. The Design of the Prediction Study	16
C. The Results of the Prediction Study	20
III. Purported Effects on the Incidence of Crime	22
A. The Shinnar Projection Model: The Rise and Fall of "Collective" Incapacitation	22
B. Greenwood's Projection Model: Altering the Shinnar Formula to Fit Selective Incapacitation	24
C. Dubious Evidence on Individual Robbery Rates: Extrapolating from Data about Incarcerated Offenders	25
D. Can More Conservative Estimates Be Made?—Restricting the Estimates to Incarcerated Offenders	28
E. Problems Even with the "Conservative" Estimate: Crime-Rate Fluctuations and Replacement Effects	29
F. Problems of Confirming Incapacitative Impact	30
IV. Ethical Issues: False Positives and Undeserved Punishment	31
A. The Morality of Overprediction	32
B. The Issue of Undeserved Punishment	34
V. Resource Issues: Prison Overcrowding and the Choice of Sentencing Rationale	37
A. Need for Selectiveness	38
B. Choosing the Criteria for Selection: The Minnesota Experience	38
C. The "Political Pressures" Argument	40

* Professor, School of Criminal Justice, Rutgers University, Newark, New Jersey. A.B., Harvard University, 1956; LL.B., Harvard University, 1960.

** Dean and Professor, School of Criminal Justice, Rutgers University, Newark, New Jersey. B.A., University of California, 1951; Ph.D., Claremont Graduate School, 1959.

We would like to thank Julia M. Mueller for her assistance in the quantitative aspects of this paper, as well as for her valuable ideas and comments. We are indebted to Donald M. Barry for his thoughts on the research-design issues discussed in the Appendix. And we would like to thank Jan M. Chaiken, David F. Greenberg, Michael D. Maltz, and John Monahan for their comments on earlier drafts of this article.

VI. Concluding Observations 42
 A. The Greenwood Study 42
 B. The State of Knowledge about Prediction 44
 Appendix: Purported Effects of Selective Incapacitation on Aggregate
 Robberies 46

I

INTRODUCTION

A familiar product, advertisers believe, sells better when given new packaging.

Sentencing on grounds of potential future criminality has been a familiar product in criminal justice. Proponents of the traditional rehabilitative ethic have long held that, while curable offenders should be treated, those who are bad risks should be identified and isolated from society.¹ In the last decade, however, predictive sentencing lost some of its cachet. Researchers raised doubts about the accuracy of forecasting techniques—particularly about their tendency to “overpredict,” that is, erroneously to classify persons as expected offenders. Questions were raised as well about the ethics of gauging convicts’ punishments on the basis of their expected future choices.² In part because of these doubts, sentencing reform efforts, such as the guidelines developed by Minnesota’s sentencing commission³ have tended to emphasize the seriousness of the defendant’s criminal conduct while limiting the role of predicted future criminality in determining the sentence.

Now, however, the idea of predictive sentencing is being revived. Several RAND Corporation prediction studies⁴ have received considerable attention in our profession and one such study, by Dr. Peter Greenwood, has obtained nationwide publicity. In the Greenwood study, the idea has been strikingly repackaged: instead of being called merely the new, improved predictive sentencing, it has been renamed “selective incapacitation.”⁵ Sub-

1. For summary of these developments, see von Hirsch, *Recent Trends in American Criminal Sentencing Theory*, 42 MD. L. REV. 1, 7-10 (1983).

2. *Id.* at 11-12.

3. For an analysis of Minnesota’s guidelines, von Hirsch, see *Constructing Guidelines for Sentencing: The Critical Choices for the Minnesota Sentencing Guidelines Commission*, 5 HAMLINE L. REV. 164 (1982). The text of the guidelines is set forth in 5 HAMLINE L. REV. 395, 395-437. For an analysis of the guidelines’ impact on sentencing practice, see MINNESOTA SENTENCING GUIDELINES COMMISSION, PRELIMINARY REPORT ON THE DEVELOPMENT AND IMPACT OF THE MINNESOTA SENTENCING GUIDELINES (1982).

4. J. CHAIKEN & M. CHAIKEN, VARIETIES OF CRIMINAL BEHAVIOR (Rand Corp. R-2814-NIJ, August 1982); P. GREENWOOD & A. ABRAHAMSE, SELECTIVE INCAPACITATION (Rand Corp. R-2815-NIJ, August 1982) [hereinafter cited as P. GREENWOOD].

5. P. GREENWOOD, *supra* note 4. Greenwood has popularized the term, but it was first used in Greenberg, *The Incapacitative Effect of Imprisonment: Some Estimates*, 9 LAW & SOC’Y REV. 541 (1975).

stantial advances have been claimed in techniques for identifying repeat serious offenders. The Greenwood study also purports to demonstrate that selective incapacitation strategies may lead to significant reductions in crime without increasing the total number of offenders incarcerated,⁶ thus promising to deal simultaneously with the intractable problems of serious crime and prison overcrowding in America.

How valid are these new claims? Does selective incapacitation really offer new solutions? Or is it a shopworn notion now decked out in flashy new guise? These are the topics of our article.⁷ We will address the following issues:

1. Traditional statistical methods of predicting individual criminality were rather crude and error-prone. Have the new prediction devices, such as Greenwood's, significantly enhanced predictive accuracy?

2. Traditional prediction studies sought only to measure the likelihood of recidivism and did not address overall crime rates. The Greenwood study, for the first time, offers supposed evidence that predictive sentencing can reduce the incidence of crime. How valid are such claims, and how sound is the research on which they are based?

3. What of the ethics of selective incapacitation? Have any new answers been offered to the moral doubts surrounding the sentencing of convicted criminals on the basis of the crimes that they are expected to commit?

4. What of resource allocation issues? Traditional prediction studies did not address the question of sentencing policies' impact on prison populations. Greenwood does, and claims that a selective incapacitation strategy is capable of alleviating prison overcrowding in a way no nonpredictive strategy can. How valid is this claim? Is there convincing evidence that selective incapacitation calls for any less use of prison resources than, say, a desert-oriented strategy?

II

THE ACCURACY OF THE PREDICTION METHOD

A. *Past Success and Failure in Forecasting Criminality*

Statistical prediction of criminality, and use of prediction methods in paroling and sentencing, have a sixty-year history in this country. As early

6. P. GREENWOOD, *supra* note 4, at xix.

7. After we completed this article, our attention was called to a valuable analysis of selective incapacitation by Jacqueline Cohen. Cohen, *Incapacitation as a Strategy for Crime Control: Possibilities and Pitfalls* in 5 *CRIME AND JUSTICE: AN ANNUAL REVIEW OF RESEARCH* 1-84 (M. Tonry & N. Morris eds. 1983). The author is critical of the predictive acuity of Greenwood's forecasting instrument, and raises doubts about Greenwood's claims regarding the impact of selective incapacitation strategies on overall crime rates. *Id.*, at 31-55.

as 1923, S.B. Warner studied offender characteristics related to parole violation.⁸ Hart, responding to Warner's work, combined such items to obtain a prognostic score, and argued that the resulting device could be used both by parole boards and by the courts in determining the penalty.⁹ In the 1930's the Gluecks' prediction studies attracted worldwide notice.¹⁰ After World War II, a number of researchers developed prediction methods further and put those methods into use. The California Base Expectancy Score, a prediction of expected parole success, was used extensively as an element in paroling and classification decisions in the 1960's.¹¹ A similar scale was adopted in the early 1970's by the United States Parole Commission for use in its guideline matrix.¹²

These studies employed essentially the same method, with varying degrees of sophistication. The researcher collected various data about a sample of convicted offenders—including the subjects' crimes, criminal records, and social histories. Using bivariate or multivariate statistical techniques, the researcher then examined how these factors were correlated with subsequent recidivism. A predictive index was then constructed based on some or all of the factors found associated with recidivism. Finally, the predictive index was tested on at least one validation sample (a sample not studied previously) to see how well it did indeed predict recidivism. Whatever the method of developing the index, the last step was considered the most important in determining how well the index works.¹³

These studies tended to show that a few familiar items—such as the offender's criminal history, age, employment and drug history—could be combined to identify subgroups of offenders having a higher probability of returning to crime than convicted offenders generally.¹⁴ Examination of these prediction studies showed, however, that the association between the items identified as predictors and subsequent offender behavior was scarcely a strong one. Where multivariate statistics were used, the percentage of the variability "explained" by all predictive factors combined tended to be

8. Warner, *Factors Determining Parole from the Massachusetts Reformatory*, 14 J. CRIM. L. & CRIMINOLOGY 172 (1923).

9. Hart, *Predicting Parole Success*, 14 J. CRIM. L. & CRIMINOLOGY 405 (1923).

10. S. GLUECK AND E. GLUECK, *500 CRIMINAL CAREERS* (1930) was one of a series of studies summarized in H. MANNHEIM & L. WILKINS, *PREDICTION METHODS IN RELATION TO BOSTAL TRAINING* (1955).

11. See, e.g., Gottfredson, *The Practical Application of Research*, 5 CANADIAN J. OF CRIMINOLOGY & CORRECTIONS 212 (1963).

12. D. GOTTFREDSON, L. WILKINS AND P. HOFFMAN, *GUIDELINES FOR PAROLE AND SENTENCING: A POLICY CONTROL METHOD* 1-80 (1978).

13. For a discussion of prediction techniques, see generally H. MANNHEIM & L. WILKINS, *supra* note 10; F. SIMON, *PREDICTION METHODS IN CRIMINOLOGY* (1971); D. Gottfredson, *Assessment and Prediction Methods in Crime and Delinquency*, in PRESIDENT'S COMMISSION ON LAW ENFORCEMENT AND ADMINISTRATION OF CRIMINAL JUSTICE, *TASK FORCE REPORT: JUVENILE DELINQUENCY AND YOUTH CRIME* 171-87 (1967).

14. D. Gottfredson, *Assessment and Prediction Methods*, *supra* note 13, at 180.

small: in the order of approximately fifteen to thirty percent.¹⁵ With such relatively weak associations, a high rate of error could be expected and was found to occur. The error was of two types.

The first type of error was that of not identifying persons who subsequently *did* commit recidivist crimes. Even though the predictive instruments spotted some recidivists, they tended to miss many others. The rate of "false negatives" or "misses"—that is, persons mistakenly identified as good risks—was substantial.¹⁶ This meant that the prediction methods could not promise to protect the public against recidivist criminality with much confidence.

Still more disturbing was the reverse kind of error: "overprediction" of recidivism. In order to spot any appreciable number of actual recidivists, the prediction instruments produced an embarrassingly large number of false alarms or "false positives", that is, persons mistakenly predicted to be recidivists.¹⁷ The rate of false positives was particularly high when forecasting serious criminality. John Monahan has compared the results of six major studies that tested the accuracy of clinical and statistical forecasts of violence. In those studies, the rate of false positives tended to be over sixty percent.¹⁸ This problem is not easily remediable because it results from the rarity of the conduct to be forecasted. Serious crimes, such as acts of violence, are statistically rare events. The rarer the event, the greater will be the tendency to overpredict.¹⁹

15. The coefficient of determination, which is the square of the correlation coefficient, measures the amount of variability "accounted for" by a predictor item or scale. Often it is not reported; but it can be estimated from the reported data. As examples: the well-known prediction method of Mannheim and Wilkins (*supra* note 10) accounted, by our calculation, for about a fourth of the variability in success or failure; the California Base Expectancy scale from ten to twenty-five percent in various samples; and prediction methods based on federal offenders about fifteen percent. The correlations reported in such studies are substantial in relation to much of social science; but they must be considered quite modest if the aim is to use them in criminal justice decisions involving the liberty of individual offenders. For a discussion of the California Base Expectancy scale, see M. GOTTFREDSON & D. GOTTFREDSON, *DECISIONMAKING IN CRIMINAL JUSTICE: TOWARD THE RATIONAL EXERCISE OF DISCRETION*, 257-79 (1980). For the data on federal offenders, see S. GOTTFREDSON & D. GOTTFREDSON, *SCREENING FOR RISK: A COMPARISON OF METHODS* (1979).

16. For an analysis of error rates as a function of cutting scores, based on a Base Expectancy scale devised by Don M. Gottfredson for the California Department of Corrections, see D. GREENBERG, *MATHEMATICAL CRIMINOLOGY* 222-29 (1979).

17. For fuller discussions of the false-positives problem and its implications in sentencing, see von Hirsch, *Prediction of Criminal Conduct and Preventive Confinement of Convicted Persons*, 21 *BUFFALO L. REV.* 717 (1972), reprinted in *SENTENCING* 148-74 (H. Gross & A. von Hirsch eds. 1981) [hereinafter cited as *SENTENCING*]. For review of studies of prediction of violence and the false-positives problem, see Monahan, *The Prediction of Violent Behavior*, in *DETERRENCE AND INCAPACITATION: ESTIMATING THE EFFECTS OF CRIMINAL SANCTIONS ON CRIME RATES* 244-69 (A. Blumstein, J. Cohen & D. Nagin eds. 1978) [hereinafter cited as *DETERRENCE AND INCAPACITATION*].

18. J. MONAHAN, *PREDICTING VIOLENT BEHAVIOR: AN ASSESSMENT OF CLINICAL TECHNIQUES* 73-80, 101-04 (1981).

19. von Hirsch, *Prediction of Criminal Conduct*, *supra* note 17, at 733-39.

Moreover, trying to alleviate the false negatives problem aggravates the overprediction problem, and vice versa. If the researcher wishes to identify a larger fraction of the actual recidivists in his forecasts, the cutting line in the predictive instrument must be lowered, that is, the predictions must be made more inclusive. This, however, will sharply increase the already high incidence of false positives. Conversely, if the researcher wants to reduce the number of false positives, the cutting line can be raised and the instrument made more selective. This, however, will mean the instrument will tend to miss a still larger number of actual recidivists.²⁰

These findings helped diminish enthusiasm for predictive sentencing during the last decade.²¹ There remained some partisans of the idea—best exemplified, perhaps, by the American Bar Association's 1979 report on sentencing.²² But the high rate of false negatives tended to discourage law-and-order advocates who wanted effective protection against recidivist criminality. And the high false-positive rate deeply disturbed civil libertarians, who disliked the idea of depriving people of their liberty for anticipated crimes which in fact they well may not commit.

In examining the selective incapacitation model developed by Greenwood, we should ask whether significant progress has been made in overcoming these difficulties. Is the study properly designed? Do the prediction instruments succeed in spotting a larger fraction of those who subsequently become the most active offenders? Has the rate of false positives been reduced significantly compared to previous studies? Progress in the predictive art has been made only if these questions can be answered affirmatively.

B. The Design of the Prediction Study

Greenwood's prediction research, as well as the companion RAND studies by the Chaikens,²³ is more ambitious than earlier forecasting efforts. Older studies sought merely to predict whether or not an offender would return to crime. The new studies focus on particular species of crime—namely, on robberies and burglaries. Moreover, they address *frequency* of offending: not merely whether but how often convicted robbers or burglars will commit such crimes in the future. In order to obtain the necessary information, the studies rely on prisoners' self reports of their past criminal activities. What problems are there in the way this research has been constructed?

20. *Id.* at 735; D. GREENBERG, *supra* note 16, at 222-29.

21. von Hirsch, *Recent Trends in American Criminal Sentencing Theory*, *supra* note 1, at 11-12, 14-17.

22. *Id.* at 14-17. For further discussion of the ABA's report, see *infra* text accompanying note 98.

23. P. GREENWOOD, *supra* note 4; J. AND M. CHAIKEN, *supra* note 4.

1. *Is this Prediction?*

Prediction involves the assessment of some future state—in criminology, typically some expected criminal behavior. It involves the independent assessments of persons, separated over time: first, a specification of predictor variables at time t_1 ; next, the inquiry into a performance at time t_2 . The criterion of that performance must be defined and tested independently of any of the predictor variables. The point of the exercise is to ascertain whether the performance occurring at time t_2 does, or does not, bear out the forecast made at time t_1 .²⁴

The Greenwood study is not, in this sense, a prediction study at all. The data concerning offender characteristics (called “predictors”) and offense rates (the behavior “predicted”) all stem from one offender questionnaire soliciting prisoners’ self-reports of their past behavior. Essentially, the study describes a cross-tabulation (based on the same data) of (1) offender characteristics with (2) self-reported past offense behavior. There are no two independent data sets, and the information has not been sequentially obtained over time. When a cross-tabulation such as this is used as a predictive tool, a number of problems arises.

An offender’s criminal habits can change with age, or, perhaps, even with the experience of incarceration. To the extent a change occurs, the factors associated with criminality in his earlier criminal career will not necessarily be associated (or associated in the same degree) with the person’s subsequent patterns of behavior.²⁵

Several of the factors found by Greenwood to be associated with past criminality are not in a form that can be used for sentencing in an operational predictive index. Two of Greenwood’s seven factors are drug abuse and past joblessness.²⁶ It is one thing to find that high robbery rates are correlated with self-reported drug use and unemployment. It is quite another to use these factors in a forecasting index for sentencing purposes. The latter would require records or other available evidence that actually show the offender is a drug abuser, or lacks a job. Such evidence may be hard to come by, with the result that the predictive utility of such factors diminishes. (Since writing this Article, we have learned from a new analysis of the Greenwood study that these problems of operationalizing the predictive index are still more severe than we had envisioned.²⁷)

24. The steps involved and the requirements of prediction methods are discussed more fully in D. Gottfredson, *Assessment and Prediction Methods*, *supra* note 13.

25. This point has been made by Michael Maltz in a conversation with Andrew von Hirsch in March, 1983.

26. P. GREENWOOD, *supra* note 4, at 50.

27. We are referring to an unpublished lecture by Jan M. and Marcia R. Chaiken entitled *Deficiencies in Official Records for Identifying Serious Offenders*, presented on November 10, 1983 at the annual meeting of the American Society of Criminology in Denver.

Moreover, the Greenwood study is not even complete as a cross-tabulation study of self-reported behaviors of robbers and burglars, for its validity has not been confirmed using a new sample. Replication to ascertain whether similar results are found with other samples is essential for determining how much the study's conclusions can be generalized, as opposed to reflecting merely the idiosyncracies of the original sample. When such replication is attempted there tends to be shrinkage: the independent variables will usually account for a smaller portion of variation in the replication sample than was true in the original sample. With replication omitted, the conclusions are thus likely to overstate the accuracy of the "prediction index."²⁸

2. *Representativeness of the Sample as a Sample of Inmates*

The Greenwood sample is composed of selected inmates in jails and prisons in California, Michigan and Texas.²⁹ Since some of the persons selected did not complete the questionnaire, an analysis was done to detect a possible resulting bias; none was found except in California, where Hispanic inmates were underrepresented. No comparable analysis was done, however, to see whether the samples could be considered representative of all California, Michigan, or Texas inmates; thus caution should be used in generalizing the results to inmate populations other than those from the sites of data collection and certainly to other jurisdictions.

3. *Reliance on Inmate Self-Reports*

The use of prisoner self-reports is critical to the design of this RAND research. The self-reports disclose the high robbery and burglary rates among a significant minority of respondents. Official reports (arrests or convictions) would show a much lower rate, thereby eliminating the possibility of identifying prospective "high-rate" offenders.

Self-reports are an established tool of social science.³⁰ They tend to be most credible when there is substantial independent corroborative evidence, when the risk of untruthfulness can reasonably be discounted, and when the self-report studies can be replicated over time. Unfortunately, these conditions are not met in this research.

There is, first, little opportunity to corroborate respondents' answers. The investigator can (and the RAND researchers did) check whether pris-

The Chaikens' analysis points out that of the seven factors utilized in Greenwood's predictive index, five either had no official records in the data or else proved to have no predictive utility when official records rather than self-reports were used.

28. For fuller discussion of the problem of shrinkage, see F. SIMON, *supra* note 13.

29. P. GREENWOOD, *supra* note 4, at 19-26.

30. M. HINDELANG, T. HIRSCHI, & J. WEIS, *MEASURING DELINQUENCY* (1981).

oners' reports of their crimes included those for which they were previously arrested or convicted.³¹ The respondents could be expected to be aware that the researchers had access to this information. The problem is that, especially for the alleged high-rate offenders, their self-reported crimes that had not led to official action exceeded by many times the crimes for which they had been previously arrested or convicted. There is no independent corroboration for these crimes.

Second, the research is not focused on the behavior of the majority of respondents, unlike much self-report research in the social sciences. If it were, one could discount the risk of untruthfulness among some respondents. Instead, the research concentrates on the behavior of those who report themselves to be a troublesome minority; it is the high-risk offenders that the instrument is meant to identify. It seems to be assumed that these individuals (like the homicidal Irish maiden in Tom Lehrer's famous song) are happy to wreak the worst mayhem, but know that lying is a sin. Perhaps persons who supposedly are so very lawless might be less than candid. Perhaps, also, many exaggerate their criminal activities in order to cast themselves in the tough role that plays such a part in macho subcultures.

These problems would not be so severe were it possible to use self-reports merely for the purpose of constructing the predictive instrument, and if the instrument then could be validated using more direct evidence of offenders' crime rates. But this cannot be done because, again, official arrests or convictions will not adequately reflect differential offense rates. Thus the accuracy of the predictive instrument could only be confirmed on another sample of prisoner self-reports.

A final problem of reliability concerns the replicability of such research. The respondents in the RAND studies may not have known it when they completed their self-reports, but their answers could vitally affect their interests: Greenwood is proposing employing those answers to develop a sentencing policy that could involve lengthy terms of imprisonment for supposed high-rate offenders.³² Once this fact becomes known (news travels in prisons and this research has received extensive publicity), the answers might change dramatically. This is not the kind of research that assuredly can be repeated and refined over time—unlike self-report research which truly does not affect respondents' own interests.

Are these problems so serious as to destroy the usefulness of these self-report data? The answer depends on the use to which they are put. Were this general criminological research, one might offer it as providing some insight

31. P. GREENWOOD, *supra* note 4, at 26. For fuller discussion, see M. PETERSON, J. CHAIKEN, P. EBENER & P. HONIG, *SURVEY OF PRISON AND JAIL INMATES: BACKGROUND AND METHOD* (1982).

32. See *infra* note 83.

into patterns of offending, subject to the caveats just explained. But the proposed use is much more ambitious. Greenwood speaks of using this research to create predictive instruments that would determine imprisonment policies. Self-report data seem a very fragile basis, indeed, for making decisions that so drastically affect people's liberty.

C. *The Results of the Prediction Study*

1. *The Predictive Factors Identified*

What factors, according to this research, identify high-risk offenders? They are, as Greenwood himself points out, similar to those typically found to be predictive of recidivism. They include the offender's criminal history, age at first involvement with criminal justice agencies, drug involvement and employment history.³³ No novel indicia of future criminality are found here.

2. *Amount of Variability Explained*

We mentioned that earlier prediction studies explained only a limited portion of the variability in behavioral outcome within the samples they studied. Does the new research account for any larger portion? Unfortunately, it does not. Greenwood reports that significance tests were made for each of his seven "predictive" factors, but does not specify measures of association. Using Greenwood's data, we calculated the degree of association with outcome of each of his seven items. None was strong.³⁴

33. P. GREENWOOD, *supra* note 4, at 50. The seven factors in Greenwood's prediction scale are:

1. Prior conviction for the instant offense type.
2. Incarceration for more than half of the preceding two years.
3. Conviction before the age of 16.
4. Time served in a state juvenile facility.
5. Drug use during the preceding two years.
6. Drug use as a juvenile.
7. Employment during less than fifty percent of the preceding two years.

Using this scale, he develops three categories: predicted low-rate, medium-rate and high-rate offenders. The categories are defined as follows:

<i>Offender Risks Category</i>	<i>No. of Factors Present</i>
low	0-1
medium	2-3
high	4 or more

Items similar to Greenwood's seven factors often have been found to be predictors in the past. See Gottfredson, *Assessment and Prediction Methods*, *supra* note 13; S. GOTTFREDSON AND D. GOTTFREDSON, *supra* note 15; D. GLASSER, *THE EFFECTIVENESS OF A PRISON AND PAROLE SYSTEM* (1964), ch. 3.

34. The degree of association ranged from .11 to .25 using Cramer's statistic. This measure can be used with any k by 1 contingency table and, like the correlation coefficient, may range from 0 to 1.00. It is calculated from the Chi-squared values and is symbolized by ϕ' .

Greenwood reports how performance on his "prediction" score correlates with low, medium or high self-reported crime rates.³⁵ However, he does not examine the efficiency of his prediction device: no attempt is made to employ statistical measures of predictive efficiency, such as the fraction of variability in crime-rate classifications "explained" by the prediction scores, measures of association, or Mean Cost Rating. Our calculation of such measures of efficiency show that the scale is comparable in efficiency to traditional recidivist prediction methods, but fails to improve on them.³⁶

3. False Positives and False Negatives

With so little of the variability explained, one can expect a high rate of false positives and negatives. Greenwood does not set forth these rates in his report in any clear or detailed fashion. Our calculations from his data, however, are not reassuring. Of those offenders in his sample "predicted" by his index to be high-rate offenders, less than half were in fact so: the remainder showed medium or low rates.³⁷ The false positive rate was thus

The values obtained for the items selected by Greenwood, all of which were statistically significant ($P \leq .001$) were: Prior conviction for current offense $\chi^2 = 19.5$, 4 d.f., $\phi' = .11$; Incarceration for more than 50% of two years preceding current arrest, $\chi^2 = 31.4$, 2 d.f., $\phi' = .20$; Conviction before age 16, $\chi^2 = 22.0$, 4 d.f., $\phi' = .12$; Juvenile commitment to state facility, $\chi^2 = 25.0$, 4 d.f., $\phi' = .13$; Drug use in past two years, $\chi^2 = 97.33$, 4 d.f., $\phi = .25$; Drug use as juvenile, $\chi^2 = 70.3$, 4 d.f., $\phi' = .21$; and Employment during less than 50% of two years preceding current arrest, $\chi^2 = 36.8$, 4 d.f., $\phi' = .15$.

35. P. GREENWOOD, *supra* note 4, at 54-61.

36. The data of Table 4.5, P. GREENWOOD *supra* note 4, at 53, permit calculation of the number of cases within cells from the percentages shown. Scores range from 0 to 5 and higher, with three categories of reported crime rates, yielding a six by three table. Our calculation of the Chi-squared value was 153.4 (10 d.f.). The value of Cramer's statistic for the combined seven-item score was .31.

The Mean Cost Rating, which measures the power of a prediction instrument by contrasting errors and correct predictions, Duncan, Ohlin, *et al*, *Formal Devices for Making Selection Decisions*, 58 AMER. J. SOCIOLOGY 573 (1953), has been used widely in criminology, F. SIMON, *supra* note 13. When the Greenwood scale was collapsed into three groups, as proposed, and the outcome variable dichotomized, the Mean Cost Rating statistic was found to be .40. Kendall's tau, a measure of rank correlation, may be derived from the Mean Cost Rating, Lancucki and Tarling, *The Relationship between Mean Cost Rating and Kendall's Rank Correlation Coefficient* in D. GOTTFREDSON, L. WILKINS AND P. HOFFMAN, *supra* note 12, at 199-206. It was found to be .30. With scores 0 to 5 and higher retained and the dichotomous criterion (high vs. low or medium), the obtained point biserial correlation coefficient was .22. Since the Greenwood study did not use multi-variate statistical methods, the total fraction of variability in outcomes explained by his predictor items in combination is not reported. In the companion Rand study, J. CHAIKEN & M. CHAIKEN, *supra* note 4, which used multiple regression in related analyses of the same data base, these statistics are reported and are within the range traditionally found: the total fraction of variability in outcomes (R^2) is about .32. *Id.* at 113.

37. These figures are readily calculated from Table 4.5 of Greenwood's report. P. GREENWOOD, *supra* note 4, at 59. However, the text of the report, *id.* at 59, fails to make clear that the false-positives rates are so high. For a critical discussion of Greenwood's

fifty-six percent, which shows scant improvement over previous studies. The false-negative rate was also substantial: of those "predicted" to be low or medium rate offenders, one-sixth (sixteen percent) in fact showed high rates. In sum, Greenwood's technique for predicting potential robbers shows little improvement in accuracy over forecasting methods of the past.

III

PURPORTED EFFECTS ON THE INCIDENCE OF CRIME

Let us turn, next, to Greenwood's assertions about crime rates. Traditional prediction studies made no special claims about being able to reduce the overall incidence of crime. Identifying the potential recidivist and imprisoning him was said to safeguard one from injury at *his* hands; but since other felons remained at large; it did not necessarily reduce the net danger of victimization. Greenwood makes the novel promise that a selective incapacitation strategy will reduce that larger danger. He asserts that by lengthening terms of confinement for convicted robbers who are identified as high risks and shortening terms for medium and low risk robbers, the robbery rate can be reduced as much as *fifteen* percent.³⁸

This is the assertion that has attracted so much public attention. Crime rates—and most importantly, rates of serious, predatory crimes such as robbery—having defied crime-control efforts for decades, supposedly can be brought down at last, using a simple sentencing technique. Or rather, they can *if* the projection model underlying Greenwood's crime-rate calculations is correct. Is it correct?

A. *The Shinnar Projection Model: The Rise and Fall of "Collective" Incapacitation*

Greenwood's model for projecting the impact on crime rates is borrowed from the "collective incapacitation" research done in the mid-seventies, and revised. Collective incapacitation was a notion popularized by Professor James Q. Wilson in his influential book, *Thinking About Crime*.³⁹ The strategy called for uniform prison sentences to be imposed on all defendants convicted of specified felonies. No attempt would be made to identify which felons were high risks. By removing from circulation all those convicted of such crimes for a portion of their potential criminal careers, one would prevent all of them who were so inclined from offending again

manner of reporting the false-positive figures, see Blackmore and Welsh, *Selective Incapacitation: Sentencing According to Risk*, 29 CRIME & DELINQ. 504, 516-17 (1983).

38. P. GREENWOOD, *supra* note 4 at xix.

39. J. WILSON, THINKING ABOUT CRIME 162-82, 198-209 (1975).

while they were confined.⁴⁰ The strategy would thus interfere with their recidivist proclivities: It would confine both the seemingly high-risk offenders and others who would in fact have recidivated. Wilson promised great crime-control gains from such a strategy; speaking of its impact on the incidence of robbery, he said: "Were we to devote [our] resources to a strategy that is well within our abilities—namely, to incapacitating a larger fraction of the convicted serious robbers, then not only is a twenty percent reduction [in robbery] possible, but even larger ones are conceivable."⁴¹

Wilson derived these optimistic estimates from the projection technique developed at the time by Reuel Shinnar.⁴² Shinnar invented a new variable named Lambda (λ), representing the average annual rate at which individual offenders commit offenses in the community. Given an appropriate estimate for Lambda, a simple formula can be derived for calculating the average fraction of time an offender spends on the street. This fraction, Shinnar showed, depends on (1) the average offense rate (λ), (2) the average probability of being arrested, convicted and incarcerated for a crime, and (3) the average duration of confinement if incarcerated. If one knows the fraction of time offenders spend at liberty, and has data on the number of offenders now confined, one can readily estimate the total number of persons committing crimes. Next, one can calculate the resulting total number of crimes committed: it will be the number of persons committing crimes multiplied by the fraction of their time they spend on the street, multiplied by their average crime rate. A collective incapacitation strategy, if adopted, will alter one of the relevant variables; for instance, it could raise the likelihood of incarceration if convicted for given offenses to something approaching one hundred percent. Given this change in policy, one can then recalculate the resulting total crimes. The resulting decrease in the number of crimes, according to Shinnar, measures the preventive effect of the strategy. For the mathematically inclined, Shinnar's formula is set forth in the Appendix.⁴³

In his calculations, Shinnar assumed a high average rate of offending.⁴⁴ With this high Lambda, his model suggested (not surprisingly) that the crime-reduction effect of a collective incapacitation strategy could be large.⁴⁵ Wilson, in turn, used similar estimates of average offender crime rates to support his promise of a twenty percent reduction in robbery rates.⁴⁶ When other researchers recalculated the effects using lower estimates for

40. Advocates of incapacitation strategies typically do not count crimes committed by prisoners against fellow-inmates and guards.

41. J. WILSON, *supra* note 39, at 199.

42. Shinnar and Shinnar, *The Effects of the Criminal Justice System on the Control of Crime: A Quantitative Approach*, 9 L. AND SOC'Y REV. 581 (1975).

43. See *infra* text accompanying notes 125-27.

44. Shinnar and Shinnar, *supra* note 42, at 590.

45. *Id.* at 605-08.

46. J. WILSON, *supra* note 39, at 199-201.

Lambda, however, the resulting crime-reduction effects of the strategy shrank dramatically.⁴⁷

In 1978, these varying estimates were examined by a panel of well-known econometricians and criminal justice researchers.⁴⁸ The panel, organized by the National Academy of Sciences, found that the crime-rate effects of a collective-incapacitative strategy depended critically on the estimates that were made of Lambda, the average offense rate. If Lambda was assumed to be high, the incapacitative pay-off appeared substantial, but crime-rate effects were much more modest when a lower offense rate was assumed. The panel noted both the absence of adequate empirical evidence of offense rates and the difficulty of estimating those rates.⁴⁹ The panel's report tended to deflate the then-current enthusiasm for collective incapacitation.⁵⁰

*B. Greenwood's Projection Model: Altering the Shinnar
Formula to Fit Selective Incapacitation*

While interest in collective incapacitation declined, Shinnar's model was not forgotten. Greenwood revised the Shinnar formula to fit a selective incapacitation strategy. His idea was to establish not one but three Lambdas: the Lambdas would be the estimated average robbery rates, respectively, of low, medium, and high-risk robbers.⁵¹ Greenwood asserts these rates can be obtained from his data: one simply uses the average self-reported annual robbery rates of the individuals falling into the three risk categories established by his "predictive" instrument. With these three Lambdas, Greenwood calculates (in much the same way Shinnar did in his formula) the aggregate number of robberies contributed by each of the three groups under present sentencing policy. Next, the proposed new selective-incapacitation policy can be mathematically described: for high-rate offenders, it would increase the probability of imprisonment given conviction for robbery, the duration of confinement for robbery, or both; for low and medium rate offenders, it might keep present practices or decrease either the likelihood or duration of confinement. With these policy changes thus specified, Greenwood figures for each of the three risk groups the effect on their total estimated crimes. The extent of the reduction in estimated crime

47. These studies are summarized and discussed in Cohen, *The Incapacitative Effect of Imprisonment: A Critical Review of the Literature* in DETERRENCE AND INCAPACITATION, *supra* note 17, at 187-243.

48. Panel on Research on Deterrent and Incapacitative Effects, *Report* in DETERRENCE AND INCAPACITATION, *supra* note 17, at 64-75.

49. *Id.*, at 67-69, 80; Cohen, *supra* note 47, at 228-29.

50. von Hirsch, *Recent Trends in American Criminal Sentencing Theory*, *supra* note 1, at 17-20.

51. P. GREENWOOD, *supra* note 4, at 74-76.

represents the crime prevention impact of the new strategy.⁵² Greenwood's calculations suggest the crime reduction could be substantially larger than a collective-incapacitation strategy using a uniform policy for imprisoning convicted robbers.⁵³

*C. Dubious Evidence on Individual Robbery
Rates: Extrapolating from Data about
Incarcerated Offenders*

Shifting from one Lambda to three still leaves us with the question of the soundness of Greenwood's method of estimating average offense rates for the members of his three risk groups. Greenwood claims that he has obtained average annual robbery rates for low, medium and high-risk robbers. Based on these rates he "finds" that a limited number of high-risk robbers are responsible for a very large number of robberies. Hence, he argues that identifying and incarcerating such persons would yield substantial crime prevention benefits.

In fact, Greenwood has no basis for these average robbery-rate estimates. Given the research design, he can only study the purported robbery habits of incarcerated robbers: those who happen to be in prison or jail at the moment, and whose self-reports can thus be obtained. He reports that the probability of being incarcerated for a robbery is low. (In California, the probability of arrest and conviction for robbery is .03, and the present probability of incarceration if convicted is .86. This means that the probability of arrest, conviction and incarceration for a given robbery is a mere .0258.)⁵⁴ If this probability is so low, then a sample of robbers such as that selected by Greenwood—by virtue of its members' happening to have a current conviction and incarceration for robbery—may be a highly unrepresentative one. No data is available in the study to help ascertain whether incarcerated offenders' self-reported robbery habits are reflective of those of active robbers generally. Greenwood's method is reminiscent of the researcher who makes "findings" about the drug habits of addicts in a given community by studying the drug history of a limited number of addicts residing in inpatient drug treatment centers. Such findings would likewise be of little or no value, because the addicts studied might be wholly unrepresentative of the general population of drug users.

A closer look at Greenwood's calculus for estimating crime-rate effects confirms the problem. The critical formula in Greenwood's model is the one that allows one to calculate the fraction of time an offender typically spends

52. Greenwood's mathematical model is described in the Appendix. See *infra* text accompanying notes 128-30.

53. P. GREENWOOD, *supra* note 4, at 78-85.

54. *Id.* at xvii, 108-13.

on the street, based on the estimated Lambda and on the probability of incarceration given an offense and the duration of incarceration.⁵⁵ That fraction, as we saw, is used to determine both the estimated number of persons offending and the incidence of offenses associated with a given sentencing policy. But one can estimate this fraction-of-time-free accurately only if one knows the true Lambda, that is, the actual average offense rate per robber. But the data in the study cannot provide this; it can furnish only average-number-of-crimes-per-offender estimates for incarcerated offenders whose self-reports can be obtained.⁵⁶ The study can provide no data on number-of-crimes-per-offender for the offenders who were not confined when he did his research, and those numbers may be very different. If so, then his estimates of the true Lambda for low, medium and high-risk offenders may be seriously inaccurate when generalized to the whole population of robbers. To the extent those Lambda estimates are wide of the mark, the number of crimes committed by the target group of "high rate" offenders may be substantially overestimated. Such overestimation will, in turn, distort the projections of the crime prevention effects of a selective incapacitation strategy⁵⁷ (see the Appendix for the mathematical specifics.)⁵⁸

Greenwood scarcely addresses this objection, except to make this comment: "Our sample has been criticized for including only incarcerated offenders. This criticism would perhaps be justified if there were evidence to suggest the existence of a significant group of offenders not subject to arrest and incarceration. There is no such evidence."⁵⁹

His point seems to be that his incarcerated-offender based research design could be invalid only if one assumes there exists large numbers of high-frequency offenders who are never caught and punished. One can, however, challenge Greenwood's design without having to make this unlikely supposition. If either of the following scenarios were true, and they

55. *Id.* at 75.

56. We note above (see *supra* text accompanying note 29) that Greenwood lacks a fully representative sample of incarcerated offenders; his sample is limited to those confined in certain facilities in the three states. By broadening the inmate survey, however, one could obtain such a sample, and hence obtain self-reported offense-rate data for persons now in confinement. Our point here is that it remains impossible, using inmate self-report data, to learn the offense rates (Lambdas) of offenders *not* now in confinement—and that this ruins his estimates of preventive impact.

57. Greenwood purports to confirm his calculations by noting that the number of robberies attributed by his calculations to members of each of the three risk groups under current sentencing policy, when added together, roughly equals the total number of actual reported robberies. P. GREENWOOD, *supra* note 4, at 117-18. However, this evidence in no way supports Greenwood's estimates of the *distribution* of robbery activity among different risk groups, which is critical to his calculations. He offers no confirmation of his essential claim that a large proportion of robberies are committed by a limited number of high-risk offenders.

58. See *infra* text accompanying note 131.

59. P. GREENWOOD, *supra* note 4, at 89.

are not the only possible ones, Greenwood's reliance on data taken from incarcerated offenders would lead to serious errors in his estimates. These scenarios do not require one to assume that there are many Dr. Moriarty's of armed robbery at work, busily committing crimes and never being detected.

Scenario 1: Overestimating Lambda Among High-Scoring Offenders.

Among offenders who scored high on his predictive index, Greenwood found a startling average robbery rate of as many as thirty per year in California.⁶⁰ This steep robbery rate may, however, have been an artifact of having considered only the incarcerated robbers. One can accumulate a score that puts one in the worst risk category in Greenwood's prediction index by being unemployed, being addicted with a history of addiction, having an early juvenile conviction and having been confined as a juvenile.⁶¹ Were it possible to examine the habits of *unincarcerated* offenders with these characteristics, it might have emerged that many such individuals never had high robbery rates or are now losing their criminal initiative. (In fact, it is because they are committing fewer robberies that they are under-represented in Greenwood's sample). Greenwood's research design, therefore, may have substantially overestimated the Lambda for robbers who fit his criteria of being bad risks. This, in turn, would lead him to exaggerate the aggregate crimes committed by such persons, and hence to overstate the incapacitative effect of confining this group for longer periods.⁶²

Scenario 2: Underestimating the Impact of Occasional Criminals.

Greenwood has assumed a uniform probability of arrest and conviction given a robbery.⁶³ In fact, he has derived that probability simply by dividing the number of robbery convictions by the number of robberies.⁶⁴ But it is easy to imagine a more complex scenario. First, the likelihood of arrest and conviction may vary considerably with the extent of the person's criminal history and other signs of "trouble" such as apparent drug addiction.⁶⁵ Second and more serious, it may be difficult or impossible to obtain any reliable estimate of what that probability is for occasional offenders with little previous contact with criminal justice agencies. The upshot is that occasional offenders may be responsible for a much larger proportion of

60. *Id.* at 57.

61. *Id.* at 50-53.

62. Michael Maltz states the point this way: the distribution of robbery-rates (Lambdas) among robbers generally may, to an unknown degree, differ from the distribution of such rates among incarcerated robbers. Hence, one cannot infer robbery rates for robbers in general from the self-report rates of confined robbers. See Maltz, *supra* note 25.

63. P. GREENWOOD, *supra* note 4, at 74-77.

64. *Id.*, at 109.

65. P. Greenwood dismisses this possibility by saying previous RAND inmate surveys suggest that the probability of arrest and conviction does not vary with the person's offense rate. *Id.* at 108. This, however, begs the same question: Whatever may be true of incarcerated offenders, Greenwood has no evidence regarding this issue among offenders at large.

aggregate robberies than Greenwood's calculations recognize. Were this the case, the crime-control effect of confining the ostensible bad risks would be much more modest than claimed.

We are speaking here of the possibility of major distortions in his estimates, because Greenwood's extrapolations from his data are so large. Consider a possibly typical jurisdiction X, where the average risk of being convicted and imprisoned for a robbery is assumed to be .025.⁶⁶ Suppose one divided the inmates with robbery convictions in the manner that Greenwood did among low, medium and high scorers on his prediction scale. Suppose the high scorers reported a fairly high average robbery rate, one well above the rates Greenwood reports for Texas, but not quite so high as his California rates. Let us assume twenty robberies per year.⁶⁷ And let us imagine that these high scorers constitute one quarter of the population of imprisoned robbers.⁶⁸ Using these figures and Greenwood's model, one would extrapolate that there are nearly *twice* as many active high-rate robbers as those he has actually identified as such in confinement and that members of this much larger group each are committing twenty robberies a year while at liberty. It is because he attributes this high rate to so greatly enlarged a group that he can "derive" large crime reductions from a policy of selective incapacitation targeted on such offenders. What we have just seen, however, is that both his estimates of the size of this enlarged group and of its average offense rate may be quite wrong—and if so, his estimates of the benefits of selective incapacitation may be seriously in error as well.⁶⁹

D. Can More Conservative Estimates Be Made?—Restricting the Estimates to Incarcerated Offenders.

Can any inferences be made from Greenwood's self-report data? Much greater caution would be necessary. On the basis of data about the robbery-rates of incarcerated offenders, one might try to project the preventive effects of selective-incapacitation policies considering only the anticipated activities of members of that inmate group. What is plainly unwarranted is to draw inferences about the behavior of robbers generally, including those whose offense patterns have not been and cannot be studied by the inmate self-report research design.

66. This is approximately the same probability as Greenwood uses for his California calculations. *Id.* at 77.

67. According to Greenwood's study, *id.* at 58, the average robbery rate for high-scoring robbers is 31 in California, and 20 and 7, respectively, in Michigan and Texas. The assumed rate of twenty robberies per year thus coincides with Greenwood's figures for Michigan.

68. This corresponds to Greenwood's initial definition of high-scoring robbers as those having robbery-rates in the upper quartile. *Id.* at 49.

69. For calculations, see *infra* text accompanying notes 134-37 of the Appendix.

The robbers now incarcerated will eventually be released. After release, some will return to jail or prison with a frequency that may be expected to depend in part on the rate at which they continue to commit robberies. A more conservative calculation of preventive impact, then, is one that considers only the future robberies of these incarcerated offenders, not the elusive criminal behavior of others never studied. The calculation would not be difficult. The three Lambdas would be the self-reported yearly average number of robberies per offender for the three risk groups identified by Greenwood's predictive index.⁷⁰ The number of persons in the three risk groups would simply be the number of offenders now incarcerated that score well, medium or badly, respectively, on the prediction score, not some larger extrapolated number that includes offenders at large. One can, by using Greenwood's formula, calculate the total number of robberies forecast for this group, were present sentencing policy continued after the members of the group returned to the community. And using the same calculation techniques, one can calculate by how much this number would be reduced if a selective incapacitation policy were introduced after their release from current confinement.

When making the calculation in this more cautious fashion, the crime-reduction effects of selective incapacitation strategies shrivel, even were one to accept Greenwood's other assumptions and his estimates of the robbery rates of the offenders he has studied. In the hypothetical jurisdiction spoken of earlier, the number of robberies prevented by a selective-incapacitation strategy would (as shown in the Appendix⁷¹) become only half of what they would be under Greenwood's calculation method. Greenwood's estimates, evidently, are at least half watered-stock.

*E. Problems Even with the "Conservative"
Estimate: Crime-Rate Fluctuations and
Replacement Effects*

Yet even this scaled-down estimate is apt to be inflated, for two reasons.

(1) *Variability of Robbery-Rates Over Time.* The calculations assume that now-incarcerated robbers will continue in the future to have the same criminal propensities (and hence show the same robbery-rates when at liberty) that they have had in the past, as shown in their self-reports. The assumption may well be mistaken. The offenders registering the worse scores on Greenwood's predictive index are those who already have exten-

70. This assumes that offenders continue robbing at the same rate they report having done recently. For a discussion of the problems with this assumption, see *infra* text accompanying notes 71-72.

71. For the calculations, see *infra* text accompanying notes 132-37.

sive criminal records. Many of these individuals may be of an age when they are past, or shortly will be past, the most active phase of their criminal careers. After release from their current incarceration, these offenders' robbery rates may begin to decline significantly. To the extent that is the case, using Lambdas based on their self-reports will overstate the expected preventive effect.

(2) *Replacement Effects.* The calculation completely ignores replacement effects: removing a high-rate robber from circulation, it is assumed, will eliminate a number of offenses exactly equal to his predicted number of robberies.⁷² This assumption is unrealistic. Albert Reiss⁷³ has suggested, for example, that the bulk (over seventy percent) of robberies committed by younger offenders are committed in groups and Greenwood himself cites statistics suggesting a substantial incidence of robbery perpetrated by several offenders acting together.⁷⁴ Removal of one member of such a group is scarcely apt to terminate the group's activities: the group can either recruit replacements or continue its robberies without a replacement; in either event no crime reduction would occur. Reiss even raises the possibility of an increase in crime, if removal of a member causes the group to split into two nuclei which continue robbing separately. Reiss' point about group robberies is only one of a variety of replacement effects that could occur. Were allowance made for possible replacement effects, this would reduce the crime-prevention payoff of selective incapacitation still further. One ends with little confidence, indeed, that predictive sentencing strategies such as Greenwood's will have any significant impact on overall robbery rates.

F. Problems of Confirming Incapacitative Impact

Crime-prevention calculations of the kind Greenwood is making are, at best, projections of what *might* be the benefits of selective incapacitation.

72. Greenwood's only defense of his having not considered replacement effects in his calculation is this hardly convincing statement. P. GREENWOOD, *supra* note 4, at 91: "Any speculation that the effects of incarcerating one offender may be cancelled out by his replacement on the streets by another must be balanced by speculation about possible deterrent effects. The existing evidence suggests these effects cancel each other out." In fact, there is no evidence whatever that the deterrent effect cancels out the replacement effect. Greenwood, in his own review of the literature on deterrence, asserts that little is now known about the magnitude of deterrent effects, *id.* at 3-6. His own cited statistic that California robberies involve an average of 2.3 offenders per offense, *id.* at 109, suggests the replacement problem may be a substantial one.

73. Reiss, *Crime Rates and Victimization*, in INDICATORS OF CRIME AND CRIMINAL JUSTICE: QUANTITATIVE STUDIES 11-17 (S. Fienberg & A. Reiss, eds., 1980).

74. Greenwood's statistics indicate that the average number of offenders per robbery incident in California is 2.3. P. GREENWOOD, *supra* note 4, at 109. However, he ignores the implications of this fact, and treats robbers in his calculations of preventive effect as though they are robbing singly.

They require verification, for they may be far mistaken for any of the variety of reasons we have mentioned. Verification means the capacity, once a selective-incapacitation strategy has been tried in a given jurisdiction, to confirm whether it has indeed yielded the promised crime-prevention benefits. If there is no way of telling whether a predictive sentencing policy actually has reduced crime, if there can only be unverifiable projections of the expected crime-rate impact, then we are dealing with elusive promises.

How could one confirm such projections? Suppose that Greenwood's projection method showed that a given selective-incapacitation policy would, in a hypothetical jurisdiction X, show a stated percentage reduction in robbery rates; and suppose our more conservative calculation method halved the projected reduction. Imagine, finally, that the jurisdiction adopted the proposed policy; and that subsequently the incidence of robberies either did or did not change. How could one tell whether that outcome confirmed or disconfirmed the projection?

An incapacitation effect would be confirmed only if two conditions obtained: (1) the robbery rate declined after introduction of the selective-incapacitation policy, and (2) it could be shown that the decline was attributable to this policy and not other influences. The difficulty lies in satisfying the second requirement. The crime rate is affected not only by sentencing policy but also by a host of other, perhaps more powerful, influences such as demographic and social changes. Techniques for identifying and adequately controlling for these other variables remain rudimentary. And there are other knotty problems, such as determining whether the direction of causality runs from changed sentences to reduced crime rates or vice versa.

The problems of confirmation are, in short, much the same as those that arise when trying to measure the deterrent effect of a change in sentencing policy. Those difficulties—of controlling for other possible influences on crime rates, determining the direction of causation, etc.—are well described in the report of the National Academy of Sciences panel on deterrence research.⁷⁵ This influential panel reached distinctly pessimistic conclusions about our ability to verify the crime-rate impact of deterrence strategies.

IV

ETHICAL ISSUES: FALSE POSITIVES AND UNDESERVED PUNISHMENT

During the first six decades of the century, few questions were raised about the morality of predictive sentencing. The Model Penal Code, pub-

75. Panel on Research on Deterrent and Incapacitative Effects, *supra* note 48, at 19-63. The Panel report focuses on the problems of correlational studies of deterrence, using multivariate statistical techniques.

lished in 1962, spoke for the consensus among American penologists when it made the offender's predicted likelihood of recidivism an important (perhaps the most important) determinant of the sentence.⁷⁶ In the 1970's, doubts began to be raised in earnest. The tendency of forecasts of criminality to over-predict became a matter for concern,⁷⁷ as did the issue of undeserved punishment.⁷⁸ If current selective incapacitation proposals represent any improvement over traditional predictive sentencing, they should provide better answers to these doubts. Do they?

A. *The Morality of Overprediction*

In predictive sentencing schemes, being classified as a high risk may have substantial consequences for the offender's liberty: the convicted robber, for example, may either be made more likely to be imprisoned than he otherwise would, or else the duration of his confinement may be increased, or both. The rationale for thus depriving such persons of their liberty is that doing so will prevent them from infringing upon the rights of others. But to the extent that the prediction is mistaken, that infringement would not have occurred: the false positive loses his liberty on account of a predicted injury that he would not have in fact committed. He is being confined merely because some persons like him commit further crimes, and because we do not know enough to tell the false positives from the true.

This false-positives argument does not directly challenge the propriety of restraining people for expected future crimes; rather, it focuses on our ability to spot future criminals accurately. The force of the objection, therefore, will vary with two factors: (1) how high is the rate of false positives and (2) how much weight does the prediction carry in the choice of sentence.

Past studies of predicting serious criminality showed very high false positive rates indeed. We have mentioned Monahan's analysis of six clinical and statistical studies of violence prediction which showed false-positive rates over sixty percent.⁷⁹ Monahan has suggested that it might be possible to reduce the false-positive rate by restricting the use of prediction to very carefully selected sub-populations with extensive histories of violence, and by focusing on short-term forecasts.⁸⁰ But the Greenwood study surely does not achieve a significant reduction. We have seen that over half of those

76. MODEL PENAL CODE § 7.01(1)(a) (Proposed Official Draft 1962).

77. See von Hirsch, *Recent Trends in American Criminal Sentencing Theory*, *supra* note 1, 11-12; von Hirsch, *Prediction of Criminal Conduct*, *supra* note 17.

78. A. VON HIRSCH, *DOING JUSTICE: THE CHOICE OF PUNISHMENTS* 125 (1976) [hereinafter cited as *DOING JUSTICE*].

79. See *supra* text accompanying note 18.

80. J. MONAHAN, *supra* note 18, at 123-28, 143-69.

classified as high-rate offenders under his proposed index would in fact show medium or low robbery rates,⁸¹ even if there were no "shrinkage" of the predictive power of the instrument upon the study's replication.

Another way of alleviating the problem might be to decide the sentence chiefly on other grounds (such as desert) and give prediction only modest scope in determining its severity. This would reduce the undesired impact of the false-positive errors. One might be more willing to tolerate errors of overprediction, so long as the consequences for the offender's liberty were not drastic.⁸² But the Greenwood study does not take this route, either. While Greenwood does not offer firm recommendations about how much weight his prediction score should be given in the choice of sentence, he offers a variety of options all of which give it a substantial impact.⁸³ And his projections of large crime reductions are based on the more ambitious of these options which would make the duration of imprisonment depend very heavily indeed on predicted risk.⁸⁴

Greenwood's main defense to the false-positive objection seems to be that the sentencing system long has relied on and will continue to rely on forecasts notwithstanding their tendency to overpredict and that his statistically-based index is likely to be less vulnerable on this ground than judicial decisions based on hunches about who is dangerous.⁸⁵ This argument assumes (1) that there is consensus about the propriety of predictive sentencing; and (2) the only choice is the practical one between individualized predictive judgments by judges and system-wide predictive policies based on statistical studies. But no such consensus exists. Recently-developed sentenc-

81. See *supra* text accompanying note 37.

82. For discussion of this issue, see A. VON HIRSCH AND K. HANRAHAN, *THE QUESTION OF PAROLE: RETENTION, REFORM OR ABOLITION?* 31 (1979).

83. Greenwood considers four alternative selective incapacitation policies. The least ambitious of these involves lengthening the prison term for high-rate robbers, while keeping punishment policies for low and middle-rate robbers unchanged. The most ambitious involves reducing terms for low and middle rate offenders to a year in jail, and extending prison terms for high-rate robbers by a percentage of their current average terms. P. GREENWOOD, *supra* note 4, at 73-74, 78-85. As these latter offenders' current average prison term is over 4 years in California, *id.* at 77, this last option would mean that the sanction for high-rate robbers would be as much as *five times or more* as severe as that for lower-rate robbery convicts.

84. It is the last most drastic option described in note 83 *supra* that, according to Greenwood, would yield a substantial decrease in robberies without any increase in the incarcerated population. *Id.* at 79.

85. *Id.* at 92; Feinberg, *Selective Incapacitation and the Effort to Improve the Fairness of Existing Sentencing Practices*, 12 N.Y.U. REV. L. & SOC. CHANGE 53 (1983). For fuller discussion, see Note, *Selective Incapacitation: Reducing Crime Through Predictions of Recidivism*, 96 HARV. L. REV. 511 (1982). The note contends that selective incapacitation presents no significant ethical problems because, chiefly, prediction has been a traditional ground for sentence in this country; the discussion of the false-positive and desert issue is superficial, at best.

ing guidelines in jurisdictions such as Minnesota⁸⁶ and Washington⁸⁷ explicitly restrict reliance on prediction. Throughout his report, Greenwood argues that such restrictions are not the appropriate direction for reform: that forecasts of future criminality should be emphasized rather than suppressed.⁸⁸ He thus is aware himself that real disagreement exists today about the role of prediction in sentencing. Given such disagreement, it surely is misleading to say that the sentencing system inherently emphasizes selective incapacitation anyway, and hence that the false-positive issue need not be worried about.

B. *The Issue of Undeserved Punishment*

A more fundamental objection to predictive sentencing, one that could not be solved by increasing the accuracy of the forecasts, is that of undeserved punishment. The objection may be stated as follows. Punishment's distinguishing feature, that which distinguishes it from civil commitment, taxation and other impositions, is that it condemns: it treats the act as reprehensible and the actor as someone to be blamed. Punishment, as a blaming institution, is warranted only for past culpable choices, and cannot justly be allocated on the basis of predictions of future criminal conduct. Unless the person actually makes the predicted choice, he cannot be condemned for it.⁸⁹

86. For an analysis of the Minnesota guidelines and their rationale, von Hirsch, *Constructing Guidelines for Sentencing*, *supra* note 3.

87. The new Washington sentencing guidelines, which were developed by the state's sentencing commission in 1982 and approved by the state legislature this year, are set forth in 1983 Wash. Laws ch. 115. The enabling legislation, as amended this year, is set forth in 1983 Wash. Laws ch. 163. The guidelines are somewhat analogous to Minnesota's, and place primary emphasis on the seriousness of offenders' criminal conduct.

88. P. GREENWOOD, *supra* note 4, at 2-4, and particularly at 5 where he states:

The lack of evidence on the effects of either rehabilitation or deterrence leaves incapacitation as the only utilitarian basis for rationalizing differences in sentence severity for different types of offenders. If we eliminate retribution as a consideration, the only reason for varying sentence lengths among different types of offenders is the predicted future risk of these offenders to society.

When addressing legislatures, Greenwood is even more unequivocal in his advocacy of prediction and rejection of desert as the principal rationale for sentencing decisions. Thus in an October 1981 statement for the Subcommittee on Juvenile Justice of the U.S. Senate Judiciary Committee, he asserts:

After reviewing the literature on prevention, rehabilitation, deterrence and incapacitation—the only four crime reduction mechanisms available to government—it is clear that only incapacitation theory provides reasonable grounds for determining the relative severity with which different convicted offenders should be sentenced. The only other basis for distinguishing among offenders is deserved punishment or vengeance and here anyone's values are as good as anyone else's.

Hearings on S. 1688 etc. Before the Subcomm. on Juvenile Justice, Comm. on the Judiciary, 97th Cong., 1st Sess. 106 (1981) (statement of Dr. Peter Greenwood).

89. DOING JUSTICE, *supra* note 78, 66-76, 124-25.

This problem of disregarding the reprobative overtones of punishment is striking in Greenwood's scheme, because his proposed prediction scale is based on matters so alien to the blameworthiness of the criminal conduct for which the convicted offender ostensibly is being punished. Factors that relate to the current crime's gravity, such as use of a weapon, and the degree of cruelty with which the offense was committed, appear to have little or no predictive usefulness.⁹⁰ Four of the seven factors in Greenwood's index concern the offender's criminal record, but are by no means direct measures of the extent and seriousness of the record.⁹¹ The other three factors do not concern victimizing criminal conduct at all: they relate to drug use and employment.⁹² The scale, in the clearest way, bears little relationship to the gravity of what the offender has done.⁹³

90. J. AND M. CHAIKEN, *supra* note 4, at 108.

91. These are the first four of Greenwood's seven factors listed in note 33 *supra*. These four factors have little to do with blameworthiness—even with the gravity of the offender's past criminal record. The first listed factor (prior conviction for the instant offense type) means the convicted robber loses a point if he has a prior robbery conviction, but not if he has a prior conviction for a worse offense of a different type. The third factor (conviction before the age of sixteen) means an offender with a conviction at an early age of a minor offense fares worse than an offender with a longer record of worse offenses where those convictions happened to occur later. The second and fourth of the listed factors (incarceration during more than half of the preceding two years, time served in a juvenile facility) depend on the disposition chosen by sentencing judges, rather than on the gravity of the prior offenses themselves. For fuller discussion of the relevance of prior criminal record to judgments of blameworthiness, see von Hirsch, *Desert and Previous Convictions in Sentencing*, 65 MINN. L. REV. 591 (1981).

92. It might to some extent be possible to substitute other predictive factors that are less obviously concerned with status and more with past choices. But even that would not solve the problem, for reasons discussed in von Hirsch, *Utilitarian Sentencing Resuscitated: The American Bar Association's Second Report on Criminal Sentencing*, 33 RUTGERS L. REV. 772, 779-83 (1981).

93. It is sometimes claimed that selective incapacitation strategies, insofar as they rely on offenders' prior criminal records, would satisfy desert requirements—because desert theory also considers prior criminality. Feinberg, *supra* note 85, at 55 thus asserts:

[T]o the extent that a policy of selective incapacitation relies exclusively on evidence of the prior criminal history of the offender in predicting future dangerousness, it can be justified independently in terms of "just deserts" i.e., since the truly high-risk offender has a more extensive previous criminal track record, he "deserves" more punishment.

This argument overlooks the fact that a desert rationale, if it considers prior criminality at all, does so in a manner that is strikingly different than predictive strategies would. von Hirsch, *Desert and Previous Convictions*, *supra* note 91, at 617-629. Thus: (1) A desert rationale relies principally on the seriousness of the current offense, and gives only limited weight, if any, to prior criminality; a predictive scheme relies chiefly or heavily on prior record. *Id.* at 621-29. (2) A desert rationale may utilize only those features of prior criminal record that bear on the blameworthiness of prior choices; a prediction scale may utilize features of the prior record that either have nothing to do with prior crimes' blameworthiness or bear an inverse relationship to such blameworthiness—e.g., the defendant's age at first conviction. *Id.*, at 620-21. (3) A desert rationale, with its emphasis on condemnation for adjudicated crimes, could consider only prior convictions. *Id.* at 607-13. A prediction scheme

To speak of blameworthiness raises, of course, the role of desert in sentencing. That has been a controversial subject, and one whose ramifications have been addressed at length elsewhere.⁹⁴ Suffice it to say here that there now seem to be three main views regarding the extent to which desert considerations should influence decisions about the comparative severities of punishment. Let us summarize these three models, and see how selective incapacitation might fare under each.

(1) *Under a Desert Model.* A desert model is a scheme that attempts rigorously to observe the requirements of proportionality in deciding the relative severity of punishments. This means that equally blameworthy conduct should be visited by equally severe penalties, and that penalties should be graded in severity to reflect the seriousness of the criminal behavior involved.⁹⁵ This view would rule out selective incapacitation⁹⁶

(2) *Under a Modified Desert Model.* A modified desert model is one in which desert would continue to play the central role in deciding relative severities of punishment. The constraints of equal treatment for the equally deserving would be relaxed, however, to allow a *modest* scope for utilitarian concerns. This would allow variations in the punishment of equally deserving offenders on predictive grounds, but those variations would have to be

could consider prior arrests and prior misconduct not leading to conviction. In fact, Feinberg emphasizes that selective incapacitation schemes should not rely only on prior convictions but arrests as well; he asserts:

Once again, the federal proposals are very narrowly drawn. Indeed, they are too restrictive. Federal proposals limit the policy of selective incapacitation to prior violent criminal activity as demonstrated by one or more convictions (no distinction is made between adult and juvenile convictions). Reliance solely on convictions poses difficulties, since convictions notoriously underrepresent the volume of reported crime. As a result, the practical value of a selective incapacitation policy is severely undercut if it is based solely on convictions. Reliance on past arrests, particularly juvenile arrests for violent crime, would seem to provide a stronger indicator of criminal potential.

Feinberg, *supra* note 85, at 57. For further discussion of these issues, see von Hirsch, *The Ethics of Selective Incapacitation: Observations on the Contemporary Debate*, 30 *CRIME & DELINQ.* No. 2 (1984) (forthcoming).

94. For diverse views on this subject, see *SENTENCING*, *supra* note 17, at 240-301; *DOING JUSTICE*, *supra* note 78.

95. This view requires one to hold that desert determines the *comparative* severities of punishments, but not necessarily the anchoring points of the penalty scale. For articulation of this distinction—that desert is a determinative principle in deciding the internal ordering of a penalty scale, but only limiting in deciding the scale's cardinal magnitude, see, von Hirsch, *Recent Trends in American Sentencing Theory*, *supra* note 1, at 28-9 and in more detail, von Hirsch, *Commensurability and Crime Prevention: Evaluating Formal Sentencing Structures and their Rationale*, 74 *J. CRIM. L. & CRIMINOLOGY* 209, 214-26 (1983); and von Hirsch, *Equality, "Anisonomy" and Justice: An Analysis of Norval Morris' Madness and the Criminal Law*, 82 *MICH. L. REV.* No. 4 (1984) (forthcoming).

96. See A. VON HIRSCH AND K. HANRAHAN, *supra* note 82, at 17-18.

kept in narrow or moderate limits.⁹⁷ Greenwood's approach would present problems even under such a model—because the options of which he is speaking (and especially those which he claims will have large crime-reduction effects) involve such striking differences in the sentences given to high and low risk robbers.⁹⁸

(3) *Under a Neopositivist Model.* Here, utilitarian aims dominate, and desert is reduced to a subordinate or even marginal role. Incapacitation (and perhaps other crime control concerns) would chiefly determine the choice of sanction, even when doing so would result in substantial inequalities in the punishment of equally blameworthy offenders. There might continue to be limits on grossly disproportionate sanctions, but desert would only provide the extreme upper (and perhaps lower) bounds on the quantum of the sentence.⁹⁹ Such a model would provide ample scope for the kind of options of which Greenwood is speaking: it would become permissible to imprison high-risk offenders for substantial terms while giving short jail stints (or less) for the low-risk offenders. It is this model, unfortunately, which is to our judgment morally the most troublesome for it distributes the blamelevying criminal sanction in a manner that largely disregards its reprobative implications.¹⁰⁰

We thus see the dilemma posed by selective incapacitation proposals. The larger the influence prediction has on the choice of sentence, the more aggravated become the problems of overprediction and undeserved punishment. On the other hand, the more the influence of prediction is scaled back in order to avoid these problems, the less significant the incapacitative effects can be even on the most optimistic calculations.

V

RESOURCE ISSUES: PRISON OVERCROWDING AND THE CHOICE OF SENTENCING RATIONALE

Greenwood asserts that a selective-incapacitation strategy is uniquely useful in solving the problem of prison overcrowding. He supplies various

97. Such a model is suggested in Monahan, *The Case for Prediction in a Modified Desert Model*, 5 INT'L. J.L. & PSYCHOLOGY 108 (1982).

98. See *supra* text accompanying notes 83-84.

99. Such a model has been proposed in a 1979 ABA Task Force on sentencing, AMERICAN BAR ASSOCIATION, TASK FORCE ON SENTENCING ALTERNATIVES AND PROCEDURES, SENTENCING ALTERNATIVES AND PROCEDURES III ABA Standards for Crim. Just. 18.8 (1980).

100. For fuller discussion, von Hirsch, *Utilitarian Sentencing Resuscitated*, *supra* note 92. It is interesting to note that the chief defender of making desert only a "limiting" principle, Norval Morris, has recently distanced himself from the ABA proposals, on the grounds that they give desert too marginal a role in sentencing decisions. N. MORRIS, MADNESS AND THE CRIMINAL LAW (1982) 202-09. For comments on Morris, see von Hirsch, *Equality, "Anisonomy" and Justice*, *supra* note 95.

calculations purporting to show that a predictive sentencing policy will best conserve prison resources.¹⁰¹ We do not think he has proven his case.

A. *Need for Selectiveness*

Greenwood begins with a true statement—indeed, a truism: to remain within the available prison resources, one must develop some policy of selection among convicted offenders. The unrestricted imprisonment of felons will simply overwhelm the prisons. It is therefore essential to supply a workable and systematically-applied standard for deciding which felony cases are important enough to warrant the prison sanction and which are of less importance. One problem with many purported sentencing reforms to date is that they have failed to supply and to apply any such standard. An example is California. The greatest deficiency of the California determinate sentencing statute is that it furnishes no meaningful standards to decide whether or not a felon should be sent to prison; it merely regulates the duration of confinement if the judge exercises his discretion to imprison. With no selection standard for the “in-out” decision, it is not surprising that there has been an uncontrolled influx into California’s already-overcrowded state institutions.¹⁰²

To say one needs to be selective on some basis does not, however, settle the important issue: deciding the criteria for selection. In terms of today’s sentencing debate, there are two major alternatives. One alternative would be to adopt a criterion of desert (in either pure or modified form). Then, the people who should be imprisoned would principally be those convicted of serious crimes—that is, crimes involving grave harm and a high degree of culpability. The other alternative is to adopt criteria emphasizing prediction. In that event, the determinant for going to prison would chiefly be the gravity and frequency of the felon’s expected future violations. It is far from obvious that the latter, predictive approach is a better way of conserving prison space than the former, more desert-oriented approach.

B. *Choosing the Criteria for Selection: The Minnesota Experience*

An illuminating comparison between desert-oriented and predictively-oriented criteria, in terms of their impact on prison resources, has already been made—in the construction of Minnesota’s sentencing guidelines.¹⁰³

101. P. GREENWOOD, *supra* note 4, at 78-85.

102. von Hirsch & Hanrahan, *Determinate Penalty Systems in America*, 27 CRIME & DELINQ. 289, 302 (1981); for fuller analysis of California’s law, von Hirsch and Mueller, *California’s Determinate Sentence Law: An Analysis of Its Structure* 10 NEW ENGLAND J. ON CRIM. & CIVIL CONFINEMENT No. 2.(1984) (forthcoming).

103. See von Hirsch, *Constructing Guidelines for Sentencing*, *supra* note 3.

That state's sentencing commission decided at the outset that the guidelines should fully take into consideration the availability of prison resources.¹⁰⁴ The Commission then developed a sentencing grid in which the seriousness of the current crime was the vertical axis, and the extent of the prior criminal record was the horizontal axis. The major policy decision shaping the guidelines became the determination of the slope of the dispositional (or "in-out") line—that is, the line on the grid separating prison from non-prison dispositions.¹⁰⁵

In making this decision on the slope of the dispositional line, the Commission proceeded to compare two major sentencing philosophies. One alternative would be a more desert-oriented line, which would be relatively flat: the primary emphasis would be given to the seriousness of the current crime. The other alternative would be what the Commission characterized as an incapacitation line: it would be much steeper, and give primary emphasis to the extent of the criminal record.¹⁰⁶ (The Commission was treating the criminal record as an approximate indicator of risk on the grounds that numerous prediction studies have shown the record to have a modicum of predictive usefulness. One could, however, approximate a selective-incapacitation scheme more closely by changing the horizontal axis of the grid so that it embodies a formal prediction score instead of only the criminal record. In that event, an incapacitative "in-out" line would still be a steep one—one in which the offender's risk score, rather than the seriousness of his offense, would principally determine whether or not he is to be imprisoned.¹⁰⁷) Ultimately, the Commission decided to adopt a line which it characterized as a "modified" desert line, and which chiefly emphasized the gravity of the offense.¹⁰⁸ The basis of the decision was a philosophical preference for a sentencing policy that stressed the blameworthiness of criminal conduct.¹⁰⁹

One can agree or disagree with the Commission's decisions on grounds of principle. But the Commission's decision to adopt a flatter dispositional line instead of a steeper one did not require increased prison space. The flatter line meant that Minnesota imprisons a larger proportion of defendants convicted of serious crimes, but compensates by imprisoning relatively few of those convicted of intermediate and lesser crimes. Had the Commission adopted instead the steeper line reflecting a predictive rationale, this would have reallocated the use of imprisonment. Such a scheme, by focusing on the risk of future criminality, would have allowed one to be more

104. *Id.* at 176-180.

105. *Id.* at 181.

106. *Id.* at 180-91.

107. *Id.* at 183-85.

108. *Id.* at 181-82.

109. *Id.*; *see also id.* at 182-91.

selective in the use of imprisonment among offenders convicted of robberies and other serious crimes. However, it would also have required one to imprison a larger proportion of offenders convicted of intermediate-level crimes where the length of their criminal record or their prediction scores indicated they were bad risks.¹¹⁰ Either approach would, in theory, have involved approximately the same aggregate prison resources.

C. The "Political Pressures" Argument

If an incapacitative rationale is not intrinsically the more parsimonious in its use of prisons, that rationale's advocates are left with the argument that a selective-incapacitation approach is somehow less vulnerable to political pressures for escalating punishments than are other approaches. We find this argument unconvincing.

A desert model, as articulated in *Doing Justice*¹¹¹ and other recent writings,¹¹² would limit the use of imprisonment to specified serious crimes. In Minnesota's guidelines, this conception has been implemented (with modifications) as a policy that restricts state imprisonment chiefly to crimes involving threatened violence such as robbery and worse offenses and that uses non-prison sanctions for intermediate-level offenses such as burglary (except where the offender's criminal record is quite lengthy).¹¹³

As a practical matter, can the line be held at this level? Given pervasive public fears of crime and the political benefits of taking tough anti-crime postures, will there not be strong political pressures to make the scheme much more inclusive: to imprison the burglars as well as the robbers? Certainly such pressures exist, and resisting them will be no easy matter. The most useful safeguards of which we know are those which Minnesota has developed, to wit: (1) have the sentencing standards written by an indepen-

110. Greenwood disregards this important last point in his calculations of prison-population impact. He takes his sample of convicted robbers, and compares the impact of these two alternatives: (1) a policy of imprisoning or presumptively imprisoning all such robbers, and (2) a policy (which he identifies as predictive) of imprisoning only those robbers whom he identifies as high risks. It is scarcely surprising that the latter policy uses less prison space, because it imprisons only some, not all, convicted robbers. But this comparison is misleading. A predictive standard aimed at restraining potential high-risk offenders would not necessarily limit imprisonment only to a selected subgroup of those convicted of such serious crimes as robbery. It would also include as candidates for imprisonment persons convicted of lesser crimes (e.g., burglary) to the extent that they are high risks (that is, to the extent one can predict either that they are likely to commit robbery or other serious offenses in the future, or are likely to commit numerous intermediate-level offenses, or both). Once one begins including these latter offenders in the prison-bound population, the prisons will begin to fill.

111. *DOING JUSTICE*, *supra* note 78.

112. These views are summarized in von Hirsch, *Recent Trends in American Criminal Sentencing Theory*, *supra* note 1, at 29-30.

113. von Hirsch, *Constructing Guidelines for Sentencing*, *supra* note 3, at 181-91.

dent rulemaker which is less vulnerable to law-and-order politics than the legislature, and (2) tie the standard-setting process closely to the availability of prison resources, so that punishment levels may be increased only if the necessary sums are appropriated to build more prisons.¹¹⁴ Escalation in punishment levels has been most in evidence in states where these safeguards have been absent: where, as in California, the legislature sets durations of imprisonment and is not required to give any consideration to the availability of prison space.¹¹⁵ But even in a jurisdiction such as Minnesota, demands for increased punishment will not easily be resisted; only time will tell whether that state will be able to hold the line at or near present levels.

This problem of pressures to escalate would by no means be avoided, however, were one to make the dispositional line steeper—that is, were one to shift toward an incapacitative rationale for the system. Greenwood's calculations about prison space are made on the assumption that the resource of imprisonment is to be used chiefly to isolate those predicted to commit frequent robberies or other serious crimes in the future.¹¹⁶ This would mean that potential medium-rate robbers, and potential high-rate burglars would be imprisoned only for short periods, or not at all. Such narrowly-drawn incapacitative policies are hardly likely to satisfy law-and-order constituencies for long.

It is our opinion that an incapacitative approach is, in fact, still more vulnerable to escalation than a desert-oriented system would be. The latter, at least, makes no promises to reduce crime.¹¹⁷ A system giving preeminence to incapacitation, by contrast, promises to prevent crime by isolating dangerous people. Every time such a system “misses” (i.e., fails to imprison) offenders who subsequently prove dangerous, demands will be heard to make the definition more inclusive. It should be recalled that Greenwood's prediction score shows a false-negative rate of about sixteen percent.¹¹⁸ This means one out of every six convicted robbers who are classified as medium or low risks will, in fact, commit frequent robberies. (The number of these false negatives thus may be nearly two-thirds the number of persons classified as high risks!) The high incidence of “misses,” as it becomes known, is likely to intensify demands to widen the net—and hence to increase the numbers imprisoned or the duration of imprisonments. A legislature or sentencing commission, if it opts for a selective incapacitation approach is

114. *Id.* at 168-71, 176-80.

115. von Hirsch and Hanrahan, *Determinate Penalty Systems in America*, *supra* note 102, at 299-303; von Hirsch and Mueller, *California's Determinate Sentencing Law*, *supra* note 102.

116. P. GREENWOOD, *supra* note 4, at 78-85.

117. For elaboration of this point, von Hirsch, *Constructing Sentencing Guidelines*, *supra* note 3, at 186-89.

118. See text accompanying note 37 *supra*.

likely to face the same problem that parole boards now face when they release prisoners who subsequently commit violent acts: They will be held responsible for failing to perform their self-proclaimed function of protecting the public from recidivists—and efforts will be made to mandate imprisonment, or longer prison terms, for an enlarged class of offenders.¹¹⁰

VI

CONCLUDING OBSERVATIONS

A. *The Greenwood Study*

Our examination of Peter Greenwood's study suggests it represents very much less than a breakthrough in the predictive art. Thus:

a. *Predictive Accuracy.* The Greenwood study raises the issue of forecasting the rate at which offenders will commit crimes in the future, whereas most earlier prediction studies had looked only to whether or not the offender committed any new crimes. To try to forecast the rate of future criminality, the Greenwood study was forced to resort to self-report techniques that may not be particularly reliable for the purpose. The result of the research is a "predictive" instrument that makes use of much of the same factors that have been used in prediction studies for years: the criminal record, employment history, and drug use. Greenwood's instrument shows (or might show when replicated) some success in its forecasts: that is, it performs better than flipping a coin would. This is scarcely surprising, however, since prediction methods have long had such limited achievements. The Greenwood study has not shown much progress in the strength of the association of its predictors with offenders' criminal behavior. And the rate of false positives and false negatives remains distressingly high.

b. *Purported Impact on Crime.* The Greenwood study promises substantial reductions in serious crime. A selective incapacitation strategy, allegedly, could produce as much as a fifteen percent decrease in the number

119. Another parallel comes to mind: the civil commitment of persons acquitted on grounds of insanity. Because such persons are deemed incapable of choice, the only basis for deciding commitments is the individual's dangerousness.

However one defines the degree of risk warranting commitment, there tend to be demands to make that definition more inclusive. This is because a predictive system conceals erroneous releases. Whereas the false-positive has no opportunity to show he would not have caused the injury if released, the false-negative remains at large, coming to public attention when he injures someone again. The public holds officials responsible for mistaken releases. This prompts efforts to make it more difficult to release insanity-defense acqutees from civil commitment.

What proponents of selective incapacitation suggest is that the sentencing system be modelled more closely on the existing system for confining those acquitted for insanity. Whatever possible arguments there might be in favor of this position, the political argument seems least persuasive: in the present atmosphere, this civil commitment system is exposed to particularly intense pressures for a toughening of standards.

of robbery offenses. It is this claim of being able to reduce seemingly intractable crime rates that has attracted so much attention for the study.

We find Greenwood's projections of crime-preventative effect to be founded on a flawed research design. The calculations are based on imputed offense rates for all active robbers. Greenwood has no evidence of these rates, and he tries inappropriately to extrapolate them from data about a limited and unrepresentative sample, to wit, a sample of robbers who have been incarcerated. The calculations fail to consider other factors that are likely to diminish the preventative impact, such as replacement effects, and the variability of robbery rates over time. It also would be extremely difficult to confirm such projections were a selective incapacitation strategy actually implemented because of problems of controlling for other possible influences on crime rates. We believe that no credence should be given to the study's assertions about reducing the aggregate incidence of robberies.

One reply to these objections has been that although the available evidence does not fully support the conclusions about reduced crime-rates, it is the only evidence now available. We cannot accept this argument, for we do not subscribe to the view that bad evidence supplies better support for a conclusion than no evidence at all. Were one so tolerant, other now discredited crime control claims would still be believed. After all, most rehabilitation studies did show that the treatment program was associated with lower recidivism rates than the comparison group; it was only after one looked at the studies more closely, to control for differences in risk, that these effects tended to disappear. Many deterrence studies did show that increased punishment was associated with decreased crime; it was only after one considered the possible influence of other variables, and the question of the direction of the causality, that the findings tended to wither. Until there is good evidence that selective incapacitation reduces crime, one has no right to assume the strategy works.

c. *Fairness Issues.* Selective incapacitation, in the form considered in the Greenwood study, in no way alleviates the problems of overprediction and of undeserved punishment that historically have been associated with predictive sentencing. On the contrary, these ethical difficulties are present in aggravated form—because the rate of false positives is so high, because the factors relied upon as predictors have so little to do with the blameworthiness of the offender's present or past criminal conduct, and because the difference in punishment between high-rate and low-rate offenders may be substantial.

d. *Effects on Prison Populations.* Greenwood asserts that selective-incapacitation strategies have special usefulness in alleviating prison overcrowding. Analysis does not bear this claim out. It is true, obviously, that a sentencing scheme that is selective in sending felons to prison will require fewer confinement resources than a scheme that imprisons all or most felons. However, Greenwood has failed to establish that prediction is, in its impact on prison resources, a superior criterion for selection than alterna-

tive criteria such as desert. And we fear that a predictive sentencing scheme may be particularly vulnerable to political pressures for escalating punishments.

It is our conclusion that Greenwood's study does not make the case for prediction in sentencing any stronger than it was before.

We are troubled, also, at the degree of overstatement found in Dr. Greenwood's report. While the supposed benefits of selective incapacitation are described very positively, none of the problems just mentioned are discussed seriously. Claims such as that of being able to reduce robberies by fifteen percent are stated with few if any qualifications: such claims are uncomfortably reminiscent of James Q. Wilson's earlier promise of being able to reduce robberies twenty percent through collective incapacitation,¹²⁰ and of Isaac Ehrlich's now deflated assertion that every execution will prevent eight murders.¹²¹

The other RAND study, by the Chaikens,¹²² utilizes the same self-report data as Greenwood, but has the virtue of greater caution. The Chaikens expressly state that theirs is "*postdiction*" research, not a true prediction study. They offer no claimed solutions to high crime rates, or to prison overcrowding. The Chaikens' predictive (or rather, postdictive) index accounts likewise for a small portion of the variability in behavior among the offenders studied. The false-positive rate remains extremely high: over sixty percent of those classified as high-rate robbers in fact had lower rates or did not commit robberies at all.¹²³ But the Chaikens explicitly call attention to the high incidence of false positives, and warn that this problem could be further exacerbated were official records rather than self-reports relied upon.¹²⁴

B. The State of Knowledge about Prediction

What do we really know about prediction? Essentially, what has been known for many years. Using a few familiar factors, one can spot groups of convicted offenders that are more likely than other groups to return to crime. Those forecasts, however, will use prediction variables that are rather

120. See text accompanying note 41 *supra*. Wilson now supports selective incapacitation instead, with much the same degree of optimism. See Wilson, *Dealing with the High-Rate Offender*, THE PUBLIC INTEREST, Summer 1983, 52, 59-64.

121. Ehrlich, *The Deterrent Effect of Capital Punishment: A Question of Life and Death*, 65 AMER. ECON. REV. 397, 398 (1975). For a critique of Ehrlich's claims, see Panel on Research on Deterrent and Incapacitative Effects, *supra* note 48 at 59-63; Klein, Forst and Filator, *The Deterrent Effect of Capital Punishment* in DETERRENCE AND INCAPACITATION, *supra* note 17, at 336-60.

122. J. AND M. CHAIKEN, *supra* note 4.

123. *Id.* at 179-80.

124. *Id.*

weakly associated with offenders' subsequent criminal behavior. The forecasting instruments will tend to show high rates of overprediction, especially when one tries to predict statistically rare behavior such as violent conduct.

A predictive strategy for sentencing thus can have limited, but not more than limited, effectiveness in isolating offenders who are risks. It does so at the expense of disregarding the condemnatory overtones of punishment, and of confining false positives. The more heavily prediction is relied upon, the more aggravated these problems will be. Deciding the role of prediction in sentencing is in the last instance a choice of values, not science.

We doubt, however, that crime rates can substantially be reduced through one or another sentencing strategy—especially any strategy that could realistically be employed with the limited resources available in most jurisdictions. Deterrence and collective-incapacitation strategies were once offered as solutions to crime, but now are perceived in our profession with justified skepticism. Selective incapacitation is not likely to make the streets safer when these other strategies did not.

Appendix

PURPORTED EFFECTS OF SELECTIVE INCAPACITATION ON AGGREGATE ROBBERIES

1. *The Shinnar Formula for Collective Incapacitation.* Greenwood's model is derived from Shinnar's formula for estimating collective-incapacitation effects.¹²⁵ That formula works as follows. Suppose all offenders commit a given type of crime at the same average rate, λ . Suppose q is the probability of arrest and conviction for that crime; J is the probability of incarceration given conviction; and S is the average duration of incapacitation. The critical step in the Shinnar model is, then, the estimation of η , which is the average fraction of time offenders spend on the street. According to Shinnar,

$$\eta = \frac{1}{1 + \lambda q J S} \quad (1)^{126}$$

Naturally, this estimate of the average proportion of time offenders are at liberty is valid only if the estimate for λ is a true average offense rate—one that holds for the currently unincarcerated as well as incarcerated offenders.

Given formula (1), the next steps are easy. If one knows the average proportion of time offenders spend on the street, then one can estimate the total number of offenders from the number incarcerated. Thus if R is the number of offenders incarcerated, then the total number of offenders (N) is:

$$N = \frac{R}{(1 - \eta)} \quad (2)$$

Next, one can estimate (C), the incidence of crime. It is the number of offenders times the fraction of time they spend on the street times their average individual crime rate, or:

$$C = N\eta\lambda \quad (3)$$

According to Shinnar, the way to ascertain the incapacitative effect of a given sentencing policy is to determine how much C changes when one either alters the probability of incarceration given conviction (J) or the duration of confinement (S).¹²⁷

2. *Greenwood's Modification for Selective Incapacitation.* All that Greenwood has done, in order to estimate *selective* incapacitation effects, is to assign three λ 's—one each for high, medium and low rate robbers.¹²⁸

125. Shinnar and Shinnar, *supra* note 42. For an explanation and analysis of the Shinnar formula, see Cohen, *supra* note 47, at 196-98.

126. Cohen, *supra* note 47 at 196.

127. *See id.*

128. P. GREENWOOD, *supra* note 4, at 74-8.

With these three Lambda's, he assigns different likelihoods of imprisonment (qJp) and different sentence lengths (S) to those groups.¹²⁹ Using the analogue of formula (1), he then gets three η 's, one for each group; and three N 's also, one for each group. The incidence of robbery (C) associated with the policy is:

$$C = \sum N_i \eta_i \lambda_i \quad (4)^{130}$$

3. *The Problem.* The problem is that Greenwood derives his estimate of Lambda from the robbery rates of *incarcerated* individuals, and has no evidence on the robbery rates of those not incarcerated.

The same problem would arise under the original Shinnar formula, were one trying to estimate a single average Lambda from data about incarcerated individuals. As we saw, the critical formula (1) is valid only if the λ is the average rate for *all* robbers, including those on the street. But Greenwood has no estimate of the true λ . He only has a basis for estimating λ' —the average past individual robbery rate for now-incarcerated offenders. With only that figure one cannot estimate the true average fraction of time offenders spend on the street—since that depends also on an unincarcerated robbers' average robbery rate (λ''), which may well be different.¹³¹ Thus the formula

$$\eta = \frac{1}{1 + \lambda' qJS} \quad (5)$$

129. Greenwood computes η_i , the average fraction of time an offender in a given risk category spends on the street:

$$\eta_i = \frac{1}{1 + \lambda_i qJ((1-p_i)s_i + p_i S_i)}$$

Id. at 75.

Since a selective-incapacitation policy may send some robbers to prison and others to jail, one needs a new variable (p) which describes the probability of imprisonment given incarceration.

The incapacitative effect of sending offenders to jail rather than to prison is integrated into the formula through $((1-p_i)s_i)$. The variable s_i , or the average length of a jail term is assumed to be one year. The variable S_i represents the duration of imprisonment of imprisoned offenders in a given risk category. *Id.* at 77.

130. P. GREENWOOD, *supra* note 4, at 75.

131. This can simply be illustrated as follows. Suppose one were trying to estimate the Ω , average fraction of working time the authors of this article spend in the office on a particular day. The interviewer finds Gottfredson in his office and is informed by him that he spends 75 percent of his working time there; he does not interview von Hirsch who is absent. The information the interviewer has is insufficient to estimate Ω , because the true average depends also on von Hirsch's habits and he is less readily interviewed, precisely because he spends more of his time working at home. Until the interviewer has opportunity to review both Gottfredson and von Hirsch, he can have no accurate estimate of Ω .

is false. Instead, the estimate of η would require knowledge of the unincarcerated offenders' average robbery rate as well as the incarcerated offenders' rate.

$$\eta = f(\lambda', \lambda'') \quad (6)$$

On a research design based on inmate self-reports, there can be no data about what λ'' actually is, so the estimate of η cannot be accurately made. Without this latter estimate, one cannot take the next steps of estimating N and C in formulae (2) and (3).

Shifting from this simple collective-incapacitation model to Greenwood's selective-incapacitation model does not solve the problem. Greenwood can provide no true estimates of the Lambdas for the high, medium, and low rate offenders; he can only estimate λ' for the three groups of incarcerated offenders. With only the three λ' figures, he cannot estimate the three η figures—the average fraction of time that high, medium and low risk robbers are at liberty. That means he cannot accurately estimate the incidence-of-robbery figures as per formula (4).

(4) *Making a More Conservative Estimate.* How might a more conservative estimate be made? It could be done by considering only those offenders whose robbery habits can be studied—that is, those in confinement when the self-report studies are done. Once these persons are released, they may—depending on their Lambdas—commit further crimes. One could attempt a calculation based on the activities of these robbers, after their release from their current confinement. The calculations would proceed as follows:

a. *Collective Incapacitation Effects.* To calculate the collective-incapacitation effect, one would assume λ' to be the average self-reported robbery rate.¹³² As the activities of now-unincarcerated robbers are not being estimated, there would be no need to take their average robbery rates (λ'') into consideration. The total number of potential offenders (N') would, *ex hypothesi*, be the number of robbers now in confinement.

Some of these offenders will be released to the street and then perhaps reimprisoned for the new crimes. Then:

$$\eta' = \frac{1}{1 + \lambda' qJpS} \quad (7)$$

would hold—where η' is the fraction of time that the group of currently-incarcerated offenders may (at a stated subsequent period) be expected to be at liberty. It follows the incidence of robbery for a given sentencing policy would then be:

132. This, of course, does not take into account the possibility, indeed the likelihood, that these offenders' robbery activities will decline in future. See *infra* text accompanying note 139.

$$C' = N'\eta'\lambda' \quad (8)$$

To measure the number of robberies prevented by a change in policy, one would examine the absolute number of crimes by which C' decreases when the new policy is implemented. This decrease should be much smaller than that hypothesized by Greenwood using his model.

b. *Selective Incapacitation Effects.* To measure selective-incapacitative effects, the formula should be revised as follows:

(1) The η'_i for each of the three risk groups would be the same as in Greenwood's formula,¹³³ using the self-report λ'_i figures for the three groups.

(2) The number of offenders (N'_i) for each of the three groups would be the numbers of now-incarcerated offenders in each group.

(3) The expected incidence of robbery for these offenders after their current confinement (C') is then calculated as:

$$C' = \Sigma N'_i \eta'_i \lambda'_i \quad (9)$$

5. *The Difference in the Two Estimates for Jurisdictions X.* Suppose that we have the following information for Jurisdiction X.

	Low	Medium	High	Total
R Incarcerated ¹³⁴ Populations	1290	645	645	2580
λ' Average self-reported individual robbery rates	3	9	20	
Jq Probability of arrest, conviction, and incarceration	.025	.025	.025	
p Probability of prison ¹³⁵ given incarceration	.12	.35	.57	

133. P. GREENWOOD, *supra* note 4, at 75. The formula is set forth in note 129 *supra*.

134. Notice this assumes that about a quarter of imprisoned offenders are high-scorers, as P. Greenwood initially said would be the case. If many more are high-scorers, this will undermine his claim that selective-incapacitation can focus on a limited number of high rate individuals. It is interesting to note that in Greenwood's calculations for California at the end of his report (*id.* at 77), a much larger fraction of confined offenders (nearly half) are high scorers!

135. These values for probability of imprisonment are the same as those used by Greenwood in his calculations for California. *Id.* at 77.

S			
Sentence length in prison ¹³⁶	2.1	2.5	3.0
s			
Sentence length in jail	1.0	1.0	1.0

Using Greenwood's model, we can, given the information, calculate η_i , the fraction of time members of each group are free. Then, N_i , the number of offenders in each group, and C_i , the number of robberies committed by each group, under Greenwood's method of reckoning, can be computed:

	<i>Low</i>	<i>Medium</i>	<i>High</i>	<i>Total</i>
η_i	.92	.74	.48	
N_i	16,125	2,481	1,240	19,846
C_i	44,505	16,523	11,904	72,932

Notice that in each group the estimated number of offenders (N_i) is much larger than the number of incarcerated offenders (R_i).

Using the more conservative model—one which is concerned only with the impact of certain sentencing strategies on offenders who can be studied—both N'_i and C'_i would initially be calculated in a different manner. N'_i would equal, for each offense rate group, the number of persons in that category now in confinement. C'_i would be calculated using the new estimates of N'_i as follows:

	<i>Low</i>	<i>Medium</i>	<i>High</i>	<i>Total</i>
N'_i	1,290	645	645	2,580
C'_i	3,560	4,296	6,192	14,048

Suppose that we now adopt a selective incapacitation policy similar to Greenwood's "Option 5".¹³⁷ We can calculate the changes that would occur in C'_i using each of these models.

Following Greenwood's manner of reckoning, a selective incapacitation strategy would reduce the number of crimes prevented in the high-rate offender category by almost 6,000. Using our more "conservative" estimation model, however, the reduction in crimes for the high-rate group would be about *half* as large: 3,100.

6. *How Accurate is the "Conservative" Model?* The "conservative" model just described might either underestimate or overestimate the crime-prevention effect. Let us consider each possibility in turn.

136. These values for duration assume generally shorter terms than in California, but reflect the relatively modest differences between lower and higher risk defendants under current practice.

137. P. GREENWOOD, *supra* note 4, at 73.

a. *Underestimation.* The model considers only the subsequent criminal activities of offenders now incarcerated. While there may well be substantial numbers of now-unincarcerated offenders who also contribute to crime rates, Greenwood's research design provides no way of estimating reliably either their robbery numbers or their robbery rates. For the reasons outlined above, extrapolating those rates and numbers from the self-reported data of incarcerated offenders is fallacious.

Could one develop scenarios for the possible activities of the unincarcerated offenders, and incorporate those into the calculation? That could be mathematically possible, but the results would vary with how favorable or unfavorable the scenarios might be to Greenwood's thesis. Using scenarios such as those suggested above,¹³⁸ the preventive impact might not be much larger than it would be using the "conservative" model. But in any event, *we have no sound evidence favoring one scenario over another.* Until we have such evidence, the use of scenarios to predict the activities of unstudied robbers is little more than speculation.

b. *Overestimation.* It is more likely that the "conservative" model still *overstates* the impact of selective incapacitation. The model makes two crucial assumptions: (1) the now-incarcerated offenders will continue to rob at the same rate as (or a higher rate than) they reported in their self-reports, and (2) there are no replacement effects, so that confining a high-rate robber will eliminate a number of robberies equal to the offenses he is predicted to commit. For reasons discussed above,¹³⁹ both assumptions are probably false. Hence even the "conservative" calculation is not a reliable estimate of preventive impact.

138. See text accompanying notes 60-65 *supra*.

139. See text accompanying notes 71-74 *supra*.

