

MINNESOTA
SOCIAL SECURITY BOARD
LIBRARY

DELINQUENCY AND FEEBLE-MINDEDNESS. A REPLY TO VOLD.

F. Kuhlmann

My own view in common with that of most psychologists of today, I believe, includes essentially three postulates. First, that certain mental traits characterizing the feeble-minded naturally and logically lead to delinquency. Second, that the inability of the feeble-minded to maintain a normal economic status results in their congregating in the least desirable and "slum" districts of our towns, thus in a measure creating their own environment unfavorable to normal conduct. Third, that by actual mental examinations delinquents show a much higher percentage of feeble-mindedness than exists in the general population.

In the paper read before this group Vold states the psychologists' position as follows: First, that it is a form of Neo-Lombrosianism. Second, that "the great majority of criminals are feeble-minded." Third, that "a policy of sterilization or segregation of the feeble-minded is the only effective method of preventing crime and of dealing with criminals." He includes some other postulates relative to the mental traits of the feeble-minded to which I am not taking exceptions. His other three statements are quite incorrect. Psychology's present position has been affected not in the least by the Lombrosian theory. The prefix "Neo" attached to an idea means that it was derived from something that preceded, or that it was a natural consequence. There might be some mis-informed extremist among present psychologists who would say that the great majority of delinquents are feeble-minded, but it certainly does not express a general view. That sterilization or segregation of the feeble-minded would be the only effective method of preventing crime does not follow even from this extreme position, and I am sure it is not the general view of psychology. I shall not take any further time to support these flat denials of Vold's statements on what psychologists think, for this is not the important issue. The main issue is the question of the frequency of feeble-mindedness among criminals and delinquents.

After reviewing a number of studies and giving some quotations, Vold concludes that "These studies all point to the same general conclusion, namely, that feeble-mindedness is approximately twice as frequent among criminals and delinquents confined in penal and correctional institutions as in the general population." But in the end, he states that the inmates of reformatories and prisons represent "a fair cross-section of the general population which feeds the institutions," and that "Crime would seem to be an inevitable consequence of social organization." In other words, there is no personality trait, not created by social organization, that is responsible for crime and delinquency. He comes to this conclusion, first, by rejecting the results of mental test examinations of reformatory and prison inmates. He does this on the grounds that the mental test norms are based on the performance of children and on his claim that the performance of normal adults on these tests is or would be below that of children. He goes to the Army test results to prove that adult performance is below that of children.

There are two ways of testing the correctness of a conclusion. One is to examine the evidence on which it is based. The other is to see what it leads to. If it leads to a contradiction of other known facts it becomes untenable.

35-DFH-FNK

Let us give a moment to the latter method. First, people growing up in the same environment, in fact, children of the same family, in innumerable known instances behave differently. Some remain law-abiding, while others become delinquent. Shall we say with Vold that some unobserved difference in social organization has caused the difference in behavior, or that there was an original difference in the children on which one and the same society had different effects?

Second, psychiatrists and psychologists have observed many cases where a particular delinquent act was pretty indisputably due to some particular mental state or condition. It is indisputable that the insane sometimes do commit crimes in acting out delusions; that some epileptics and insane are abnormally subject to fits of anger that lead to rash acts and assaults; that feeble-minded sometimes become delinquent because they are unable to resist a temptation that a normal person would, or because they fail to understand the act or consequence that a normal person would understand. The presumption is that these cases occur more frequently than actually observed. Dr. Vold's claim is that they do not occur at all, or at least so infrequently that they do not show up in our statistical data when that data is correctly interpreted.

Third, it is pretty well agreed that criminals are not all alike; that we should study the personality of each individual criminal, as well as his history and records, on the basis of which he should receive individual treatment. Without attempting to say why people have come to this conclusion, I am certain that most of them believe that the personality differences in question are largely original, and not simply the effects of differences in social organization that they have contacted. If there were no original personality differences we would really do better by limiting our study to the criminal's training, and environment, and whatever else could be called social organization, in order to get at the direct causes of his delinquency.

Fourth, Vold's position denies entirely any individual responsibility as ordinarily understood. If society is entirely to blame we owe much to the criminal for the wrongs we have done him, and we should make amends after the fashion dictated by the mushiest of our sentimentalists. That at least would be justice to the criminal.

One might add other instances and ways in which Vold's conclusion leads to the contradiction of accepted facts or pretty generally accepted opinions. This is the method of reductio ad absurdum, and the conclusion might be made to look quite "absurd" in this manner. But it is not the best way. Let us consider the evidence on which the conclusion is based.

The first consideration is the very high percentages of feeble-mindedness among delinquents reported by the early mental testers, and the rapid decline of these percentages in later reports to less than half the first figures. Vold with others rightly criticizes the psychologists for this result. He concludes rightly also that the difference between the earlier and later reports from mental tests is due to the fact that "The methods of measurement of intelligence have changed during this period", but in the end maintains that adults score lower I.Q.'s than children do on mental tests with norms derived from the performance of children, and that therefore the percentage of adults

falling below a given low I.Q. is much larger than the percentage of children who do so. He attempts to prove this difference between the I.Q.'s of children and adults by a comparison with the army test results.

My chief difference with Dr. Vold concerns the validity of this comparison. But before we take up this issue other matters should be cleared up about this difference in the percentages of feeble-minded among delinquents that the earlier and later mental testers reported. Goddard's explanation is quoted that the difference is due to the fact that the earlier mental testers classed everyone as feeble-minded who earned a mental age of twelve or less on the tests, and that this dividing line is too high. This, of course, is not an error in the test results, but an error in their interpretation, for which the psychologists cannot be alone held responsible. Many others help to determine at what intelligence level the line shall be drawn between feeble-minded and normal. The chief function of the psychologist in this matter is to determine the grades of intelligence in different people. The psychologist did not fix this dividing line. He merely determined the mental levels below and above this line that had already been drawn on the basis of other than mental test evidence. The psychologists, however, did make an error for which they alone were responsible, and which Vold does not mention. They overlooked the fact that the original Binet scale and earlier revisions did not extend up high enough to give correct mental ages for adults near average, and also used a small number of tests that made them more unreliable than later scales. The original Binet scale had only three tests at the thirteen year level and none beyond, four tests at twelve, five tests at eleven, four tests at ten. On such a scale persons with a true mental age of thirteen would hardly average a mental age of twelve. At a true mental age of twelve and a half most would get mental ages below twelve. Also because of unreliability due to small number of tests in the scale many with a true mental age of fourteen even would get a mental age of twelve or less on the scale. This defect alone accounts for the low average mental age of adults and the large number scoring at twelve and below. The earlier revisions of the original Binet of 1908 were Binet's in 1911, Goddard's in 1911, mine in 1912, and the Yerkes' Point Scale in 1913. None of these removed this defect in any material degree. The only Binet revisions that did so were Terman's in 1916, and mine in 1922. At this point let me quote Sutherland's summary given in Vold's paper. It gives the median percentage found feeble-minded among institution delinquents, for a number of studies during successive periods.

Period	Median Percent
1910-1914 -----	51
1915-1918 -----	28
1920-1924 -----	21
1925-1928 -----	20

This shows distinctly that immediately after the first adequate mental test scale appeared the high percentage of delinquents found feeble-minded promptly disappeared. A few years later, when the older scales had been pretty well discarded for the new this median percent remains about the same. And this has been in the main true of later results up to date, where the same mental age or I.Q. dividing line between feeble-minded and normal had been used.

This brings us to the important conclusion that the large difference in mental test results reported from different surveys of delinquents in institutions, and of which Vold makes so much, no longer exists. But the percentage of delinquents found feeble-minded is still several times as large as found among school children. At this point Vold objects to applying the same mental test standards to adults as to children. He says, "The assumption that average adult intelligence is around I.Q. 1.00 as measured by these tests seems to me exceedingly questionable and not supported by good information." He then goes to the Army test results to show that adults score lower I.Q.'s than do children on the same tests. His case, therefore, stands or falls, with what the Army tests can prove on this point.

Vold accepts the assumption, first, that the Army draft represented a fair cross-section of the general population, as regards intelligence, and second, that the Army mental test results have the same reliability as have our later Binet test scales. I disagree emphatically with both these assumptions.

There were three calls for registrations during the war. Recruits from the first two only got into service since the third Registration was made only two months before the Armistice. The first two yielded 10,679,814 registrants. There were five ways of entering the service.

1. By commissioning as an officer	230,287
2. By appointment as a field clerk	3,294 (approximate)
3. By voluntary enlistment before registration	562,760
4. By voluntary enlistment after registration	948,545
5. By drafting, or drawing by lot	2,548,930

The fifth was called the draft, and the Army test results used to show the frequency of different grades of intelligence in the Army were limited to the draft. I shall attempt to show, first, that this draft could not possible have been anywhere near a fair sample of the Army or general population as regards intelligence, and, second, that the unreliability of the mental tests used, together with the former, more than accounts for the high frequency of feeble-mindedness shown in the Army test results.

The number entering military service by voluntary enlistment and the draft by the above five methods was 4,411,753. Sixty-six percent of these were drafted, and thirty-four percent entered by voluntary enlistment before and after registration. It is stated that nearly all of the voluntary enlistments become officers. It is extremely unlikely that any feeble-minded were among these 1,744,886 voluntary enlistments. 562,760 enlisted before any were drafted. Again, 65 percent of all registrants were excluded by deferment and exemption. A study of the classes that were thus excluded can hardly lead to the conclusion that the remaining men left to be drafted represented a fair cross-section of the general population. The excluded comprised the following classes:

1. All engaged in Municipal, County, State and Federal government

2. All in the U. S. Mail Service
3. All ministers and Divinity students
4. All regarded as essential in industry and farming
5. All who had dependents requiring their support
6. Certain aliens claiming exemption, and alien enemies
7. Men with certain physical defects and disease
8. Men with certain nervous and mental disorders
9. Mental defectives

I believe it is fair to assume that there was a negligible number, if any at all, of feeble-minded in the first five classes. The remaining classes comprised only 16 percent of all exclusions. The sixth class alone, the aliens, comprised 13 percent, and this is the most likely of the remaining classes not to include feeble-minded. The total number of mental defectives rejected was only 24,514, and these together with the men of draft age already in institutions for the feeble-minded comprised less than two tenths of one percent of the registrants.

If, in a word, we assumed that only normal men were in Class 1, commissioned as officers, Class 2, appointments as field clerks, Class 4, voluntary enlistments after registration, and all deferments and exemptions, except those exempted as feeble-minded, we would have the following numbers of normal men among the 10,679,814 registrants who did not get into the draft of 2,548,930 men.

Class 1.	230,287	
Class 2.	3,294?	(Includes some duplicates)
Class 4.	948,545	
Deferment & Exemption	6,948,758	- left after subtracting 24,512 feeble-minded
Total	<u>8,130,884</u>	

This leaves 2,548,930 men for the draft, to which the army test results were limited. Assuming that there was the same percentage of feeble-minded, that is, with M.A. not over 11 years or I.Q. .69 with 16 as the maximum divisor among the registrants as I found among school children, gives 566,030 feeble-minded that went into the draft, excepting the 24,512 exempted feeble-minded which leaves 541,516. This number 541,516, makes 21 percent of the draft feeble-minded. The main sample of the army test result gave only 17.6 percent feeble-minded, that is, not over M.A. 11 yrs.

We have arrived at this figure by making some assumptions that may be questioned to some degree. But we have so far considered only the effect of the selective service draft on the frequency of feeble-mindedness in the draft that resulted. There were other factors that increased the apparent number of feeble-minded in the draft over the true number. Let us turn now to the Army tests, and how the results were secured on the men examined. The data from which an attempt is made to determine the frequency of different grades of intelligence in the Army is based on groups specially selected for this purpose. They were limited to the white draft proper. They included 93,965 men given group tests, Alpha or Beta, plus the individual Yerkes Point Scale or the Pintner-Peterson Performance Scale, and 653 men given both group tests and the Stanford Binet. No good reason appears for denying that the first and larger group was a fair sample of the draft, although it constituted only two percent of the whole draft, making a considerable error a distinct possibility. The sample of 653 men given the Stanford-Binet as well as the group tests is questionable. These were obtained by sending instructions to the various camps "asking that unselected groups of the white draft be given both examination Alpha and examination Beta (Group tests), and, when possible, the Stanford Binet." The manner of getting unselected groups is not stated, and was evidently left to a dozen or more different people. 1,047 men were thus selected and given the group tests, but only 653, or 62%, were given the Stanford-Binet. The regular procedure on all other occasions was to give the individual Stanford-Binet test to those who did poorly on the group tests. There is some suggestion that this may have occurred here to some extent, making the 653 no longer an unselected group. In this group the percentage rated below a mental age of eleven years is 6.3 percent higher than in the larger group that were given only group tests.

But more important than the question as to whether the men tested were a fair sample of the draft, which itself was highly selected as we have seen, is the question of the nature of the tests used and how mental age scores were derived from the raw scores. I shall attempt to show (1) that there is a very considerable chance that large errors were made in converting the raw scores on the group tests into mental ages, and (2) that the number of tests used was so small that they must have been very unreliable, resulting in many men getting mental age scores much below their true intelligence.

The conversion of raw scores on the group tests into mental ages involved a highly technical procedure, a review of which is out of place here. Suffice it to say that it took over 250 ordinary sized pages to tell how it was done, by attempting to overcome numerous difficulties, resulting finally in the highly theoretical construction of three "conversion tables" and a special formula by means of which raw score could be converted into mental ages. One of the difficulties that we may mention for illustration arose from the fact the raw scores on these different scales gave quite different distributions on the same men, indicating that they evidently were not measuring intelligence equally well. For instance, with one scale the scores bunched heavily at the lower end, because it failed to discriminate between different levels of intelligence at this point. Those unfamiliar with mental test technique perhaps should be reminded that mental tests do not yield mental ages until after they have been given to children of different ages and age norms are established. An earlier form of the Army tests had been given to a number of school children, but apparently these results were not used. The conversion of the group test raw scores into mental ages had

then to be done by comparing the mental ages earned on the individual tests, which had age norms with the raw scores earned on the group tests by the same men, and in this way equating the two measures on the two kinds of tests. Since we have no way of knowing how much error resulted from this procedure we can do no more on this point than to protest to the assumption that there wasn't any error at all.

The question of the number and nature of the tests used is much more important. Most people who hear of the Army test results, including many who quote them in support of this or that point, believe that they are on a par with an individual examination with a scale of the Binet type. The unreliability of a group test score as compared with the result of an individual examination has been proven time and again. The following are some of the reasons for unreliability of the group test. Inability of the examiner to (1) control the subject as to motivation, (2) to insure attention, (3) to make sure that the subject understands directions, (4) to vary the tests actually given as to difficulty required by the abilities found in the subject. The last is related to the most important of all, the small number of tests that can be used in a group test scale. Present day Binet scales contain over a hundred separate tests arranged in order of difficulty. In the examination of any individual the part of the scale that is actually used is determined by the requirements of that particular individual. You cannot get the correct mental age of a person by using only tests that are adapted for a mental age several years lower or higher than that of the person examined. In the group test this is what is attempted for all the dull and bright in the group. There were only eight tests in the Army group test battery, Alpha. The same was true of the Beta Battery, which was used on non-English speaking men. I have not been able to determine from the data at what point in mental age the group test score was discarded and an individual test score substituted. There was evidently no very fixed rule about it, making it more difficult still to say how many were classed, say below mental age eleven, on a group test score alone.

When we turn to the individual examinations we find the same fault of small number of tests used. The individual tests used were the Stanford-Binet, the Yerkes-Bridges Point Scale, and the Pintner-Paterson Performance Tests. But each was used in much abbreviated form. The Yerkes-Bridges Scale had twenty tests, but was abbreviated to nine. The Stanford-Binet had six tests per age-group, but were abbreviated to two tests per age group. To give an individual Binet test examination unabbreviated takes from one to two hours. The abbreviated individual test examinations in the Army were made in twelve to fifteen minutes.

Now how does this affect reliability, and the number of normal persons who are likely to score below normal? To answer this question, let us remember that people, either normals or feeble-minded, differ very widely one from the other in the particular tests that are easy or difficult for them. If only a small number of tests are used some individuals will get much too low a mental age score, simply because the particular tests used happen to be tests that are especially difficult for those individuals. With more tests used the chances for a correct mental age are correspondingly increased. Suppose now we consider a hundred persons who with an adequate number of tests would all score with a mental age of eleven, and that we called all below mental age eleven feeble-minded. With an inadequate number of tests about half of them would score below eleven and would now be classed as feeble-minded, while most of the others would score above mental age eleven. This is what

we mean by unreliability of a test result due to a small number of tests used or inadequacy of sampling of the function tested. And this is what indisputably happened in the Army testing. To be sure this would happen with the actually feeble-minded as well as with the normal, so that many truly feeble-minded near the borderline would score as normal. But it could not happen as often with the feeble-minded as with the normal simply because there are not as many truly feeble-minded as there are normal. In this way the apparent number of feeble-minded in the draft was increased because of unreliable test scales. That increase could easily have added another several percent to the number classed as feeble-minded. And with this addition to the twenty-one percent, caused by the elimination of two thirds of the normals from the draft, the Army test results are more than explained. They agree substantially with expectation based on the frequency of feeble-mindedness found among school children. The following table brings together the essential data for comparison. It assumes for convenience that all adults with a mental age not over eleven years, and all children whose final mental age will not exceed this at age sixteen are feeble-minded. The first line gives 5.3 percent of public school children found feeble-minded, from a survey of twenty-six Minnesota towns. The 17.6 is the percentage with a mental age not over eleven as found in the main sample of the Army draft that were given group tests and abbreviated individual test scales. The 23.9 is the percentage in the Army draft found not to exceed mental age eleven in a special group given the abbreviated Stanford-Binet tests. All the other figures in the table give the percentage found feeble-minded by our individual test examinations in Minnesota institutions, as reported in the Bienniums of 1926 to 1934. These figures are for the new admissions during each biennium, except the first figure, 42 percent, for the Prison which is the result of the initial survey of the whole institution population. There is also a slight difference in the basis of classification for the institutions on the one hand and the public schools and Army data on the other. The data from the institutions is based on an I.Q. below .75, when 15 is used as a maximum divisor. This corresponds to a mental age of 11.1 years instead of just 11 years. The discrepancy is quite too small to affect comparisons.

Schools	Army	Army	Prison	Ref. Women	Ref. Men	Tr.Sch. Boys	Home Sch. Girls	Average age	Biennium
	93,965	653							
5.3	17.6	23.9							
			42	36	25	24	26	31	'24-'26
			45	21	21	21	19	25	'26-'28
			36	33	21	18	25	27	'28-'30
			34	34	18	23	23	27	'30-'32
			32	38	14	14	26	23	'32-'34
Average			38	30	20	20	25		

The examinations in the public schools and in the institutions were all made with the same and unabbreviated Binet scale using eight tests per age group and taking from one to two hours for an examination. The Army examinations were made with the Alpha or Beta group tests and one of three different individual test scales taking from twelve to fifteen minutes per examination.

My contention has been that feeble-mindedness is from four to eight times as frequent among inmates of corrective and penal institutions than it is in the general population, in accordance with these results from mental test surveys in the public schools and institutions. I have shown that when the facts and circumstances about the Army test results are taken into account they are in agreement with these findings. Vold by accepting the Army results at their face value, as though they were based on a fair sample of the general population, and as though they were obtained with tests and procedures on a par with other tests with which they are compared, has concluded that feeble-mindedness is no more frequent among delinquents in institutions than in the general population. To arrive at this conclusion it has been necessary for him to maintain that adults score so much lower on mental tests than school children do that the apparent frequency of feeble-mindedness among adults is increased four to eight times the frequency found among school children. Our results in the different institutions alone disprove that contention. At the time of examination the age of the inmates of the Training School for Boys, and of the Home School for Girls averages near 16. At the Reformatory for Men, the average age is around 22. At the Reformatory for Women it is between 30 and 35, and at the State Prison it is between 35 and 40. The results do show a higher frequency of feeble-mindedness among the last two and oldest groups. This may be the effect of age, or due to the fact that they are more truly criminal, or to the fact that they do not cooperate as well in a mental examination as the younger cases do in the other institutions. But one would hardly contend that in the other institutions, where the average age is around sixteen, age could have already been a handicap in the examinations. Psychology has plenty of evidence to prove that mental deterioration as measured by mental tests does not begin before twenty.

REFERENCES

Second Report of the Provost Marshal General on the Operations of the Selective Service System. Psychological Examining in the United States Army. Memoirs of the National Academy of Sciences. Vol. XV.