

Summer 1934

## Reliability of Criminal Statistics

Edwin H. Sutherland

C. C. Jr. Van Vechten

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

Edwin H. Sutherland, C. C. Jr. Van Vechten, Reliability of Criminal Statistics, 25 *Am. Inst. Crim. L. & Criminology* 10 (1934-1935)

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.

## THE RELIABILITY OF CRIMINAL STATISTICS<sup>1</sup>

EDWIN H. SUTHERLAND<sup>2</sup> and C. C. VAN VECHTEN, JR.<sup>2</sup>

The social information contained in police records and prison records is generally based on the unverified statements of the prisoners. Notwithstanding the lack of verification these records are used as data in many extensive and expensive research projects and are the most general source of statistical information regarding the social characteristics of criminals. These statistics are sometimes important in the formation and development of social policies and of theories of criminality. If the social information in such records is unreliable the conclusions drawn from them are necessarily unreliable, also. Furthermore, from the administrative point of view, the work of collecting, compiling and publishing this information is not justified unless the data are fairly reliable.

Two attempts have been made previously to determine the extent to which these records and the statistical reports based on them are reliable. In 1915 Miss Hinricksen compared the inconsistencies in the successive arrest records of recidivists in Illinois towns. She reported many inconsistencies in regard to age, nativity, and other social characteristics. Of twenty repeaters in Rockford, Illinois, during 1915 only two gave the same birthplaces, occupations, and ages at each arrest.<sup>3</sup> This method of determining the reliability of criminal statistics is not entirely satisfactory, first because identification of an individual generally depends on the name when the other items, such as age, address, and birthplace, differ and it is quite possible that the investigator may be mistaken in a judgment that all of the records refer to one individual; secondly, even if the data regarding recidivists are proved to be inaccurate, the data regarding the much larger number of prisoners who in a particular year are not recidivists may be accurate.

A second method of evaluating the social information contained

---

<sup>1</sup>This report was made possible by a grant from the Social Science Research Committee of the University of Chicago, and by the hearty cooperation of the Chicago Police Department, the Illinois Institute for Juvenile Research, the Illinois State Penitentiary at Joliet, the Illinois State Reformatory at Pontiac, the Illinois State School for Boys at St. Charles, and the Cook County Juvenile Court.

<sup>2</sup>University of Chicago.

<sup>3</sup>Hinricksen, Annie. Pitfalls in criminal statistics. Proceedings of the American Prison Association, 1916 pp. 393-401.

in prison records was used in a study made recently under the direction of the Institute for Juvenile Research of Illinois. A report has been made for this Institute by Edward A. Conover regarding a study at the Illinois Penitentiary at Joliet in 1931.<sup>4</sup> He selected at random 200 records of prisoners who had been examined by the mental health office in the penitentiary. In all of these cases supplementary information had been secured by correspondence with members of the families, with employers, school authorities, and others. He compared the prisoners' own statements to the mental health office with the information secured by correspondence from the other sources. The extent of the disagreement between the statements of the prisoners and of the correspondents varied widely from one item to another. The general conclusion reached was that substantial divergence was found between the inmates and the correspondents in twenty-four per cent of the comparisons which were made; in twelve per cent of these divergences the correspondence was less favorable to the prisoner than his own version; in seven per cent the correspondence was more favorable than his own version, and in five per cent neutral.

The original plan of the present study was to take a consecutive series of records of arrests by the Chicago police and by independent investigations made by social agencies or by our own efforts to determine the reliability of the social information on the police records and, in cases resulting in imprisonment, on the prison records. This method was abandoned after brief experimentation. Not more than ten per cent of the persons in a sample of fifty arrests could be located in the social service exchange in Chicago. The expense of field investigations of the other ninety per cent would be very great. The police authorities were opposed to such field inquiries. Finally, it would probably be impossible to locate many of the persons even if field studies were made and those who could be located would presumably be a selected group differing in many respects from the ones who could not be located.

The following method was therefore substituted. An effort was made to measure the consistency of all the social information on the records of all the agencies which had had contacts with 507 inmates of state prisons and reformatories in Illinois. The 507 prisoners had been selected for a different purpose and prior to the inaugura-

---

<sup>4</sup>Conover, Edward A. An evaluation of the materials secured by social case work procedures in the Illinois State Penitentiary at Joliet. This paper was read before the American Orthopsychiatric Association in 1931 and is to be published in the Proceedings of that Association.

tion of the present study by assistants at Joliet, Pontiac and St. Charles. The general principle used by these assistants in the selection was consecutive admissions of Chicago-born criminals to these institutions. The principle of consecutive admissions was violated to some extent because of the difficulty or ease of assigning prisoners to the investigators for life histories. Also, a group of prisoners designated "Joliet Special" was added. This group consisted of thirty-two prisoners recorded as born elsewhere than Chicago but committed to the Joliet prison from Chicago. The location and composition of this group of 507 prisoners is shown in Table I.

TABLE I.  
LOCATION AND RACE OF 507 PRISONERS

<i>Institution</i>	<i>White</i>	<i>Colored</i>	<i>Total</i>
Joliet Penitentiary .....	135	7	142
Joliet Special.....	20	12	32
Pontiac Reformatory.....	219	12	231
St. Charles School for Boys.....	71	31	102
Total .....	445	62	507

The following records were secured regarding this group of prisoners: (1) A front office record is taken by the prison clerk at the time of admission to the institution. This furnishes the data from which the published statistics of the Institution are compiled. (2) A mental health office record is taken by the psychiatrist ordinarily later than the day of admission but within the first week. In Pontiac, however, the mental health office record is merely a copy of the front office record and is not utilized in this study. (3) A special schedule was filled out for the Institute for Juvenile Research by the investigators who selected the 507 cases. (4) The records of the arrests which resulted in the present commitments to the institutions were secured for 355 of the 405 inmates in Joliet and Pontiac but were not located in the other 50 cases in those institutions and were not secured in the 102 St. Charles cases, since the arrest file of the police department does not contain the names of juveniles. (5) Previous arrest records of 150 of the 405 Joliet and Pontiac inmates were located. This does not represent a complete search of the police records of arrest. The prior arrest record was investigated only for the period 1930-32 unless the prisoner's name was not found in the files for that period in which case the files for the period 1916-30 were consulted in addition. (6) Juvenile court

records of 57 of the 174 Joliet inmates, 139 of the 231 Pontiac inmates, and 97 of the 102 St. Charles inmates were located. (7) Social service exchange records of 86 of the Joliet inmates and 138 of the Pontiac inmates were secured, but because of the present pressure of work in the social agencies the records of the constituent agencies were not utilized. Similar social service exchange records were presumably available in all of the St. Charles cases since the juvenile court cases are automatically cleared in the exchange, but they were utilized in only two cases which could not be located in the files of the juvenile court. (8) Previous prison records were secured for 44 of the Joliet inmates and 1 of the Pontiac inmates. In addition to the sources listed above, in which an effort was made to locate each of the 507 cases, other sources were utilized for a few extraordinary cases. These include 5 cases located in the files of the Chicago Bureau of Identification and 7 located in the court records which could not be located in the regular arrest records of the Chicago police department, and 4 cases located in the files of the Chicago and Cook County School which could not be located in the juvenile court files. This gave a total of 2622 records of 507 prisoners, or an average of a little more than 5 records per prisoner.

The following items of social information were taken from the records of these agencies and compared for the purpose of determining the consistency of the records: name, age, address, nativity, race, marital status, and occupation. These records were taken independently except that in principle the name of an offender was transmitted in writing from the police department to the court, from the court to the institution, and from the front office of the institution to the mental health office and to the investigators who made the personal interview. In fact, however, the name is not always thus transmitted in writing.

The general method in this study was to compare these various and relatively independent records of identical prisoners for the purpose of determining the frequency of inconsistencies in the records. Inconsistencies in individual records are different from inconsistencies in statistical series. It is possible, for instance, that the five records of a particular prisoner may show five different dates of birth. In that case at least four of the five must be inaccurate. Nevertheless, because of the balancing of errors, the statistical distribution of the ages of all prisoners may be approximately the same in the five types of records. The effort has been made to determine both the inconsistencies in the several records of

a particular prisoner and the inconsistencies in the statistical distribution of records of all of the prisoners on each item of information on which distributions can be secured.

Inconsistencies were found, on the average, in 34.1 per cent of the cases. This is somewhat higher than the average of 24 per cent reported by Conover. In both studies, however, this general average is relatively meaningless, for it combines items on which the inconsistencies are practically zero with items on which inconsistencies are found in the present study in 70 per cent of the cases. Consequently the average is not a stable or necessary symbol but is affected decidedly by the inclusion or exclusion of a few items of information.

The general average in this present study is the result of four principal types of variations. (a) A very large proportion of the variations are in the form of minor differences due to clerical errors, ordinary human carelessness or misunderstandings. The criteria of consistency which were used were extremely rigorous, including variations in one letter of a name, one figure in an address, one day in the age, and any verbal difference in the description of occupation. (b) Factual changes during the interval between the taking of the several records account for changes in some of the items. This is a much less important source of variation than the clerical errors, and is not found at all in certain items such as place of birth, date of birth, or race. (c) The several agencies are seeking different information on some points, though the information is entered on the records under approximately the same rubrics. The police make a record of last occupation, the penal institutions of principal occupation; the police make a record of address of the offender at the time of arrest, the penal institutions make a record of the address of relatives to whom communications may be sent by the institution; the police classify prisoners by marital status as either single or married, while the penal institutions use five classes for the same data. (d) Purposive deception is the motive in a certain part of the inconsistencies.

If correction is made for the minor variations due to carelessness, for the factual changes in the interval between arrest record and subsequent record, and for differences in the conception of items, the inconsistencies are reduced to about 10 per cent. In other words, approximately 10 per cent of the inconsistencies in records are probably due to intentional deception on the part of prisoners. This is not a completely objective summary of the data but involves an element of subjective appraisal. Moreover, it does not necessarily mean that the unreliability of the data would be 10

per cent if the corrections listed above were made, for the reason that the records might be unreliable due to consistency in deceptions. It was found in a study of the aliases of prisoners that 17 per cent of the prisoners in the present study had at some time or other used fictitious names evidently for the purpose of deception. An attempt to deceive in regard to the right name is probably made more frequently than on any other item of social information.

The percentage of cases in which the records of the several agencies were not entirely consistent is shown in Table II.

TABLE II.  
PER CENT OF PRISONERS WHOSE RECORDS WERE INCONSISTENT  
ON SPECIFIED ITEMS OF SOCIAL INFORMATION

<i>Item</i>	<i>Per Cent of Inconsistency</i>
Race .....	0.8
Nativity .....	10.3
Marital Status .....	17.3
Name .....	39.3
Address .....	41.7
Age .....	60.8
Occupation .....	69.7
Average .....	34.1

This is based on 507 prisoners for all items except marital status and occupation, in which only 305 prisoners are used, due to the exclusion of the St. Charles boys who are generally below the age of marriage and of regular occupation. This Table shows a range of inconsistencies from 0.8 for race to 69.7 for occupation. One would expect the greatest consistency in regard to race and age, which are more completely susceptible to check by observation. This expectation is met in regard to race, but not in regard to age. The failure of age to rank next to race is probably due to the fact that variations of one day in the age, when correction is made for the passage of time, is counted as an inconsistency. In other respects, also, the ranking of the several items in regard to the percentage of inconsistencies does not represent a necessary or fixed relationship. The low rank of nativity is explained primarily by the method of selection of cases for the present study. The low rank of marital status is explained in part by the inclusion of a large group of Pontiac Reformatory inmates who are generally in fact and according to the records single. The items which rank highest in inconsistencies are those in regard to which the smallest units or classes are used in

measuring inconsistencies, namely, one letter, one figure, or one occupational term. The record of occupation may be least consistent either because prisoners make false reports regarding occupation in order to secure desirable labor assignments in prison, because the conception held by the prison in recording occupation is different from the conception held by the police, because occupations are changed readily, or because occupational terminology is not standardized. The last is presumably the principal reason for inconsistencies. If broader criteria of occupational similarity could be utilized, it is probable that the rate of inconsistency would be greatly reduced. If the slight errors in spelling are eliminated, inconsistency in regard to name is reduced to 3 per cent of the cases. If the addresses which show a variation of less than two blocks are not counted as inconsistencies, the percentage of cases in which the addresses are not recorded consistently is reduced from 41.7 to about 20 per cent. Similarly if the minor variations in age are eliminated the percentage of inconsistency is reduced to 30. In general, it is fairly safe to conclude that substantial differences between the records of the agencies are found, on the average, in about 10 per cent of the cases.

The rate of inconsistency is 27.9 per cent for the inmates of St. Charles, 32.8 per cent in the Joliet cases and 37.8 in the Pontiac cases. The Pontiac cases have the highest rate of inconsistency of any institutional group in regard to name, age, address and nativity, the Joliet cases have the highest rate in regard to marital status and occupation. The lower rate in St. Charles is presumably based in part on the fact that field studies were made in securing the records. The lower rate of inconsistency in Pontiac in regard to marital status is probably due to the fact that the Pontiac group is younger and has a much larger proportion of inmates who have not been married and who consistently and truthfully report themselves on each record to be single.

The rate of inconsistency in the records of white prisoners is 33.8, of colored prisoners 36.2. This difference is not great enough to be very significant, but the rate is higher for the colored than for white prisoners in regard to nativity, address, marital status, age, race, and occupation. The only item on which they have a lower rate—and it is at this point significantly lower—is in regard to the name and this is due to the greater ease of spelling the names of colored prisoners than the names of many prisoners of foreign stock. The higher rate of inconsistencies of records of colored prisoners is probably not due to greater frequency of efforts at deception.

The rate of inconsistencies in items which are necessarily or customarily stable and fixed (date and place of birth, race, and name) is lower than, the rate in the items which more readily change (address, marital status, and occupation). Nativity is excluded from this comparison because of the method used in the selection of cases for this study. The comparison of these groups of items with regard to inconsistency in the records is shown in Table III. The rate of inconsistency is significantly higher for the items

TABLE III.  
COMPARISON OF THE FIXED ITEMS AND THE CHANGEABLE ITEMS  
IN REGARD TO PER CENT OF INCONSISTENCIES

<i>Items</i>	<i>White</i>	<i>Colored</i>	<i>Total</i>
Items which are fixed (age, race, name).....	33.6	30.1	33.2
Items which change (address, marital status, occupation) .....	43.9	54.9	45.3
Total .....	39.5	42.5	39.6*

\*This total is different from the total presented previously due to the exclusion of the data regarding nativity from the present table.

which are easily changed than for those which are fixed and stable. The rate of inconsistencies for the unchangeable group would be not more than 5 per cent if the minor variations in the spelling of the name and errors of two years in age were eliminated. But similarly a large part of the inconsistency in the record on items which change would be eliminated if the approximate neighborhood rather than the specific address, and the general range of occupational skill rather than the particular occupational terminology were used as the criteria of consistency.

Variations in the records tend to balance when the data are presented as statistical tabulations. The general conclusion derived from the records of one agency would be, statistically, almost the same as the conclusions derived from the records of another of the agencies used in this study. The statistical distribution of ages shows the same general form of curve for each agency, with a variation in mean ages of about a half year. The police records tend to exaggerate the ages of offenders, in comparison with the other records, and this is true particularly in the younger age groups and is probably explained by the regulation that no child under 16 years of age is to be recorded. The statistical distribution of nativities, with reference here only to the comparison of foreign born and native born,

similarly shows an exaggeration in the police reports of the proportion of foreign born prisoners in comparison with the other records. In the selected group used for this study with a very small number of foreign born offenders, the disagreements were due almost entirely to the report by the police that a prisoner was foreign born, while the other records and, in case of recidivists, many of the prior arrest records report him to be native born. The police records in this sense exaggerate the proportion of prisoners who are foreign born by approximately 10 per cent. This is due primarily to the record form used by the police, which makes no differentiation of nativity and nationality. Similarly, the records taken within the institutions were consistent statistically in regard to marital status though they showed many individual variations. There was wide variation between the records taken within the institutions and the police records, due to the two-fold classification used by the police and the five-fold classification used by the institutions. A statistical distribution of occupations was not attempted because the information regarding occupations was not sufficiently specific to justify a classification on the basis of a single principle. The impression derived from inspection of the entries regarding occupation is that the variations are almost entirely within narrow ranges of skills and that if fairly broad class intervals could be constructed objectively there would be little difference in the proportions of prisoners in the several occupational classes. The addresses in a group of cases selected for more intensive analysis were exactly identical according to the several records in about two-thirds of the cases, and in the other third in which variations did occur, the distances between the addresses was, on the average, about one mile. These variations, moreover, tend to balance each other so far as delinquency areas are concerned. These various comparisons tend, in general, to justify a generalization that the statistical distribution derived from the records of one agency are highly consistent with the statistical distributions on the same item derived from another agency.

The records of the police department gave rise to more inconsistencies when compared with other records than did any other agency. This is due in part to the higher average age recorded by the police department, in part to the higher proportion of foreign born prisoners, but principally to the two-fold classification in regard to marital status, and to the inquiry regarding occupation at time of arrest by the police while the inquiry had reference to general or principal occupation in the other agencies. On the other hand, the

inconsistencies in the several police records of recidivists are less than the inconsistencies in the records of all agencies included in this study. The rate of inconsistency for fixed items in the police records of recidivists is 24.6 per cent, for the changeable items 27.9 per cent. This lower rate of inconsistency in the police records is presumably due to the uniform definition of terms and the uniform criteria used in the several police stations as compared with different definitions and criteria used by the several agencies.

The following general conclusions may be derived from this analysis: (1) The several agencies used in this study are exactly consistent in their records of any item of social information about a particular prisoner, on the average, in about two thirds of the cases, and are substantially consistent in about nine tenths of the cases. (2) The records of any one of the agencies yield about the same average and statistical distribution of data on any one item regarding all prisoners as do the records of any other agency. (3) The consistency of the several records of a particular prisoner and the statistical consistency of the several records of a group of prisoners may not properly be translated into terms of reliability or truth. The records may be consistently untrue. However, it was found that at least 17 per cent of the prisoners had used fictitious names, as checked by finger print evidence, at some time or other. (4) Approximately two thirds of the inconsistencies in the records of the several agencies could be eliminated by greater care in recording and by standardization of definitions and classifications.

The following recommendations<sup>5</sup> develop from this analysis: (1) The terms which are used in recording data regarding criminals should be standardized. Address, occupation, nativity, and the other items should mean the same thing to the several agencies. It is particularly important that nativity and nationality should not be confused. (2) The classifications on each item should be identical. The classification used in recording marital status could easily be made uniform in all the agencies. It is highly important but much more difficult to develop a similar classification of occupations. (3) The names should be filed in every agency phonetically rather than alphabetically. (4) Other studies of the reliability of criminal statistics should be made. The method used by Conover in his Joliet study of comparing prisoner's statements with correspondent's statements is the easiest method from an administrative point of view,

---

<sup>5</sup>In the re-organization of the Chicago Police Department some of the modifications recommended have been made already.

but has limitations in the form of inadequate replies and biased replies. The method used in the present study has values in the preliminary study of the reliability of the statistics but can not yield positive conclusions regarding reliability in most states. It could perhaps be used to best advantage in a state like Massachusetts where a large proportion of the prisoners have records verified by field investigations. It is probable that the only adequate method of testing these records will be by a system of registration of inhabitants of the United States, either in the form of European registration systems or in the form of general finger print records which have been recommended by some of the police associations. With the extension of finger-printing to the army, the postal-savings accounts, hospitals, banks, immigration departments and other groups of people, we shall soon have finger-print records only slightly short of universal.