

Summer 1935

Prediction Methods Applied to Problems of Classification Within Institutions

George B. Vold

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>

 Part of the [Criminal Law Commons](#), [Criminology Commons](#), and the [Criminology and Criminal Justice Commons](#)

Recommended Citation

George B. Vold, Prediction Methods Applied to Problems of Classification Within Institutions, 26 *Am. Inst. Crim. L. & Criminology* 202 (1935-1936)

This Article is brought to you for free and open access by Northwestern University School of Law Scholarly Commons. It has been accepted for inclusion in *Journal of Criminal Law and Criminology* by an authorized editor of Northwestern University School of Law Scholarly Commons.

PREDICTION METHODS APPLIED TO PROBLEMS OF CLASSIFICATION WITHIN INSTITUTIONS

GEORGE B. VOLD¹

Two things will be attempted in this paper: (I) an outline of the problem involved in prediction of criminal behavior, and (II) an account of an application of prediction methods to the problem of classification within penal institutions.

I

The first developed plan for the application of statistical methods of prediction to the problem of the probable future conduct of criminals appears to have been that suggested by Dr. E. W. Burgess in 1927 in connection with a study of the indeterminate sentence in Illinois. Since then the work of the Gluecks, Vold, Tibbitts, Monachesi, and others, has indicated a number of variations in technique and suggested many problems for further consideration.

1. *The development of criteria of parolability.*

The data for parole predictions which have been made are usually those found in official records. There is an important field for experimentation in the effort to develop a more exact and scientific procedure in obtaining information which will throw light on probable future conduct.²

2. *Problems of reliability and follow-up.*

In his Illinois study, Professor Burgess used the title, "Factors original Illinois study, Professor Burgess used the title, "Factors Determining Success or Failure on Parole," Similar titles have been used by others, including the present writer. The words carry their own challenge. What factors are important in determining probable future conduct? How can one be sure they are important? These questions are of vital importance both to theory and application of prediction methods.

¹University of Minnesota, Minneapolis.

²An experiment of this nature is described in this Journal by Ferris F. Laune, on "A Technique for Developing Criteria of Parolability." See XXVI, 1, 41-45, May, 1935.

Professor Burgess relied upon inspection of the percentage distributions in relation to "violators" or "non-violators" of parole as they appeared for the various subclasses under the several categories or factors. As a result he constructed his prediction tables on the basis of twenty-one factors selected as important. Obviously, all of these factors were not of equal importance or significance as indicators of probable future conduct. It is possible that some were of little or no real importance and the prediction tables might have been equally satisfactory, or even more satisfactory, without their inclusion.

The Gluecks, on the other hand, attempted to determine the significance of their factors by statistical means, utilizing for this purpose the coefficient of mean square contingency. As a result they constructed their tables on a smaller group of factors, (from 6 to 13), presumably those determined to be significant by showing the highest contingency values in relation to post-parole criminality. Their method also makes use of a system of weighted scores for the subclasses of the selected factors while the Burgess method gave to each an equal value, namely, one point for every subclass having a rate "better than the average for the institution."

There can be no question of the desirability of utilizing only such factors in prediction tables as are clearly important as indicators of future conduct.^{2a} The problem is how to discover the importance or significance of a factor.

The present writer, working with Minnesota parole cases, made use of both the Burgess and the Glueck methods on the same group of cases. The results were not very different as indicated by the correlation between the two methods of $r = .92 \pm .006$. The opposite alternative from the Glueck method was also attempted in which only the factors having the lowest contingency values were utilized. The prediction tables thus devised were not as satisfactory (evident from inspection) as the tables resulting from the other methods. Nevertheless it was true that those factors which had contingency values of zero or practically zero (less than .10) gave a moderately satisfactory prediction table whether they were used with the Burgess' method or with that of the Gluecks.

These results tend to raise serious doubts as to the adequacy of the contingency method in revealing the importance or significance of factors. If factors showing low contingency values have considerable significance, it may likewise be true that factors having

^{2a}For contrary conclusions see:

"One Thousand Juvenile Delinquents—A Critique," by Elkind, Henry B. and Taylor, Maurice, *Mental Hygiene*, Vol. 18, No. 4, October, 1934.

fairly high contingency values actually have little significance. Stouffer and Tibbitts³ have addressed themselves to the technical problem here involved and have sought to apply other statistical methods for determining the significance of factors. The most recent work of the Gluecks⁴ uses a method for determining the importance of factors which combines a simple numerical index of significance (maximum percentage difference between any subclass of a particular factor and the average expectancy rate for the group) with a common-sense evaluation of practical considerations such as completeness of information, apparent overlapping, etc. This combination of common-sense inspection with a simple numerical index seems to be the most readily applicable answer to date of the question of how to determine the significance of factors. Yet, from the standpoint of theoretical completeness in perfecting a tool of analysis, much work remains to be done.

Three additional questions need to be considered in this connection but can only be stated, with no detailed presentation of the problems involved. First, how do the results compare when two or more persons work on the elaborate descriptive data of case records, reducing these to numerical categories? In other words, are there any important differences in the results obtained when different persons score the same items of information? Second, how much better than common sense "guesses" or "estimates" of responsible and experienced officials are the outcomes "predicted" by present techniques? Third, how well do prediction tables work when checked against the actual outcome of the same cases?

The writer's experience with these questions has been unsatisfactory and somewhat disconcerting. Reliability of scoring by two or more persons is relatively high for all categories of definite factual information but decreases rapidly when the factors employ subclasses calling for interpretative judgment. A group of 282 Minnesota State Prison cases gave a correlation of around .5 ($r = .51 \pm .08$) between predicted outcome on parole and the "guess" or "estimate" of

³See: Stouffer, Samuel A., and Tibbitts, C., "The Method of Expected Cases Applied to Sociological Data," *Papers and Proceedings of the American Sociological Society*, 26: 185-186, No. 3, August, 1932. "Tests of Significance in Applying Westergaard's Method of Expected Cases to Sociological Data," *Journal of American Statistical Association*, 28: 293-302, September, 1933. "Testing the Significance of Comparisons in Sociological Data," *American Journal of Sociology*, 40: 357-363, No. 3, November, 1934. See also: Stouffer, Samuel A., "A Technique for Analyzing Sociological Data Classified in Non-Quantitative Groups," *American Journal of Sociology*, 39: 180-193, September, 1933.

⁴Glueck, Sheldon and Eleanor, *500 Delinquent Women*, Alfred A. Knopf, New York, 1934, pp. 278-296.

a responsible parole official. When compared with the actual outcome on parole of the same cases, both correlations dropped considerably. The correlation for predicted outcome and actual outcome was about .4 ($r = .38 \pm .08$) while that of parole agent's estimate and actual outcome dropped to nearly .3 ($r = .302 \pm .07$). Thus it appears that the prediction method and the parole agent's guess agree more closely with one another than either agrees with the actual recorded outcome. These results must not be taken too seriously, however, for the simple reason that during the time interval involved important changes in both personnel and policy of the Parole Board took place.

The empirical test of prediction methods measures the extent to which actual outcomes agree with expected outcomes. Two ready sources of error in interpretation of follow-up results must be guarded against. When actual outcomes and expected outcomes are compared only in terms of the number of cases that fall in the respective percentage groupings, it is possible to have the totals in substantial agreement without having any assurance that the expectancy rate for individual inmates corresponds. Scatter diagrams and statistical measures of association will help give a more nearly correct picture. A second difficulty develops in connection with the interpretation of differences between actual outcome and expected outcome. Discrepancy here need not necessarily mean unreliability of the prediction method. It may also indicate improved methods of selection, or change in procedure, or changes in the situation in other ways. It seems probable that the empirical test of the reliability of prediction requires consistency in performance over a period of time with a fair realization of the requirement that "other conditions remain the same."

3. *Prediction methods in courts and in penal institutions.*

The theoretical possibility of using prediction methods to aid the court has been apparent for some time. The different kinds of treatment available to a court must somehow be made to fit the different types of individuals who appear before it. It is probable that relations can be established between types of individuals and the several treatment alternatives now available, such as fines, probation with supervision in homes, probation with supervision outside of home, incarceration in penal or correctional institutions. The work of the Gluecks,⁵ based on a post-treatment period, has indicated defin-

⁵Glueck, Sheldon and Eleanor, *500 Criminal Careers*, pp. 278-296; *1000 Juvenile Delinquents*, pp. 185-190; *500 Delinquent Women*, pp. 284-298.

ite possibilities in this direction. Monachesi demonstrated that outcome on probation under the conventional arrangements prevailing in Ramsey County, Minnesota, could be forecasted from the probationer's record.⁶

Similarly there is need for classification of inmates of penal institutions and adaptation of the institution's program to the different types of men that enter. A natural question arises: Is the consistency of human behavior so great that prediction methods can be adapted to this problem? In the light of present changing penal methods, are the results obtained from such a procedure worth the effort?

The following section of this paper will be devoted to a brief report on the writer's efforts to apply prediction methods to the different potential treatment groups at the Massachusetts State Prison at Charleston and the Massachusetts State Prison Colony at Norfolk, Massachusetts.

II

Fundamental to the whole idea of penal treatment is the fact of individual differences, but no institution can completely individualize its treatment program. Its best hope lies in being able to select and diagnose groups that will be homogeneous for treatment purposes. That is to say, the treatment for a particular type or group of individuals will be approximately the same. The problem confronting every progressive prison administrator is that of discovering or selecting groups of individuals for different kinds of treatment.

The modern prison of the better type offers many opportunities for self-improvement to the inmate, such as educational advantages, opportunity for sports and recreation, participation in debate, oratory, music, and other cultural activities. Some inmates take advantage of the existing facilities of the institution and develop their personalities to a considerable extent while in prison. Others are indifferent to the institutional program and are satisfied merely to serve time, still others are actively hostile and are continually getting into trouble with the officers and prison administrators. Are there any distinguishing characteristics whereby one can recognize in advance the man who is going to take advantage of the institution's facilities for the development of a better balanced and well-rounded personality? Similarly, is there any way by which one can pick out in advance the man who is going to be a trouble maker? A study

⁶Monachesi, E. D., *Prediction Factors in Probation*, pp. 97-110.

of these "natural" groupings of a present prison population seemed to the writer to offer possibilities for the experimental approach to the problem of prison classification. If definite items of information about the man's past life, known to the prison officials at the time of his incarceration, can be related to his conduct within the institution then a sufficient basis for the application of prediction methods should be at hand. That, broadly speaking, was the program attempted in the Massachusetts' institutions.

A short explanation of procedure followed will help to clarify the results. The cases of 290 inmates of the Massachusetts State Prison Colony at Norfolk (population as of August 13, 1931) and 289 inmates of the Massachusetts State Prison at Charleston (admissions between July 1, 1930 and June 30, 1931, less certain immediate transfers), a total of 579 adult male felons, constituted the principal group studied. The State Prison Colony (Norfolk) was in process of construction at that time and the group was small, with unusually close relations between inmates and officers. The simple plan was therefore followed of having all responsible officers coming in contact with a man, rate him in one of three general classifications:

- (1) As among the best one-fourth at the institution, that is, cooperative and seriously interested in taking advantage of the facilities offered by the institution for his own further development;
- (2) as among the worst one-fourth, that is, a continual trouble maker, disagreeable and generally non-cooperative;
- (3) as belonging to the middle half, neither a conspicuously good prisoner nor a conspicuously bad one.

Since the institution was small, it was possible to get ratings on every inmate by from three to eight officers. The average of all ratings available for each man was taken as that inmate's final rating as to desirability as a prisoner. These ratings were then arranged in rank order for the entire institutional group and the distribution studied for "natural" breaks or divisions. Case record information for each man was then checked and made quantitative under appropriate categories and subclasses similar to those employed in the published parole studies. From this point on the familiar parole prediction procedures were followed.

In the State Prison at Charleston a slightly different procedure was adopted. The institution was at that time following the prac-

tice of having the various officers in charge make monthly written reports on each inmate, grading him as A, B, C, or D on such items as attitude at work, industrial proficiency, leisure time activity, and general ability. These monthly reports were given numerical values, totaled and averaged to give an individual score, making possible the rating or placing of every man in the group in rank order from the "most desirable" to the "least desirable." From that point on the procedure was the same as that followed with the State Prison Colony cases.

Prediction tables constructed from these data gave results surprisingly similar to the familiar parole tables. Both the Glueck and the Burgess methods of scoring were tried with quite similar results. Tables based on many factors (29) were tried and also those based on a few (10) selected ones (importance determined by inspection). The results were not markedly different in either case, though using a larger number of factors tended to give sharper discrimination at the ends of the scale.

Table I gives a fair sample of the results. It is based on twenty-nine factors, using the Burgess method of scoring. It is clear from this that a definite relation may be established between items of information in the inmate's record and his conduct in the institution, as that conduct is judged by officers who know him well. As soon

TABLE I
PROBABLE PRISON CONDUCT, EXPERIENCE AND EXPECTANCY TABLE. GROUP OF
290 CASES, MASSACHUSETTS STATE PRISON COLONY
(Emphasis on effort to select the "Good" inmates)

Score	No. of Cases in Each Class	Prison Conduct Desirability Groupings			
		"Good" ⁷		"Average" and "Poor" ⁸	
		N	%	N	%
23 and over	32	21	65.6	11	34.4
20-22	54	27	50.0	27	50.0
17-19	61	23	37.7	38	62.3
14-16	59	17	28.8	42	71.2
11-13	56	4	7.1	52	92.9
10 and below	28	2	7.1	26	92.9
Totals	290	94		196	

⁷"Good" refers to the group of inmates showing the most promise for constructive use of the period of confinement in the institution.

⁸"Average" refers to those inmates of which nothing much can be said—they are neither constructively nor destructively using their time, but merely "soldiering on the job"; "Poor" refers to those inmates who are hostile to the efforts of the institution and either actively or passively oppose the institutional program.

as the new inmate's case history has been prepared, it is possible to predict his probable reaction to the established institutional routine. If he falls in the score class "23 and over" it may be said that his chances of being rated as giving promise of making constructive use of the institution's facilities are 65.6 in 100, since in the past 65.6 per cent of men with scores in this group were so rated. Similarly, if he falls in the score class "10 and below," his chances of being rated "good" are only 7.1 in 100. The general trend in this table and in the other tables for Norfolk and the State Prison are sufficiently consistent to justify a conclusion that behavior in prison can be predicted.

But the question must be faced frankly, Is this the really important or desirable basis to use in classification of prisoners? The answer is probably negative. Treatment programs should be developed by experiments in common-sense ways under the guidance of an intelligent personnel. Prediction methods will have much greater utility as an instrument of verification for such experimentally determined treatment classifications than as a device for discovering these classifications.