

Winter 1933

## Association Test As a Measure of Delinquency, The

A. R. Gilliland

J. C. Eberhart

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

A. R. Gilliland, J. C. Eberhart, Association Test As a Measure of Delinquency, The, 24 Am. Inst. Crim. L. & Criminology 736 (1933-1934)

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 ASSOCIATION TEST AS A MEASURE OF DELINQUENCY

A. R. GILLILAND and J. C. EBERHART<sup>1</sup>

In 1925 H. R. Laslett published the results of an attempt to diagnose delinquency by means of the word-association test technique.<sup>2</sup> His results indicated that the list of words he had chosen might be expected to distinguish individual delinquents from non-delinquents. The present paper is a report of the application of Laslett's test to 546 Chicago boys, selected so as to represent various gradations of delinquency and non-delinquency.<sup>3</sup>

The word-association test technique involves the presentation to a subject or a group of subjects of a series of words, one at a time, with instructions to respond to each word with the first word that comes to mind. The technique was originally devised for use with psychopathic patients as a help in discovering complexes, etc., and has recently been used with considerable success by Crosland<sup>4</sup> and others in the detection of guilt. The method had not been used so far as we know prior to 1925 as a tool in the diagnosis of general delinquency, although Eastman and Rosanoff in 1912 had reported the results of some general work done on delinquent and feeble-minded children with the association test.<sup>5</sup> Their groups were not specifically defined enough as to delinquency, and their words not specially chosen, so that their results have little bearing on the present problem.

The advantages in the development of a technique which would distinguish the delinquent from the non-delinquent, or the potential delinquent from the non-delinquent, or the "delinquent-minded" (if that term can be so defined<sup>6</sup> that the group it designates is sociologi-

<sup>1</sup>Department of Psychology, Northwestern University.

<sup>2</sup>Laslett, H. R. Preliminary notes on a test of delinquent tendencies, *Jour. Delin.*, 1925, 9, 222-230.

<sup>3</sup>This study was undertaken as a part of a larger research centering around the differences in attitude between delinquent and non-delinquent boys. Both this and the larger study were made possible by a grant from the Payne Fund of New York.

<sup>4</sup>Crosland, H. R. The psychological methods of word-association and reaction-time as tests of deception. Uni. of Oregon Pub. Psych. Series No. 1 (1929).

<sup>5</sup>Eastman, F. C. and Rosanoff, A. J. Association in feeble-minded and delinquent children. *Amer. Jour. Insan.*, 1912, 69, 125-141.

<sup>6</sup>Laslett defined "delinquent-mindedness" as "the tendency to reflect on,

cally significant) from the non-delinquent-minded, are obvious. The primary value of such a technique would be, of course, in the prevention of delinquency. The task of prevention is itself a difficult one, but it appears difficult if not impossible to attack it until some diagnostic tool is developed. Many attempts have been made in this direction, but no one has achieved outstanding success. In this connection the authors, after a survey of the literature on tests of delinquency, concluded that the most feasible channel to follow was that chosen by Laslett. It was accordingly decided to find out the adequacy of Laslett's test to differentiate between delinquent and non-delinquent groups of boys in Chicago.

The specific advantages the association test has for diagnostic work with delinquents have been well stated by Laslett,<sup>7</sup> and will not be repeated here.

#### I. *The nature of Laslett's test.*

In selecting his words, Laslett first picked out of a standard desk dictionary 1,200 words which seemed capable of arousing two, three or more lines of association, at least one common to delinquents and one or more not. To this list he added from various local sources, and then with the help of two other judges he reduced the total number of words to 360. These 360 words were given to 150 boys, half of whom were delinquent and half non-delinquent. On the basis of the responses of this group he reduced the list to 96 words.

The scores for the responses to each stimulus word were obtained from another group of 300 boys. The list was given to 150 delinquents in the state schools and to 150 boys from public schools of a high character, and scores were given to the various responses on the basis of the relative proportions of each of the above groups that responded to the stimulus word in similar fashion. The scores given bore a plus sign when more non-delinquents than delinquents gave a certain response, and bore a minus sign when the opposite condition held true.

The range of scores Laslett obtained as from +173 to -161. The reliability of the test was calculated to be +.82. Correlations of test scores with teacher ratings on morality, with I. Q.'s with mental ages and with the percent of time delinquents in the state school had spent in the "No Privilege Cottage" all were under .10. Chronological age correlated with association test scores +.446 for

wish for, and approve of acts which, if performed, would result in severe social disapproval" (op. cit., p. 222).

<sup>7</sup>Op. cit., p. 223.

49 cases, showing that for this group gain in age was marked by gain in delinquent vocabulary. The scores for one hundred of the cases used to establish the scoring-system—fifty delinquent and fifty non-delinquent—were plotted, and only four scores overlapped. This result pointed to an excellent diagnostic capacity for the test.

Three criticisms must be made of Laslett's work: (1) His definition of delinquency is so broad that it makes the validation of any diagnostic test practically impossible. He says,

"I shall take delinquency to mean not only those acts which are unapproved by society to the extent that children committing them on several occasions are sent to the reform schools but also to mean delinquent-mindedness, and all degrees of misconduct and evil habits of thought that lie between these two extremes. By delinquent-mindedness I mean the tendency to reflect on, wish for, and approve of acts which, if performed, would result in severe social disapproval."<sup>8</sup>

It is easy enough to estimate the validity of a test of delinquency by applying it to a group of convicted delinquents in a state school and to a group of non-delinquents wherever available. This has been the method generally followed. But if we so define "delinquent" that the term includes even individuals who think about delinquent acts, we have no criterion by which to pick out critical groups. An analogous situation would be that of attempting to devise a test diagnostic of tubercular tendencies, and with the further provisions that no method was available for isolating individuals with tubercular tendencies from those without. There would thus be no definitely non-tubercular individuals, and hence no way to validate the test. If a test, either of delinquency or tuberculosis, be devised so that it distinguishes well between two critical groups it can be used to pick out individuals like either of those groups from an unselected population. When two groups cannot be set up, however, and when no other measure of the trait in question is available, then the task of attaching significance to the test scores, i. e., the task of validating them, is impossible.

(2) Laslett's evidence that his test has diagnostic value is in the bimodal nature of his distribution of 100 scores—50 non-delinquent and 50 delinquent. As was mentioned above, only four scores overlapped. As was also mentioned above all of these hundred cases were among the 300 whose responses were used to work out the scoring system. The fact that striking differences between the scores of these two groups were obtained is no indication that the test would

<sup>8</sup>Op. cit., p. 222.

distinguish equally well between new groups of delinquents and non-delinquents. The practice of testing the validity of a scoring system by applying the test to groups which had been used to work out that scoring system, although rather widely used, cannot be logically justified. Laslett's distribution indicates that there is an internal consistency to his scoring system, but it gives little evidence as to the validity of the scores.

(3) Laslett's statement that his non-delinquent subjects were from "public schools of a high character" is the only clue we have as to possible differences other than delinquency between his two groups. Unless a non-delinquent group is carefully selected it will generally represent a much higher social level than will any delinquent group, and it is reasonable to suppose that Laslett's groups were typical. It is important to inquire, therefore, how much of the association test scores is due to differences in social and economic level, and how much to the factor of delinquency itself. No test depending largely on vocabulary can be considered exclusively a test of delinquency until it has been demonstrated that the factors of social and economic level are not responsible for the differentiating capacity of the test.

## II. *The present experiment.*

The present experiment was designed to check the diagnostic effectiveness of Laslett's list of words for Chicago boys. It was desired also to secure groups of subjects who differed greatly as to social level, in order that the possible effect of social level on responses to the test might be measured. The easiest method of measuring the validity of the test was to compare test scores of two groups, one delinquent and one non-delinquent. The greater the separation of the scores of the two groups, the greater would be the diagnostic value of the test. The alternative method—that of relating association test scores of a group of subjects to some other measure of delinquency—would probably involve the uncertainties of ratings, and should be avoided if possible. Accordingly the group method of validation was chosen for the present experiment.

Subjects were taken from four schools in the Chicago area.<sup>9</sup>

---

<sup>9</sup>The authors wish to express their appreciation of the cooperation given by Miss Isabella Dolton, Ass't. Supt. of Chicago schools; Vernon L. Bowyer, principal of Skinner School; E. H. Stullken, principal, and Miss Charlotte Lowe, psychologist, of Montefiore Special School; Mr. Colwell and Mrs. Rogers, principal and psychologist of Sullivan Junior High School; Supt. Havlik and Miss Milly E. Patton, principal of the academic division of St. Charles School for boys; and Dr. Paul L. Schroeder, Director of the Institute for Juvenile Research, and the Institute Staff at St. Charles.

The extreme delinquent group comprised 208 boys from the St. Charles School for Boys, a reform school that draws from both Chicago and downstate. Of this group 112 were from Chicago, and hence comparable as far as previous environment goes with the other groups used. All of this group of 208, of course, were convicted delinquents, but it is commonly believed that the Chicago boys in the school are as a rule much worse than the boys sent in from other parts of the state. This is generally ascribed to the fact that in Chicago all other agencies for the care of delinquents are exhausted before a boy is sent to the state school, while the rest of the state has no such agencies, and often commits to St. Charles on the first offense.

The extreme non-delinquent group was made up of 133 seventh, eighth, and ninth grade boys from the Sullivan Junior High School. This school was chosen because it is located in and draws pupils from an area with one of the lowest delinquency rates in the city of Chicago. The region is known locally as Rogers Park, and is one of the better residential districts of Chicago. The delinquency rate where the school is located and in nearby square mile areas ranges from 0.0 to 0.5 offenders per hundred of the same age and sex.<sup>10</sup> This may be compared with a rate of 21.8 near the stockyards. Of these 133, 44 were in 7B, 48 in 8B, and 41 in 9B. Age and mental test status will be given below. It should be mentioned here, however, that these grades were chosen so that the subjects would be as similar to the St. Charles subjects in mental age as possible.

In between these two groups as far as delinquency goes are 93 subjects from Skinner grammar school (sixth, seventh, and eighth grades), and 112 from Montefiore Special School, a school where incorrigibles and truants from all over the north side of Chicago are sent. The delinquency rate in the square mile area in which Skinner is located is 9.0. Neighboring square mile areas have rates of 12.6, 18.0 and 21.1. Skinner draws from all of these. The delinquency rate data have no bearing on the Montefiore subjects because their homes are scattered over such a wide area, but many of the boys are known to have committed delinquent acts. Many are committed to the school for behaviors which might, in other communities, cause them to be sent to the reform school. We are justified in classing the Montefiore group of subjects as less delinquent than the Chicago

---

<sup>10</sup>These figures are taken from *Delinquency Areas*, Clifford R. Shaw, University of Chicago Press, 1929. They are based upon the 8,591 alleged male juvenile delinquents (10-16 years) dealt with by juvenile police probation officers during the year 1927. Later figures being worked on by Dr Shaw are not yet available.

group at St. Charles, but more delinquent than the Skinner or Sullivan subjects.

Laslett, in the article cited, spoke of his test as being composed of 96 words. His published list of stimulus words contained 95 words. The scoring key for the list supplied the authors through the kindness of Professor L. M. Terman of Stanford University, gave scores for just 88 words, and it is these 88 that were used in the present experiment. They are as follows:

queer	touchy	trusty	blue
big	odds	nuisance	cell
mad	club	muscle	frisk
blind	palm	felony	promise
egg	fare	Little Joe	cheat
judge	lay-out	break	bully
camel	work	funny	shame
extravagant	rods	blood	good
dare	knife	hungry	kill
faith	cunning	kind	flirt
case	lawyer	beautiful	crooked
fight	ring	outlaw	pug
liberal	term	fairy	sheriff
false	beat	box car	love
tough	yellow	rude	deny
elegant	luxury	accuse	steal
safe	attack	minor	mash
flush	bar	brag	gun
beg	game	forge	graft
bounce	switch	bet	blow
happy	gamble	rough	chip
gunner	box	pick	cattle

The list was given as a group test. All subjects other than those at Skinner were provided with mimeographed blanks numbered from 1 to 88. Instructions given were the same as those used by Laslett. As each stimulus word was pronounced to the subjects it was held up on a hand-lettered card so that they could see as well as hear it. The lettered words were three centimeters high.

Between five and ten per cent of the papers could not be scored. The reasons were: copying stimulus words, writing responses before the stimulus words were pronounced, inability to think of and write down association in the time allotted (about 8 seconds per word), associating objects in the room rather than the stimulus words, etc. Where possible these subjects were given the list of words again individually, and a scorable paper obtained.

The delinquent, or St. Charles group, was slightly older than any of the others. The average age of the Chicago boys at St. Charles was 16.12 years; the downstate group at St. Charles averaged 15.42. The Montefiore group was next with a mean age of 14.66, and Sullivan followed with 13.25. Age data were not secured for the Skinner group, but the probable average age for this group was less than that for Sullivan.

Mental test data were available for all groups except Skinner, and showed that the groups differed much less widely in mental age than in chronological age. The average mental ages of the various groups were:

Chicago boys at St. Charles.....	14.27
Downstate boys at St. Charles.....	13.93
Montefiore .....	12.15
Sullivan .....	14.31

The subjects from Sullivan Junior High as would necessarily follow from the above data showed a much higher average I. Q. than did those from St. Charles or Montefiore. The means are:

Sullivan .....	108.02
Montefiore .....	82.10
Chicago boys at St. Charles.....	85.77
Downstate boys at St. Charles.....	90.82

### III. *Results.*

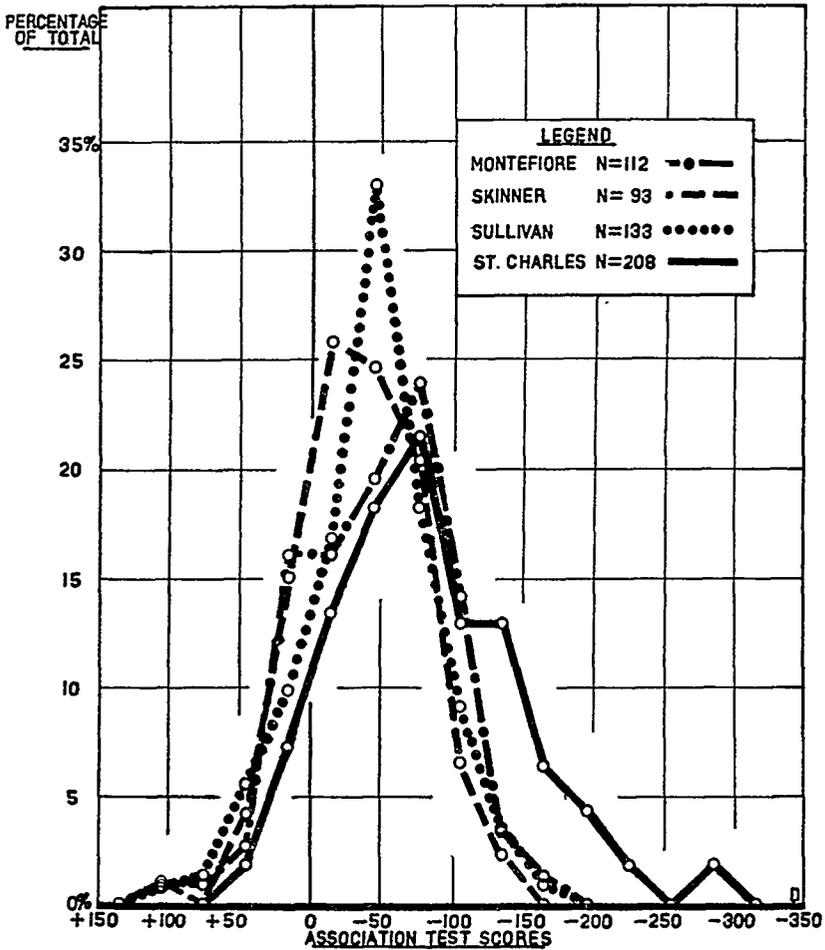
The test scores of all groups of subjects used are given in Table I; the distribution is further illustrated in Figure I.

TABLE I  
ASSOCIATION TEST SCORES OF SUBJECTS

Scores	St. Charles	Montefiore	Skinner	Sullivan
91-100		1	1	1
81-90				2
71-80		1		
61-70				
51-60				2
41-50	1	1	2	1
31-40	3	2	2	5
21-30	4	4		3
11-20	4	7	3	5
1-10	7	7	11	6
-9-0	5	3	5	4
-19-10	10	7	13	8
-29-20	13	8	6	12
-39-30	10	8	6	12
-49-40	13	6	11	9
-59-50	15	8	5	12
-69-60	18	16	9	12
-79-70	17	7	4	11
-89-80	10	4	6	3
-99-90	14	5	2	4
-109-100	7	9	2	7
-119-110	6	3	2	2
-129-120	8	1	2	2
-139-130	8	3		3
-149-140	11			
-159-150	5			1
-169-160	4	1		1
-179-170	4			
-189-180	4			
-199-190	3			
-209-200	2			
-219-210				
-229-220				
-239-230	1			
-249-240				
-259-250				
-269-260				
-279-270				
-289-280				
-299-290	1			
Number	208	112	93	133
Mean	-76.37	-43.70	-33.30	-41.99
S. D.	58.8	48.02	40.61	50.3
$\sigma$ (Mean)	4.078	4.55	4.21	4.36

FIGURE 1.

DISTRIBUTION OF ASSOCIATION TEST SCORES FOR ALL SUBJECTS  
 Frequency of Each Interval Expressed as a Percentage of  
 Total Population of the Group



The test scores show, briefly, that (1) Skinner scores are highest, and that Sullivan, Montefiore and St. Charles follow in order; (2) the differences between the first three means are so slight as to be almost negligible; (3) there is a significant difference (differences between 5 and 7 times the standard error of the difference in each case) between the mean of the St. Charles scores and the means of the scores of the other three schools. Differences as much larger than their standard errors as are these tell us that we have found true differences, and that in comparable experiments similar differences would be obtained in 99 out of 100 instances.

It has been pointed out that the Chicago boys at St. Charles and the Sullivan group represent the extremes of delinquency and non-delinquency. If the association test has a diagnostic value, there should be a considerable difference between the scores of these two groups. The mean scores are, respectively,  $-84.1$  for the St. Charles Chicago group, and  $-42.0$  for Sullivan. The difference of 42.1 points is 5.9 times its standard error, and as a difference 3 times its standard error is generally considered reliable, we are safe in concluding that in the present instance the obtained difference is not due to chance.

Thirty-five of the Chicago group at St. Charles were listed as repeaters, while the other 77 were first offenders. The recidivist group had a mean association test score of  $-98.94$ , while the score of the group of first offenders was  $-76.80$ . The difference of 22.14 points has a standard error of 11.23, which is contained in the difference 1.97 times. The chances are thus 98 in 100 that the true difference between the two groups is greater than zero.

Laslett found practically no relation between I. Q.'s and association test scores, and the same for mental age and test scores. But he did obtain a correlation of  $+.446$  between age and delinquency scores for 49 subjects. By the use of the mental test and chronological age data summarized for our subjects above, the following correlations were obtained:

TABLE II.

CORRELATIONS (product moment  $r$ 's) OF TEST SCORES WITH DEVELOPMENTAL DATA

Subjects	N	$r$	PE
<i>M. A. &amp; test scores</i>			
Montefiore	109	$-.05$	$\pm .07$
St. Charles, Chicago group	106	$-.05$	$\pm .07$
St. Charles, downstate group	85	$-.27$	$\pm .08$
<i>I. Q. &amp; test scores</i>			
Montefiore	109	$.01$	$\pm .07$
St. Charles, Chicago group	106	$-.07$	$\pm .07$
St. Charles, downstate group	84	$-.17$	$\pm .08$
<i>C. A. &amp; test scores</i>			
Montefiore and St. Charles	296	$-.20$	$\pm .04$

The tendency in all these correlations is for high chronological age, high mental age, and high I. Q. to be associated with large negative, or delinquent, association scores. The coefficients are so small, however, that no one of these factors can be pointed to as being primarily responsible for the association test scores.

The reliability of the test was found by Laslett to be  $+.82$ . He used the split half method with 100 cases, 50 delinquent and 50 non-delinquent, and corrected his raw coefficient by the Spearman-Brown formula. In repeating this process with the present groups, the writers selected at random fifty papers from the St. Charles group and fifty from Sullivan. The coefficient of reliability, corrected from the raw correlation between the split halves of the test, was  $+.39 \pm .06$ . The variation from Laslett's result is considerable. A reliability of  $+.39$  indicates an instability that may explain why the test did not distinguish between the groups used any better than it did.

In order to discover the reasons for the low reliability of the test, the individual responses given to the first 25 stimulus words were tabulated. The results are revealing. For these 25 stimuli Laslett provides scores for 966 different response words. Tabulation of the results of the present study show that more than 40% of these 966 words were not given as responses by a *single one of our 546 subjects*. In other words, 40% of the scores on the test were not used at all.

When Casselberry<sup>11</sup> restandardized Laslett's test as a part of a battery for the prediction of delinquency, he gave weights to every reply that appeared at least four times in a hundred papers. Our analysis shows that only 9.7% of the words Laslett gave weights to were given by as many as four out of a hundred of our papers, so that if we had been restandardizing on the same basis, we could have used only 9.7% of the response words Laslett used. Further, only 15.6% of Laslett's response words were given by as many as *two* out of a hundred of our subjects.

It is evident that many of the response words given by our subjects were not among those listed by Laslett. The low reliability was undoubtedly caused in large measure by the few response words which could be scored.

#### IV. *Conclusion.*

Laslett's results indicated that the association test might be expected to pick out delinquent cases from an unselected group—in-

<sup>11</sup>Casselberry, Wm. S. Analysis and prediction of delinquency. *Jour. Juv. Res.*, 1932, 16, 1-21.

dividual diagnosis. Our results from application of his test to Chicago groups do not confirm that prediction. The great disparity in the nature of the response words given by Laslett's groups and by the Chicago groups suggests a considerable difference in general vocabulary.<sup>12</sup> Whether this is a function of the time intervening between the two experiments, or of a difference in environment in Chicago and California cannot be determined from the present data. Whether any group of words could be found to which significant associations would be given by boys from all parts of the United States remains to be seen. At least Laslett's test as now constructed is not satisfactory for this purpose.

It must not be forgotten, of course, that statistically significant differences were found between the delinquent and the non-delinquent groups, and between the first offenders and the recidivists. These facts, however, neither justify the hopes Laslett expressed for the test nor are of particular value in attacking the problems of delinquency. In individual diagnosis lies the key to a successful beginning on the solution of delinquency's problems.

The question of the relative influences of actual delinquency and of social and economic level on association test scores remains unanswered. There were marked differences in both social level and delinquency between the Chicago boys at St. Charles and the Sullivan Junior High group. Yet the difference in test score was only 5.7 times its standard error. This relatively small difference made unnecessary the writers' proposed plan of selecting two groups of boys with identical social level but with a wide difference in delinquency, for the purpose of isolating the actual influence of delinquency on test scores.

Finally the writers wish to emphasize that the present study has not proved the impracticability of the association technique in the study of delinquency. Their data have shown, however, that this method, along with other similar-aimed methods, must be carefully used and cautiously interpreted. It seems quite probable that Casselberry's method of trying out every possible measure of delinquency and using the best ones as a battery may turn out to be the only successful approach to the problem of "measuring" general delinquency.

---

<sup>12</sup>That such a difference does exist was the contention of Dr. L. S. Selling, of the Institute for Juvenile Research, Chicago, in conversation with the writers. Dr. Selling based his statement on data collected by himself and others on the vocabulary of Chicago delinquents.