

## **Profile Analysis of the Minnesota Multiphasic Personality Inventory in Differential Diagnosis**

**Paul E. Meehl**

*University of Minnesota*

A personality test may be employed in several kinds of clinical situations. These include: overall differentiation of normals from abnormals or persons predisposed to abnormal developments as in "screening" in the military, industrial, educational, or general medical out-patient situation; differential diagnosis among abnormals; prognosis; evaluation of changes and results of therapy; and the assessment of certain components for other than strictly diagnostic purposes such as the detection of important paranoid trends in a reactive depression even though diagnosis presents no problem.

The present paper presents preliminary data on the use of the Minnesota Multiphasic Personality Inventory (MMPI) with respect to differential diagnosis, with secondary findings upon the subject of overall identification of "abnormals" from people in general. Since MMPI has been described elsewhere (Hathaway & McKinley, 1943), we may merely state that this device is a structured personality test which yields scores on nine components of abnormality, namely Hs (Hypochondriasis), D (Depression), Hy (Hysteria), Pd (Psychopathic deviate), Mf (Masculinity-femininity), Pa (Paranoia), Pt (Psychasthenia), Sc (Schizophrenia), and Ma (Hypomania). In addition, there are four scores which indicate "validity," in the sense that they attempt to detect test-records which for reasons such as confusion, language difficulty, or non-cooperation cannot be accepted as adequate samples of the patient's verbal behavior. These scores are ? (Cannot say), L ("Lie"), F (carelessness and misunderstanding), and a recently devised suppressor called K (Meehl & Hathaway, 1946). For the details of function and interpretation of these validity indicators, the reader is referred to previous articles. In the present study, the scale Mf (Masculinity-femininity), has been excluded from consideration throughout, so that there are only eight personality components involved. All of the scores are expressed as T-scores, the general normal sample having a mean of 50 and a S.D. of 10.

The purpose of the present study was to evaluate MMPI as used in the differential diagnosis of three main categories of hospitalized psychiatric patients: psychosis, psychoneurosis, and "conduct disorder." Gough (1946) and Schmidt (1945) have stressed the importance of considering the "pattern" or configuration of the profile in addition to the elevation of single scores. An elevation on a single component, even if it is the highest or "peak" score of the profile, does not imply

that the patient should be so diagnosed. For example, the most frequent peak score on abnormal profiles of all sorts is D (Depression). It is clinically known that many different kinds of psychiatric difficulties involve degrees of depression, and the test reflects this fact. Again, a peak of 75 on Sc might suggest a schizophrenic picture, whereas if it occurs together with markedly elevated scores on the neurotic triad (Hs, D, Hy) and a Pt of, say, 85, it may better be taken to indicate a psychoneurosis with poor prognosis (Harris & Christiansen, 1946). It must be emphasized that the patterning of a profile cannot be neglected in the case of structured tests any more than we would think of interpreting one determinant column of the Rorschach without considering anything else.

As yet, these configurational criteria on MMPI have not been adequately treated in the literature. Locally, the Minnesota group have tended to form more or less crude clinical judgments and global impressions based upon accumulated experience. The articles by Schmidt and Gough have contributed materially to the objectification of procedure, although neither of these investigators published results in the form of percent correct identifications for clinically diagnosed groups, a kind of treatment which is in many ways more meaningful than establishment of significant differences between central tendencies (Dvorak, 1935, p. 19). Furthermore, in both of these articles the similarity of "psychosis" to "severe psychoneurosis" in MMPI profile is too close for comfort, a drawback of MMPI which has been informally reported by a number of military clinical psychologists through personal communications.

In the present investigation, an attempt has been made to determine the approximate accuracy of a very rapid, inspectional diagnosis from the MMPI profile alone, using the more or less poorly defined criteria which have so far seemed valuable clinically. Naturally, it is not suggested that the profile be used in this way, but we want to know how much the test can contribute entirely on its own when so used. Because of the fact that recently hospitalized cases were not diagnosed independently of MMPI, it was necessary to utilize old cases, before July 1941, on whose response sheets the present scales had been subsequently scored. At that time, the MMPI had not been published and was still in process of development. The only scales which appeared on the profiles then in use were H (a relatively less valid, uncorrected key for hypochondriasis) and D (Depression). For all practical purposes, it may be assumed that the clinical diagnoses made on these cases at that time were almost wholly unaffected by the presence of these scores on the chart. Of course, none of the present "pattern" criteria could have been employed at that time; further, knowledge of and confidence in the test were negligible among the psychiatric staff.

The procedure of blind diagnosis was as follows: profiles of male abnormals were leafed through in the order of their appearance in the files (roughly chronological). Any profile showing a ? (Cannot say) or L (“Lie”) score as great as 70 was recorded as “invalid,” except that if any abnormal score reached a standard score of 80, an elevated L score was ignored, since defensive lying could hardly be the reason for such a positive elevation. F was allowed to reach a raw score of 16 ( $T = 80$ ) before the profile was considered invalid. The terms “valid” and “invalid” are used hereafter to indicate the acceptability of the profile as an adequate measure in terms of ?, L, and F, and have no reference to the question of accuracy of identification. When it had been decided that a profile was “valid” by these criteria, it was classified as either *normal* or *abnormal*. Actually, of course, it was known that all of the cases were abnormal, so that the criteria of classification had to be made wholly objective and hence more rigid than would be the case in practice. Profiles were called abnormal under the following four conditions:

1. Any of the eight components showed  $T \geq 90$ .
2. Any of the eight components showed  $T \geq 80$ , *unless*  $K < 40$ .<sup>1</sup>
3. Any of the eight components showed  $T \geq 70$ , *unless*  $K < 50$  and  $L < 60$ .
4. Any of the eight components showed  $T \geq 65$ , *unless*  $K < 65$  and  $L < 60$ .

It can be seen from the above criteria that the classification into normal or abnormal is a matter of spotting the highest T-score, then reading to the right to see if the restrictions on K and L throw the profile into one group or the other. The profiles consist wholly of MMPI scores and a code number, so that there is no other source of information in making the diagnosis.

Application of these criteria to 294 profiles from our general population male sample<sup>2</sup> yields 10% “invalid” records on the basis of ?, L, and F scores set as above. Of the records which can be accepted as valid, there are 9% indicative of abnormality by the criteria, which may be considered the upper limit of “false positives.” Actually, of course, an unknown proportion of these false positives are profiles of persons who, although not under psychiatric care at the time of testing, were at least as psychiatrically deviant as some of the hospitalized abnormals. The figure 9% is to be contrasted with the 3%

---

<sup>1</sup> The scale K is a correction scale or suppressor variable which may be used to correct for certain test-taking attitudes which tend to invalidate a record. If the K score is low, it indicates that the testee was overly self-critical and obtained spuriously abnormal scores, hence is probably less deviate than his profile suggests. If K is high, it indicates a defensive tendency, and suggests that the profile is too low—a more subtle form of the old L scale. For further discussion see Meehl & Hathaway (1946).

<sup>2</sup> The data are for males only.

to 5% found previously for single scales. It is to be kept in mind in what follows that the differentiations achieved among the hospitalized abnormals occur at the expense of almost 1 in 10 among the normal population. The remainder of this paper deals only with the differentiation among actual abnormals.

When a profile had been classified as abnormal by the above criteria, a quick inspectional classification was made using three categories. The three employed were psychosis, psychoneurosis, and "conduct disorder." The last category is used to cover cases diagnosed constitutional psychopathic inferior, psychopathic personality, criminalism, alcoholism, except psychoses or deterioration, simple adult maladjustment, or "primary behavior disorder" such as the adolescent conduct problems not otherwise classified. The criteria employed in this subdivision of abnormal records were intentionally vague and subjective, since it was this sort of inspectional judgment which was to be evaluated. No "computations" of any sort were performed on the scores. In general, the criteria insofar as they were explicit, were those described by Schmidt and Gough, and such personal impressions as the examiner had acquired from considerable clinical work with MMPI. Psychosis was suggested by markedly elevated profiles, high F, Sc greater than Pt, Pa or Ma markedly elevated, the "psychotic" (right-hand) end of the curve reaching the level of the "neurotic" (left-hand) end, or a distinct spike on D, with the Hs and Hy scores on either side falling far below the D. Psychoneurosis was suggested by a less elevated profile, lower F, Pt greater than Sc, Pa and Ma not much elevated, the neurotic triad clearly elevated more than the rest of the curve, and the three scores of the triad closer to one another. Conduct disorder was suggested by elevations on Pd, Ma if not too high and especially with secondary peak at Pd, neurotic triad low except for some Hy, psychotic end running about 60. The examiner restricted himself to 10 seconds per profile in making his decision, and in most cases the judgment was made in less than five seconds. After having made the classifications, these were compared with the diagnoses of the psychiatric staff. All cases were eliminated in which the staff diagnosis was indicated as highly questionable or based upon insufficient study or cases of organic C.N.S. disease or feeble-mindedness. The actual composition of the abnormal group as subsequently determined was as follows: Psychosis, 57 cases (Schizophrenia 26, Manic-depressive 21, Paranoid condition 8, and Involutional melancholia 2); Psychoneurosis, 53 cases (Hypochondriasis 14, Hysteria 13, Reactive depression 9, Psychasthenia 7, Anxiety state 5, Mixed or unspecified 4, Neurasthenia 1); and Conduct Disorder, 37 cases (Psychopathic personality 21, Psychopathic personality pathological sexuality 8, Alcoholic 5, Behavior disorder 2, Adult maladjustment 1).

Of this entire group of 147 clinical abnormals, 25 (17%) invalidated their records on the basis of the validity indicators. Seventy-

eight (53%) were correctly called abnormal, while the remaining 44 cases (30%) were (erroneously) classified as normal. The following table represents the data in various convenient breakdowns:

Table 1  
Classification by Profile Inspection of 147 Records  
of Hospitalized Abnormals (Male)

A. Percentages based upon all 147 records			
	Called Abnormal	Called Normal	Invalid Record
Total group ( $N = 147$ )	53%	30%	17%
Psychotics ( $N = 57$ )	60%	21%	19%
Neurotics ( $N = 53$ )	47%	36%	17%
Conduct disorder ( $N = 37$ )	51%	35%	14%
B. Percentages based upon the 122 valid records (based on ?, L, F scores)			
	Called Abnormal	Called Normal	
Total group ( $N = 122$ )	64%	36%	
Psychotics ( $N = 46$ )	74%	26%	
Neurotics ( $N = 44$ )	57%	43%	
Conduct disorder ( $N = 32$ )	59%	41%	
C. Percentages based upon the 78 cases called abnormal			
	Called Abnormal	Called Normal	Called Conduct Disorder
Psychotics ( $N = 34$ )	56%	29%	15%
Neurotics ( $N = 25$ )	24%	68%	8%
Conduct disorder ( $N = 19$ )	16%	16%	68%

From these tables, we see that in employing a criterion of abnormality which holds our false positives down to one in ten among general normals, we are able to detect only about half of the known abnormals (Table 1, A). This figure is not quite fair to the test, however, since those with invalid records would *not* under these conditions be erroneously classified as normal, but would either be requested to take the test over again with more precautions or their profiles disregarded. If we confine our attention to records in which the validity indicators are satisfactory, we find that about two-thirds of the abnormals can be identified (Table 1, B). It should be pointed out that the disappointingly large proportion of apparently invalid records among the abnormals (about one-sixth of all the records) is in part due to the time at which these tests were administered. At that time patients were allowed to invalidate their testings by sorting large

numbers of the cards into the "Cannot say" category, sorting the cards at random, and so on. More systematic supervision now eliminates many of these uninterpretable profiles.

Setting up a contingency table for the 78 cases correctly classed as abnormal, we obtain a chi-square of 34.016, which with 4 d.f. is highly significant ( $P < .001$ ). This corresponds to a contingency coefficient of .55, with the upper limit possible for a  $3 \times 3$  table being .82.

In comparing the accuracy of identification for the three diagnostic groups, we shall consider only the valid testings, since the percentages of invalid records differ insignificantly among the three. While three-fourths of the psychotics were identified as being abnormal as contrasted with between one-half and three-fifths of the neurotics, a test of significance in proportion of false "normals" in the three diagnostic categories fails to show a significant difference (Chi-square 3.233, 2 d.f.,  $P > .14$ ). This being the case, most of the further subgroup differences in identification were not statistically analyzed. Mere inspection of Table 1, C, however, would suggest that the chief confusion occurs between neurotic and psychotic curves, rather than between either of these and the class of conduct disorders. Once having correctly classed a profile as being abnormal, the probability of its being thrown into the appropriate one of the three categories is about two in three.

Detailed inspection of the table of actual-real classifications does not indicate much because of the small numbers of cases in various subcategories. In the case of the psychoneuroses, however, inspection suggests that some clinical subgroups are more likely to show apparently "normal" profiles than are others. The differences in proportion called abnormal were tested by grouping the cases into four classes: hypochondriasis, hysteria, psychasthenia, and all others, and running a chi-square test on the resulting  $4 \times 2$  table. This chi-square was barely significant at the 5% level (Chi-square 8.303, 3 d.f.,  $P < .046$ ). Inspecting the table for the source of the differences, we find that 11 of the 13 hypochondriacs were identified as abnormal, as compared with only half of the ten hysterias, and only one of the six compulsives. It has been recognized for some time that the Pt scale is relatively ineffective clinically, and the use of K as a suppressor for Pt in this crude way tends to increase the false negatives by leading to an under-interpretation of profiles because K is highly correlated (negatively) with Pt. More detailed treatment of the actual subcategory tables is not warranted by the numbers involved.

### Summary and Conclusions

The adequacy of the MMPI in differential diagnosis employing a rapid, inspectional method of pattern analysis of profiles was investigated by making "blind" diagnoses from records of 147 hospitalized psychiatric cases into three major categories of psychosis, psycho-

neurosis, and conduct disorder. The criterion was the clinical diagnosis of the psychiatric staff, made at a time when the present scales of MMPI, with one exception, were not yet in existence. The findings were as follows:

1. Setting up arbitrary criteria for the overall distinguishing of normal from abnormal persons, we find that about 1 in 10 persons from the general population sample is called abnormal (false positive).

2. Approximately 2/3rds of actual abnormals are identified as such by these criteria, if we exclude records obviously invalid on the basis of the validity indicators  $\alpha$ , L, and F.

3. Of the abnormal cases identified as abnormal, about 2/3rds are placed in the appropriate category of the three employed. The contingency coefficient for the agreement between blind diagnostic grouping and the actual diagnosis is .55.

4. There is a suggestion that some varieties of abnormality are more readily identified than others. Hypochondriasis is fairly easily identified, whereas hysteria and psychasthenia are less so.

In general, while the discriminations achieved are very much better than chance in the statistical sense, especially considering the fact that *no* skilled clinical time is involved in giving or scoring the test and less than 10 seconds was used here in "interpreting" it, it must be admitted that the proportion of false classifications is considerable. Two developments can be expected to reduce materially this margin of error: first, the more mathematically precise utilization of the suppressor K; second, the greater formalization of pattern interpretation.

*Received July 6, 1946.*

### References

- Dvorak, B. J. (1935). *Differential occupational ability patterns*. Employment Stabilization Research Institute Bulletin No. 8, Vol. 3. Minneapolis: University of Minnesota Press.
- Gough, H. G. (1946). Diagnostic patterns on the Minnesota multiphasic personality inventory. *J. clin. Psychol.*, 2, 23-37.
- Harris, R. E., & Christiansen, C. (1946). Prediction of response to brief psychotherapy. *J. Psychol*, 21, 269-284.
- Hathaway, S. R., & McKinley, J. C. (1943). *Manual for the Minnesota multiphasic personality inventory*. New York: The Psychological Corporation.
- Meehl, P. E., and Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota multiphasic personality inventory. *J. appl Psychol*, 30, 525-564.
- Schmidt, H. O. (1945). Test profiles as a diagnostic aid: The Minnesota multiphasic personality inventory. *J. appl. Psychol*, 29, 115-131.