

"NEUROTICISM" AND HANDWRITING

BY H. J. EYSENCK

Psychological Department, The Maudsley Hospital, London

INTRODUCTION

THE history of the scientific investigation of graphology, and of the relation between psychologists and graphologists, is marked by acrimonious debate and almost complete divergence of viewpoint. The psychologist has criticized the graphologist for his impressionistic methods, his failure to apply strict experimental and statistical checks to his findings, and the use of reasoning by analogy instead of properly scientific inductive-deductive methods. The graphologist, conversely, has criticized the psychologist for his "atomistic" tendencies, i.e., his method of correlating isolated graphological "signs" with temperamental variables, for his failure to devise experimental methods which would make possible the exercise of the particular skill claimed by graphologists, and for his failure to provide criteria acceptable to the graphologist against which his comparative success or failure could be measured. More recently, attempts have been made to bring the opposing points of view more closely together, by such devices as "matching" experiments (1, 2, 5), and the results of such work have on the whole been favorable to the graphologist in the sense that extra-chance success has usually attended his or her efforts.

Work of this kind, however, does no more than touch the fringe of the problem. In particular, it does little to settle the crucial question of the criterion. As long as the graphologist's diagnosis or assessment is validated against the usual type of self-rating, clinical opinion, character sketch, or other subjective estimate, the possibility remains that low validity coefficients are due to the imperfections of the criterion rather than failure on the part of the graphologist. Alternatively, when graphological assessments are validated against objective tests (other than intelligence tests), the question of the meaning and interpretation of the test scores arises, and cannot be dismissed by arbitrary naming of the function supposed to be tested—persistence, suggestibility, dominance, or whatever it may be.

EXPERIMENTAL PROCEDURE

In the present paper, an attempt is made to overcome in part the difficulties mentioned and to provide as the criterion a score of known reliability and validity. The variable assessed by the graphologist, and measured by the psychologist, was "neuroticism"; the writer has shown elsewhere how this dimension of personality was isolated by means of a large-scale factorial study and how a battery of tests was constructed in an attempt to measure this variable (3). An operational definition of the concept is given in this book and will not be repeated here. Details of all the objective tests used in the present study are also given there; they will be described only very briefly here.

The population tested consisted of 198 male army patients at the Southern Hospital; 105 of these were diagnosed as "psychoneurotic" and sent to the special neuropsychiatric section of the hospital, while the other 93 were surgical cases without any notation of "neurosis." This differentiation into neurotic and non-neurotic subjects will be referred to as "psychiatric diagnosis," or as the subjective criterion. This criterion is, of course, far from perfect; a few so-called "neurotics" referred by the Command Psychiatrist are not very neurotic, and several of the surgical cases in a hospital of this type are quite severely neurotic (compensation neuroses, hysterics, and the like). Nevertheless, there is little doubt that on the whole the two groups are differentiated with respect to the quality under consideration.

A battery of tests was given to these two groups of patients, made up in the main of tests which had previously been shown to correlate well with subjective estimates of neuroticism. This battery is described in detail in (4) and contained tests of Primary Suggestibility, Persistence, Personal Tempo, Perseveration, Fluency, Static Ataxia, Level of Aspiration, Rigidity, Manipulation, Dark Vision, and one Psychoneurotic Inventory. Scores of the seventeen tests used were inter-correlated, and correlations were also run

with the subjective criterion. A factorial analysis was carried out on these correlations, which showed the presence of one strongly marked general factor, identified with "neuroticism." The battery of tests was found to have a (corrected) split-half reliability of .75, to correlate with the subjective criterion to the extent of .73, and to have a theoretical validity of .87 (3, 4).

All the subjects, in addition to being subjected to this battery of tests, were required to write a short essay on the subject "My breakfast this morning," to which they were asked to append their signatures. They wrote on unlined, standard paper, using either their own fountain pens or a pen selected by themselves from three standard thicknesses. Free composition rather than copying was used as the possibility that, in copying, a person's temperament may show up less well could not be ruled out. The possibility that the content might have given the graphologist certain clues is infinitesimal; all patients had the same kind of breakfast, and the writer, in reading through the essays, could not find any difference in style, in choice of words, or in any other respect between the two groups.

The sheets containing the essays were thoroughly shuffled and handed over to the graphologist, Mrs. F., to be rated on a five-point scale of "neuroticism," ranging from "more normal than average," through "comparatively normal," "relatively mild indications of abnormality," and "definitely neurotic," to "extremely neurotic." Mrs. F. is a well-known continental graphologist, established in that profession, who has worked in this country for many years; she generously agreed to collaborate in this experiment. The results are obviously valid only for the particular graphologist who was taking part in the experiment. The system used by Mrs. F. is by and large a simple adaptation of that advocated and taught by Klages.

RESULTS

The total number of cases for which correlations were run between the two criteria and the graphological assessment was 176, distributed between the five graphological categories in the ratio of 25:37:44:46:24. A few cases had to be eliminated because indications of their ward number, or other identifying

signs, had been included in their essays.

The biserial correlation between the subjective criterion and the graphological assessment was $r_{bis} = -.02$; this correlation is clearly insignificant and indicates that there is no relation between the two variables. The product-moment correlation (corrected for coarseness of grouping) between the objective criterion and the graphological assessment was $r = +.21 \pm .07$; this correlation is significant, although rather small, and indicates that the two variables are definitely related to a slight extent.

The existence of a significant correlation between a combination of subjective and objective criteria and graphological assessment was established by means of the following procedure. Twenty-two neurotic subjects (subjective criterion) who also had very neurotic scores (objective criterion) were compared with 19 non-neurotic subjects (subjective criterion) who also had very non-neurotic scores (objective criterion); each person was given a score on the basis of the graphological assessment, from 1 (more normal than average) to 5 (extremely neurotic). The significance of the difference between the two groups on the graphological assessment scale was determined by means of student's t , which reaches the value of 2.16 ($P = .05$).

DISCUSSION

The results show clearly that, while the graphologist succeeded in diagnosing "neuroticism" from handwriting with better than chance success (objective criterion), the predictive accuracy of her diagnosis would be extremely low. It is worthy of note that the graphologist failed entirely to assess "neuroticism" when the subjective criterion was used. This failure may be contrasted with the success of the battery of objective tests in diagnosing "neuroticism" (subjective criterion): twelve tests showed significant correlations with the psychiatric diagnosis, some correlations being as high as .57. When it is further remembered that these objective tests are of short duration (2 to 5 minutes each), and that they can be combined and multiplied at will, while the graphological analysis is extremely time-consuming, the relative superiority of the tests as opposed to the analysis will perhaps become apparent.

From a methodological point of view, this contrast appears to the writer of great importance. The graphologist, when shown in actual experiment that his assessment is not very reliable or valid, often retorts that, even if this be so, yet no other psychological tests of temperament and character show even the small success which his own efforts evince, and that if only the criterion could be improved his success might be much more apparent. Yet the present study demonstrates that even with an imperfect criterion short, objective tests show much higher validity coefficients than does graphological analysis; this would seem to invalidate the claims of the graphologists. It would also seem to argue against the idiographic view that "global" appraisals of personality are superior to nomothetic or "atomistic" tests; handwriting analysis is a typical idiographic or global procedure, while the series of tests used in this study might be considered representative of the nomothetic approach.

While on the whole the present study has not resulted in conclusions which can be regarded as favorable to graphology, it would be unjust to forget that significantly extra-chance results were in fact obtained. This indicates that personality *does* reveal itself in a person's handwriting; it also indicates the desirability of improving by the usual scientific methods the interpretation of handwriting, and of establishing correlations between unitary personality trends and certain expressive qualities of movement.

No very clear reason emerges which might account for the fact that handwriting analysis correlates significantly with the objective criterion but fails to correlate with the subjective criterion. Two possible explanations spring to mind. In the first instance, the objective criterion may be the more reliable and valid one; that this may be so is indicated in the original analysis of the experimental data (4). Secondly, it is possible that there is a group factor common to all *motor* expressions of "neuroticism" which would be found both in the handwritings and in the battery of objective tests, but which would be missing in the subjective criterion. It is conceivable that both explanations apply.

As a last point it should not be forgotten

that the results reported here are based only on the analysis carried out by one handwriting expert; the possibility that others might have been more successful cannot be ruled out, although in our view it is not a very likely one. As the material collected in the course of this experiment appears to be unique in combining two separate sets of criteria, objective and subjective, and in making possible assessment of personality in a dimension securely established by experimental and statistical means, the writer believes that the cause of science would be served by making the handwritings available to any graphologist who believes that the conclusion drawn in this paper with respect to the relative failure of handwriting analysis in personality diagnosis is too severe. By thus submitting himself to the experimental test the critical graphologist would find an obvious way to disprove the conclusion indicated above.

SUMMARY AND CONCLUSIONS

One hundred and seventy-six handwriting specimens were judged by a graphologist on a five-point scale with respect to the writer's degree of "neuroticism." Two validating criteria were available, an objective criterion, consisting of the results of a battery of personality tests, and a subjective one, consisting of a psychiatric diagnosis. The two criteria showed a high degree of intercorrelation. The correlation of graphological assessment with the objective criterion was low, but significant ($r=+0.21$); with the subjective criterion it was insignificant ($r=-.02$). Possible reasons are given for this difference.

REFERENCES

1. ALLPORT, G. W., & VERNON, P. E. *Studies in expressive movements*. New York: Macmillan, 1933.
2. EYSENCK, H. J. Graphological analysis and psychiatry; an experimental study. *Brit. J. Psychol.*, 1945, 35, 70-81.
3. EYSENCK, H. J. *Dimensions of personality*. London: Kegan Paul, 1947.
4. HIMMELWEIT, H., DESAI, M., & PETRIE, A. An experimental investigation of neuroticism. *J. Personality*, 1946, 15, 173-196.
5. WELLS, F. L. Personal history, handwriting, and specific behavior. *J. Personality*, 1946, 14, 295-314.