McKinley, J. C., Hathaway, S. R. & Meehl, P. E. (1948). The Minnesota Multiphasic Personality Inventory: VI. The K scale. *Journal of Consulting Psychology*, *12*, 20-31.

#012

THE MINNESOTA MULTIPHASIC PERSONALITY INVENTORY: VI. THE K SCALE[1]

BY J. C. MCKINLEY,[2] S. R. HATHAWAY AND P. E. MEEHL
UNIVERSITY OF MINNESOTA

THE K SCALE of the Minnesota Multiphasic Personality Inventory (MMPI) was developed in an attempt to correct the scores obtained on the personality variables proper for the influence of attitudes toward the test situation. The rationale of the approach as well as the empirical procedure employed in deriving K has been presented in a previous publication [7] and will not be dealt with here except very summarily. The present paper is to be read as a sequel to the original and aims chiefly to present norm data on K for various groups, an improved technique for applying K statistically, and certain miscellaneous observations such as its effects on the validity and intercorrelations of the other scales of MMPI.

The K scale was derived by studying the item response frequencies of certain diagnosed abnormals who had *normal* profiles. It was here assumed that the occurrence of a normal profile was suggestive of a defensive attitude in the patient's responses. The response frequencies were contrasted with those from an unselected sample of people in general ("normals"). The differentiating items were then scored so that a high K score would be found among abnormals with normal curves, whereas a low score would be found in clinical normals having deviant curves. In this operational sense, it can be said that a high K score is indicative of a defensive attitude, and a low K score suggests unusual frankness or self-criticality ("plus-getting"). The extremes of defensiveness and plus-getting may be called "faking good" and "faking bad" respectively.

The earlier procedure for applying K was one of subjectively correcting profiles on the basis of K score. Thus, a given borderline curve would be "under-interpreted" if K was considerably below the mean, since the examinee would be presumed to have achieved a bad curve because of his plus-getting tendency. If the same profile occurred in the presence of an elevated K, the clinician would assume that the curve ought to be "over-interpreted," since the examinee showed evidence in his high K of having been defensive.

In the following presentation we will first give the more practical data referent to the routine use of K. Following the description of the determination of K correction factors and specific data on validity we will return to the more general facts bearing on clinical interpretation integrated with the whole profile.

The original method of using K was admittedly vague and inspectional, and would require considerable experience on the part of the individual clinician. It was clear that the influence of the K

[1] Supported by graduate research grants from the University of Minnesota Graduate School.

[2] Doctor McKinley's name appears here in honorary recognition of the fact that his last research work before he became disabled was in connection with the development of the K scale.

factor upon scores was not the same for all MMPI variables, so that the optimal interpretation of the personality scales proper on the basis of a given K deviation varied. It is obvious that the amount of experience required to make a satisfactory use of K in profile interpretation would be very great, even assuming that the clinician would be able subjectively to record, retain, and analyze the welter of impressions with reference to the nine personality components. For this reason, it seemed that a more rigorous and objective procedure for taking account of the K score would be desirable.

Since high K scores represent the defensive or "fake good" end of the test attitude continuum, the most obvious approach to the problem is to add K (or some function of K) to the raw score on each personality variable, i.e. increase the score in the direction of abnormality. Thus, a psychopath who is very defensive in taking the test is presumed to have attained a lower raw score on the Pd scale than he "should" have, i.e., than he would have had he been less defensive. This defensiveness will also tend to reflect itself as a high K score. The obtained score on Pd should accordingly be corrected by adding some amount, the amount added being dependent upon the degree of defensiveness present as indicated by K. The problem is simply one of determining the optimal weight for the K factor with respect to any given scale, taking a linear function as an adequate approximation for practical purposes.

Our first attempt was crude in that it treated K as what may be called a "pure" suppressor, whose only contribution lay in its correlation with the noncriterion components of the personality variable [5, 6]. In a preliminary study of the Hs scale, using an unusually carefully selected group of diagnosed hypochondriacs, the Hs score was increased by a fraction of K proportional to the regression weight of Hs on K among the normals. In other words, in place of Hs alone we now were using the *residual* of Hs regressing on K, i.e., that part of Hs which is K-independent. This procedure is inexact since it assumes that K itself is uncorrelated with the dichotomous criterion, and also because it neglects the correlation of Hs with K among the abnormals. In spite of this crudeness, it was encouraging to find that the corrected Hs score now enabled us to detect 89% of the hypochondriacs as contrasted with about 70% of the same sample using Hs alone. This separation was achieved on a test group which had not entered into the derivation of either Hs or the K-weight, and involved no increase in the number of false positives among normals (about 5% in both cases).

The desirability of taking account of the correlation of K with the personality scales both among normals and abnormals, as well as any differentiating power of its own which K might have on certain sorts of cases, suggests the use of the discriminant function for determining the optimal weight. In the present problem, the variances among normals and abnormals were not always alike, nor was it convenient to restrict our analysis to the usual case of equinumerous groups. We experimented with a modification of the discriminant function which added variances rather than sums of squares, but decided to reject this also for the following reason: The region in which differentiation is clinically most important is around 60 to 80 T-score. There is little or no basis at present for interpreting the personality scores which are below the mean. All methods which are based upon maximizing the ratio of the variance of criterion group means to some type of pooled variance *within* groups will be taking account of the entire distribution. This results in a K-weight based upon information which there seems to be no reason to include. The skewness of MMPI variables and the obvious doubts one might have as to the influence of K at different points of the distribution led us to determine the optimal weight by a study of a more restricted region, within which refinement was of greatest consequence. It is unfortunate that this decision entails procedures which are mathematically inelegant and in sore need of analytic justification but we have not been able as yet to devise acceptable alternatives. It is hoped that the procedure now to be described will seem reasonable, and that others will attempt a formally simple solution and will study the sampling distribution of the test employed. In the present case, there is reason to suspect that a general maximizing solution is impossible without making assumptions regarding distribution form which are empirically inadmissible.

Consider a given personality variable, represented in deviate score form by x, where the deviation is from the mean of normals. Let the K deviate score be represented by z. Let λ be an arbitrary weight, whose optimal value is to be determined. *Optimal value* refers here to the λ which achieves the best differentiation between a criterion group of abnormals diagnosed as having the abnormality in question (e.g. hypochondriasis) and a sample of unselected normals. In other words, we are here considering the personality variables singly, by specific diagnosis, rather than "abnormals" as a whole. Then the deviate *corrected* score on the given abnormal compenent is

$$y = x + \lambda z$$

Let us now restrict our attention to the cases scoring above the mean of *normals* on y, i.e., consider only cases such that $x + \lambda z > 0$. We now define a sum of squares for those abnormals whose corrected score is above the *normal* mean. That is, for cases such that $x + \lambda z > 0$, we define a sort of "half sum of squares,"

$$S\,S_a = \Sigma_a (y)^2 = \Sigma_a (x + \lambda z)^2$$

The same quantity is computed for the normals,

$$S\,S_n = \Sigma_n (y)^2 = \Sigma_n (x + \lambda z)^2$$

The ratio of these two sums of squares, which we shall call the *differential ratio,*

$$\frac{S\,S_a}{S\,S_n} = \frac{\Sigma_a (x + \lambda z)^2}{\Sigma_n (x + \lambda z)^2}$$

is then taken as an index of the degree of differentiation achieved by a given value of λ.

It can almost be seen by inspection that a straightforward analytic solution for the optimal λ cannot be carried through by maximizing this ratio, since the number of cases involved in numerator and denominator will occur in the resulting derivative and will itself fluctuate with the choice of a λ in a manner that cannot be known without special specifications of the joint distribution of x and z. Even if special assumptions are made, such as normal bivariate surface and equal correlation for normals and abnormals (neither being true in this sort of material), the solution of the problem presents serious mathematical complications. We hope to be able to make further progress in this direction and invite more mathematically competent readers to attack the general case. We fell back upon a straight trial-and-error method. We assigned arbitrary vaules of λ (= .1, .2, .3, .4, etc.) and for each of these values we distributed y for normals and criterion cases separately. The ratio SS_a/SS_n was then calculated for each of these λ values, and these ratios plotted as a function of λ. A smooth curve was drawn by inspection through the plotted points, and a rough maximum was estimated therefrom. Where several different samples of abnormals were available, such curves were drawn separately for each, in the hope of having more confidence in the estimated maximum on the basis of agreement in curve "trend."

One further qualification needs to be mentioned. Since squares emphasize extreme deviations, and in view of what has been said above concerning "clinically important range," it was felt desirable to limit the influence of extreme deviations upon the ratio. Therefore, after the distribution of y for a given λ had been obtained, all scores of the normal and abnormal groups lying above three standard deviations on the basis of a given $x + \lambda z$ normal distribution (corrected T score of 80) were arbitrarily reduced to that value. A change in λ which produced further elevations of abnormals already at three sigma would therefore not result in further improvement in the differential ratio. It is possible that four sigma should have been chosen instead, since recent work on pattern analysis in differential diagnosis among abnormals suggests that elevations above three sigma may be important. In fact, we would not be prepared to vigorously defend the use of this restriction at all.

The graphical method used gave opportunity to observe the behavior of the differential ratio as λ was varied, and to check the degree of disparity with other indicators of separation. In general, it was found that the λ which maximized the ratio tended to agree fairly well with that selected by such measures as per cent abnormals above the top decile of normals. Research on a different problem suggests that the d.r. gives results similar but not identical with the critical ratio. In the present study, the λ's finally chosen were sometimes based upon compromises between the curve maxima of d.r. for various criterion groups, as well as counting measures of overlap.

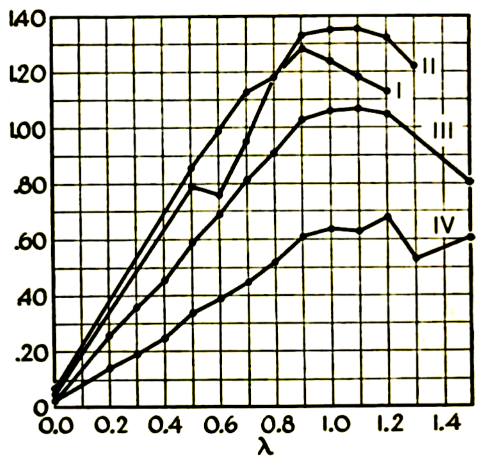

Fig. 1. Values of the "differential ratio" as a function of the size of the K-weight (λ), for the Sc scale.

Figure 1 shows typical data on the d.r. as applied to Sc + λ K. Groups I and II are composed of 25 and 28 males diagnosed schizophrenia and groups III and IV of 24 and 14 female cases respectively. There were some minor differences in the clinical constitution of the four groups but since the curves were similar in maximum points these differences can be disregarded. From these data we chose the λ weight for Sc to be 1K.

TABLE I

THE K WEIGHTS OF THE SCALES AFFECTED BY THE K CORRECTION

Hs + .5 K
Pd + .4 K
Pt + 1.0 K
Sc + 1.0 K
Ma + .2 K

Table I gives the K-weights which were finally adopted by these procedures. It must be emphasized that these weights are optimal, within our sample, for the differentiation of largely in-patient psychiatric cases of full-blown psychoneurosis and psychosis from a general Minnesota "normal" group. For other clinical purposes it is possible that other λ-values would be more appropriate. Thus, it seems likely that for the best separation of "maladjusted normals," such as those which abound in a college counseling bureau and would be formally diagnosed in a psychiatric clinic as *simple adult maladjustment*, other weights might be better.

The mode of applying these weights has been described already in the supplementary manual for the MMPI published by The Psychological Corporation. This manual contains a set of tables to be used in making the K-correction, and new test blanks are also available to be used with K. Briefly one determines the weighted K-value by referring to the table, which is based directly upon the proportions just cited. Thus, in correcting Hs for K, one begins by determining .5K either mentally or from the table (K here is the *raw* score). This quantity is then added to the original *raw* score on Hs, to yield Hs + .5K. This sum is called the *corrected raw score* on Hs. This corrected raw score is then entered in a second table of Tc (corrected T scores). This T table is of course based upon the mean and SD (on general normals) of the quantity Hs + .5K. Similar processes are involved in the case of the other scales.

It will be noted that only five scales receive a K-correction. The scales D, Hy, and Pa are uncorrected. (Mf was not studied in this respect.) It may seem paradoxical that in the original article on K, we found only three scales on which K would "work." This finding was based on a crude test using a single cutting score and a chi-square analysis. In the case of Pt, the sample studied originally was very small and significance hard to attain. The present K weights are to be accepted as the best indication of the "improvement" achieved by K, rather than the chi-squares cited previously.

We see then that Pt and Sc are most affected by K, Hs and Pd mediumly, Ma least and D, Hy, and Pa not at all. It is interesting to speculate upon this hierarchy. It will be recalled that the scale D already contains a correction scale, called Cd [4]. Hy already contains the K-factor in the Hy-subtle items [7, 8]. While Pa has no explicit correction scale, almost one-fourth of the Pa items are of the Hy-subtle type (eight being actually "0" items on the scoring key). These items, with the indicated response scored for Pa, are as follows:

B-54 "My mother or father often made me obey even when I thought that it was unreasonable." (F)

D-46 "Most people inwardly dislike putting themselves out to help other people." (F)

D-50 "Most people are honest chiefly through fear of being caught." (F)

D-52 "I think most people would lie to get ahead." (F)

D-53 "I think nearly anyone would tell a lie to keep out of trouble." (F)

D-54 "Most people will use somewhat unfair means to gain profit or an advantage rather than to lose it." (F)

D-55 "The man who provides temptation by leaving valuable property unprotected is about as much to blame for its theft as the one who steals it." (F)

G-50 "Some people are so bossy that I feel like doing the opposite of what they request, even though I know they are right." (F)

H-9 "I tend to be on my guard with people who are somewhat more friendly than I had expected." (F)

TABLE II

The Effect of the Final K Corrections on Test Case Groups Contrasted to a Standard Sample of 200 Normal Cases.

The values given are the per cent of cases at or above the given T score points.

T	Hs 200 Normals	Hs 101 Hs Test Cases	T	Hs + .5K 200 Normals	Hs + .5K 101 Hs Test Cases
70.0	3	59	70.0	5.5	74
69.8	5	62	70.3	5	72
65.0	10	69	62.9	10	89
T	Hy 200 Normals	Hy 101 Hs Test Cases	T	Hs 200 Normals	Hs 74 Hy Test Cases
70.0	4	64	70.0	3	38
67.5	5	74	69.2	5	42
63.3	10	79	65.0	10	55
T	Hs + .5K 200 Normals	Hs + .5K 74 Hy Test Cases	T	Hy 200 Normals	Hy 74 Hy Test Cases
70.0	5.5	54	70.0	4	53
70.3	5	51	67.5	5	62
62.9	10	69	63.3	10	66
T	Pd 200 Normals	Pd 89 Pd Test Cases	T	Pd + .4K 200 Normals	Pd + .4K 89 Pd Test Cases
70.0	5	52	70.0	3.5	55
70.0	5	52	67.5	5	65
65.0	10	65	62.7	10	76
T	Pt 200 Normals	Pt 36 Pt Test Cases	T	Pt + 1K 200 Normals	Pt + 1K 36 Pt Test Cases
70.0	6.5	42	70.0	4	61
71.5	5	40	68.5	5	67
67.2	10	47	64.0	10	67
T	Sc 200 Normals	Sc 91 Sc Test Cases	T	Sc + 1K 200 Normals	Sc + 1K 91 Sc Test Cases
70.0	4.5	31	70.0	2	59
69.0	5	31	64.0	5	69
62.5	10	43	61.2	10	75
T	Ma 200 Normals	Ma 89 Ma Test Cases	T	Ma + .2K 200 Normals	Ma + .2K 89 Ma Test Cases
70.0	8	62	70.0	2.5	65
66.3	5	72	65.7	5	74
61.8	10	79	63.1	10	84

If one calculates the percentage of "O" items for each of the eight personality scales (excluding Mf), the proportion of such items per scale is Hs = 0%, Sc = 3%, Pt = 4%, Ma = 15%, Pd = 16%, Pa = 20%, D = 27%, Hy = 33%.

These figures at least suggest that the proportion of zero items per scale tends to be negatively associated with the K-weight found to be optimal. One way of looking at this finding is to say that scales which are more *subtle* are less subject to distortion by such test-attitudes as K, and hence cannot be improved much by application of a K-correction. It cannot be decided on present evidence whether this is the correct view rather than the view that the subtle items, although not derived as suppressors, already contain "suppressor" components for the obvious items.

THE EFFECT OF THE K CORRECTION ON VALIDITY AS RELATED TO DIAGNOSIS

Table II gives an idea of the diagnostic effect achieved by the K correction. The cases designated "test cases" were not always clear cases of the given diagnostic category but represented patients who were noted by the psychiatric staff as being at least in part characterized by traits belonging to the category. Hence it is probably fair to assume that the percentages of these cases lying above the three given T values are smaller than would be true of more carefully selected patients. The 200 standard sample normal records used as reference were made up of 100 males and 100 females from the general normative files who were specially selected to be representative of the whole population.

For Hs and Hs + .5K, the data are given on both Hs and Hy test groups. The figures for these groups as distributed by Hy are also included. (See also Table III.) One may compare not only the Hs with Hs + .5K but also Hs + .5K with Hy. It is apparent from Table II and from correlational data that the addition of K to Hs makes it act more like Hy. This could be predicted from the communality of K and Hy-subtle [7]. In terms of the group data of Tables II and III, one is justified in using both scales. As was argued in an earlier publication [8] clinical evidence is at present in favor of the continued use of both scales because they are complementary when operating in the individual case. For example, Table III indicates that the joint use of both scales results in the identification of more hypochondriacs and hysterics than would the use of either separately. We hope soon to publish further data relative to the clinical significance of the two scales used together.

TABLE III

COMPARISON OF THE ACTION OF HS + .5K AND HY ON TEST CASES DIAGNOSED PSYCHONEUROSIS, HYPOCHONDRIASIS AND PSYCHONEUROSIS, HYSTERIA

	Hs Test Cases Per Cent with T score 70 and above	Hy Test Cases Per Cent With T score 70 and above
On Hs + .5K alone	16	9
On Hy alone	6	12
On both scales	58	41

The gains for Pd + .4K over Pd are most marked at the 5th and 10th percentiles. This results from a flattening of the frequency curve for the normals in the range of 60 to 70 T-score. We have already tended to interpret Pd as having clinical significance at around T = 65 when it appears as a clear "spike" or when certain other values (especially the neurotic triad) are below 50. The above data probably add justification to this interpretation.

The increased validity of Pt + 1K is a function of both increased normality in the frequency curve for normals and relatively higher scores for the test cases. The Pt + 1K is likely to be more clearly a "clinical" scale than was the Pt. We have pointed out [7] that Pt is a rather good measure of the K factor and one would expect partial removal of this variance to result in a remainder "purer" for the real clinical component. Good clinical data on psychasthenia relatively independent of schizophrenia are difficult to obtain and we can give no further evidence at this time.

Sc was never a very satisfactory scale in terms of the number of schizophrenic patients identified, although when it is elevated Sc is quite valid. When Sc + 1K is used, a very gratifying improvement is apparent. These gains with a K correction are from all standpoints the best of the five scales.

The improvement of Ma + .2K over Ma is not great but if the effect upon the frequency curve for normals is combined with that on the test group, it is definitely worthwhile to use the correction. Ma is the most common single deviate score in both high and low directions. Among the profiles of unselected normals, Ma occurs as a "peak" score more often than does any other scale, and it also occurs as a lowest score more often than any other. This is presumably a statistical consequence of the fact that Ma correlates with the other scales less than they tend to correlate among themselves. Ma probably has more independent clinical significance than any other single scale. These facts add to the importance of any gains in validity.

THE GENERAL INTERRELATIONSHIP OF K AND K CORRECTED SCALES

Table IV shows the means and standard deviations for various groups. We attribute no certain significance to the variations that can be observed in these statistics. The normals designated in this table are the general normals that have been described elsewhere in publications on the MMPI as a reasonably satisfactory cross section of Minnesota residents. While there are several possible sex difference trends as seen in the separate age groups, a grand compilation of all these normal males contrasted to all of the females shows no appreciable differentiation.

TABLE IV

THE K MEANS AND STANDARD DEVIATIONS OF VARIOUS GROUPS

	No.	Sex	Mean	Sigma
Normals age 16-25 inc.	115	F	12.61	4.96
Normals age 16-25 inc.	73	M	13.79	5.27
Normals age 26-35 inc.	153	F	12.82	5.21
Normals age 26-35 inc.	105	M	12.41	5.85
Normals age 36-45 inc.	105	F	10.41	4.60
Normals age 36-45 inc.	69	M	12.49	5.14
Normals age 16-45 inc.	373	F	12.08	5.07
Normals age 16-45-inc.	247	M	12.84	5.64
Mixed Psychiatric	372	M	14.57	5.85
Mixed Psychiatric	596	F	14.34	5.21
University	50	M	16.10	5.15
University	50	F	15.66	5.01
University (Drake, Wisconsin)	379	F	15.58	4.20
High School (Capwell)	78	F	14.96	5.46
Reform School (adolescent) (Capwell)	88	F	12.77	4.99
Reformatory (adult) (Capwell)	84	F	14.18	4.86
Graduate Electrical Engineers (Minneapolis-Honeywell)	100	M	16.72	4.19
Miscellaneous Employed (American Airlines)	100	F	15.38	6.05

The two groups referred to as "mixed psychiatric," included all diagnoses observed in the psychiatric unit and are not necessarily typical of a psychiatric hospital of the usual type. Many of the patients presented behavior problems of types that would not be committed to an institution for the insane and in general the group would be a borderline group between the obviously psychotic and the normal. The moderate rise in the means for these groups is chiefly contributed by the psychopathic personality and criminal individuals who would make up about 20% of the whole number. University students have a relatively higher mean as contrasted to general normals of their age range. An interesting point is evident in the means for the Capwell

[1] girls. Here the reform school cases obtain a higher mean than otherwise similar adolescents in high school. This low mean is contradictory to the tendency that we observed in adult offenders which is illustrated by the reformatory women whose mean score is somewhat higher than the general norm. It is of interest in this connection that a K correction slightly decreased the differentiation of the Capwell cases from their matched partners when one used the Pd scale as a discriminator. We have no explanation at present for this finding.

The largest mean that we have observed was obtained from the graduate electrical engineers. These men were studied during the war and were mostly around 30 years of age. They were exempted from military duty in order to carry on aviation research and at the time of testing were applying for special airplane control testing at high altitude. The final group of miscellaneous employed was obtained from a sample of airline employees most of whom were college graduates or had several years of college work. These were in more skilled clerical or minor administrative type positions.

We have described elsewhere [7] experiments in which ASTP men and several other groups were asked to fake good and bad profiles on the MMPI. In these experiments half the class faked a good or bad profile and the other half took the Inventory in a supposedly honest way. At a subsequent session of the class, the roles of these two groups were reversed. All of the subjects were naive in regard to personality inventories and in regard to the Multiphasic in particular.

This procedure afforded a check upon the action of F and L as well as K. In brief, it was found that F was very efficient in distinguishing faked bad records but L was not at all effective in detecting a faked good record among the men and was only moderately effective with women. We at first presumed that the failure of L for men was in part due to the relatively obvious items of which L is composed.

Among 48 student nurses asked to fake a good record, 16, or 33%, obtained a raw score L greater than or equal to 7 (T score greater than or equal to 60) in contrast to only one out of 48 when the same girls took the test with a supposedly honest attitude. If a raw score L greater than or equal to 6 (T score greater than or equal to 56) is used, these figures become 54% identified as faked for the faked records, as contrasted to 10% "false positive" among the honest records. This finding accords with our clinical experience that it is profitable to begin interpretation of L at $T = 60$ or even lower [7].

When we turn to the K distributions for these two groups, the most interesting findings are that the mean K score for the 48 nursing students taking the Inventory "honestly" is 18.3, standard deviation 3.80 and the corresponding statistics for 107 ASTP men are 19.8 and standard deviation 4.10. These two means correspond to general normal T values of about 61 and 63 respectively. These means are definitely larger even than the means of college students in general as given in Table II. Some factor seems to have operated on these two experimental groups to produce an unusually high average value of K when they were supposedly taking the test with an honest attitude.

As might be expected, when data were obtained from the 54 of the ASTP men faking a bad profile, the mean K values shifted markedly downward. The statistics for this group of faked bad data are a mean of 8.1 and standard deviation 4.04. The mean corresponds to a normal T value of 41. We have no corresponding statistics for women. K is in this case equally sensitive with F in differentiating the faked bad profiles. By contrast, the average K score for the 53 ASTP men who attempted to fake a good score was very little different from their normal mean as given above. The mean of this group's faked good records was 20.2 and the standard deviation 3.66. The mean would correspond to a normal T of 64. This result was similar to the finding among the student nurses who obtained a mean K score of 19.7 and a standard deviation 3.90 on faked good records. The normal T score for this mean is about 64. It should be kept in mind that when the ASTP men attempted to fake a bad profile the resulting profiles were very severe, differed to a remarkable extent from the individual's "honest" profile and could be recognized as invalid from the profile form alone. In contrast again to this, neither the men's nor the women's faked good profiles could readily be distinguished in any consistent way from their honest profiles, nor from the ordinary profiles of normal persons in general. The obvious experiment in which one would take a group who had deviant profiles and ask them to attempt to fake good was not performed. Further evidence on the behavior of K and F in the "fake bad" situation can be found in a recent article by Gough [2].

In consideration of these data, it seems justifiable to postulate that in these experiments the differentiation of the faked good profiles by the use of K is impossible because the "honest" was already in some sense faked good. The evidence that the "honest" represented something already related to faking can be derived from the fact that the "honest" means of both these groups were more than a standard deviation elevated in terms of the general normal mean statistics for K and at least half a standard deviation in T score above the means obtained from other college data. This elevation over the three other means given in Table II would be even greater in standard scores on the basis of the college data considered as norms. Some unidentified factor related to K must have operated in the experimental situation where the faking data were obtained. This latter assumption would be more certain if the rise in the means had not been observed from such different groups as student nurses and ASTP men. It is possible, however, to link these two groups provisionally in one significant element. Both the nurses and the men were under impulsion not to jeopardize in any possible way their continuance in the war-related programs that they were following. This pressure would contrast to the situation of the miscellaneous college students who were tested either before or after the war and probably in even greater degree to the attitudes of the general MMPI norm groups that provided the normative statistics for the T table of K.

The nearest approach to data on faked good scores as obtained from persons with *initially* deviant profiles is embodied in some incidental data obtained from our records where psychiatric hospital patients repeated the Inventory for one reason or another. By searching the duplicate records, we were able to find a few cases where patients had taken the Inventory twice and where the K raw score for the second test was four or more points higher than that for the first test. Most of these patients had originally deviant profiles. The obtained differences are not worthy of statistical analysis but all scales show a tendency to decrease in T score under these conditions. The most marked changes occurred on Hs, D, Pt and Sc. Naturally, since these patients were not asked to fake a good score, the finding yields only presumptive evidence.

CORRELATIONAL DATA

The test-retest correlation of K is available on two groups. For a group of 85 high school girls (Capwell data) retested at an interval of 110 to 410 days the correlation was .72. For a group of miscellaneous normals retested after four days to one year the correlation was .74. It is of course impossible to say to what extent these coefficients are to be viewed as indicators of "reliability."

A second question that may be raised regards the effect of the K-correction upon the intercorrelations of the other MMPI scales. Table V gives the correlation coefficients for the same group with and without the K-correction having been made before correlating. The intercorrelations for the original scores are indicated in ordinary type, while the corresponding coefficient upon the same sample after making the K-correction is indicated immediately to the right of the originals in bold-face. These coefficients are based upon a sample of 100 normal males, all college graduates, employed as engineers in an industrial concern (Honeywell cases of Table IV). We see that some of the correlations are raised by the K-correction, that others are lowered, and that this is true whether they are considered in the absolute or algebraic sense. The increases preponderate over the decreases. Inspection suggests that the greatest shifts occur in the case of pairs of scales one of which suffers a considerable K-correction and the other none (e.g. Hy and Pt). We are not prepared to give any special interpretation of this table and include it here only for the sake of completeness.

SUMMARY AND CONCLUSIONS

Specific arguments and data are presented establishing the rationale of using the K factor as a suppressor on certain MMPI clinical scales. Five scales seem to be improved by the correction, as indicated by increased correspondence between scores and clinical status. The scales Pt, Sc, Hs, Pd, and Ma receive K-corrections of varying amounts. The scales Hy, D, Mf and Pa are not so treated nor is it established that the K-score should be taken into account subjectively in evaluating them. A new statistic was used to determine the K correction factors. This statistic, called the differential ratio, is described as appropriate to establishing maximal differentiation between two distributions with emphasis upon the region of their over-

TABLE V
CORRELATIONS AMONG SCALES BEFORE AND AFTER K-CORRECTION. LATTER ARE IN BOLDFACE. N = 100 EMPLOYED MALE COLLEGE GRADUATES

	Hs		D		Hy		Pd		Mf		Pa		Pt		Sc		Ma
Hs																	
D	33	**33**															
Hy	33	**65**	31														
Pd	28	**37**	36	**37**	25	**47**											
Mf	27	**17**	26		21		28	**22**									
Pa	08	**22**	17		37		15	**24**	33								
Pt	47	**47**	28	**45**	–17	**38**	33	**43**	41	**44**	16	**45**					
Sc	51	**59**	20	**26**	–03	**51**	31	**44**	45	**32**	19	**39**	72	**66**			
Ma	25	**04**	–07	**–05**	–13	**00**	32	**01**	30	**34**	06	**12**	50	**26**	53	**21**	
K	–25		08		53		–09		–08		22		–65		–46		–42

lap.

Normative statistics on the distributions of K for various groups are presented.

The chief finding of interest here is a tendency for college and college-educated persons to deviate in the upward direction between one-half and one standard deviation. It was suggested in the original article on K that this difference is chiefly a function of socio-economic status.

Some evidence was presented to show that K behaves in the expected manner when persons attempt to fake a "bad" profile, although the corresponding effect in faking "good" was not demonstrated on any experimental group. Some clinical support for this latter effect has been found.

The addition of K had a variable effect on the intercorrelations of clinical scales. There seems to be some indication that the optimal amount of K-correction for a given clinical scale is inversely related to the proportion of "subtle" items the scale already contains.

It is suggested that the K-correction should be made routinely by users of the MMPI and that old records should be scored and redrawn if any research or validation study is to be carried on.

References

1. Capwell, D. F. Personality patterns of adolescent girls: II. Delinquents and non-delinquents. *J. appl. Psychol.*, 1945, 29, 289-297.
2. Gough, H. G. Simulated patterns on the Minnesota Multiphasic Personality Inventory. *J. abnorm. soc. Psychol.*, 1947, 42, 215-225.
3. Hathaway, S. R. and McKinley, J. C. *Manual for the Minnesota Multiphasic Personality Inventory.* New York: Psychological Corporation, 1943.
4. Hathaway, S. R. and McKinley, J. C. A multiphasic personality schedule: III. The measurement of symptomatic depression. *J. Psychol.*, 1942, 14, 73-84.
5. McNemar, Q. The mode of operation of suppressant variables. *Amer. J. Psychol.*, 1945, 58, 554-555.
6. Meehl, P. E. A simple algebraic development of Horst's suppressor variables. *Amer. J. Psychol.*, 1945, 58, 550-554.
7. Meehl, P. E. and Hathaway, S. R. The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *J. appl. Psychol.*, 1946, 30, 525-564.
8. McKinley, J. C. and Hathaway, S. R. The Minnesota Multiphasic Personality Inventory: V. Hysteria, hypomania, and psychopathic deviate. *J. appl. Psychol.*, 1944, 28, 153-174.