

## Section of Psychiatry

President—Professor AUBREY LEWIS, M.D.

[November 12, 1946]

### **The Measurement of Personality.** [*Résumé*]

By H. J. EYSENCK, Ph.D. London

*Psychological Laboratory, The Maudsley Hospital*

BOTH terms in the title of this paper, “measurement” and “personality” need definition and discussion, as common usage of these words is rather indeterminate. Personality we shall define, following Warren (1934), as “the integrated organization of all the cognitive, affective, conative, and physical characteristics of the individual as it manifests itself in focal distinctness to others.” This rather inclusive definition distinguishes “personality” from such more narrowly circumscribed terms as “temperament” and “character,” which refer more specifically to the affective and conative organization of the individual respectively, from “intelligence,” which refers to the cognitive aspects, and from “constitution,” which refers to the physical aspects.

If we accept this definition, it becomes clear at once that we cannot in any intelligible fashion talk about the *measurement* of personality. Without going into a detailed discussion of the principles of measurement (Campbell, 1928, 1938; Scates, 1937; Thomas, 1942; Young and Householder, 1941; Gulliksen, 1946), it is apparent that measurement can only proceed along one dimension at a time; we can measure the height of a building, or its length, or its depth—we cannot measure all three at the same time.

This difficulty lies at the back of Henderson and Gillespie’s (1943) statement that “if it is doubtful what we measure with ‘intelligence’ tests, it is still more uncertain what we would try to measure if we tackled ‘emotions’ in a similar way”. In other words, while in the cognitive sphere we have succeeded to some extent in isolating one “dimension” of personality for measurement, there is still no agreement about the “dimensions” existing in the affective sphere, and consequently the very basis for any kind of measurement is absent.

Under those conditions, it behoves us to study the methods used in establishing the cognitive dimension, and to inquire whether the same methods might not with advantage be applied to the orectic sphere as well. Two main contributions may be discerned in the development of the science of intelligence measurement: one, the creation of a large number of objective, reliable tests validated against external criteria, such as judgments by teachers, parents, officers, psychiatrists, and so forth; and second, the elaboration of statistical methods, more particularly the method of factor analysis (Burt, 1940; Thomson, 1939; Guilford, 1937; Thurstone, 1935; Spearman, 1927), which ensure unidimensionality and internal validity to these tests.

In noting these two contributions, we find that both claim "validity" of testing as part of their achievement. This indicates that there are two different types of "validity". Validity is usually defined as a measure of the extent to which a test agrees with a criterion; e.g. the extent to which an intelligence test succeeds in measuring intelligence. But the example itself shows immediately that this definition is almost valueless; how *can* we know how well a test measures a quality, such as intelligence, unless we already have a perfect measure of that quality? True, we have a number of obviously imperfect measures of intelligence, such as teachers' ratings, success in school and college, or earning capacity, but as these criteria themselves are unequal, the question arises, which are we to choose? What criterion shall we use for choosing our criterion? In this way we become involved in an infinite regression, and our definition, so simple and obvious at first glance, is seen to be swallowed up by a metaphysical hydra.

As opposed to this "external" type of validity, psychologists have elaborated an "internal" validity, derived from the pattern of interrelations obtaining among groups of tests. Thus, fifty tests presumed to measure a great variety of aspects of intellectual ability are given to large numbers of subjects, the intercorrelations of the tests are calculated, and from the pattern or matrix of their intercorrelations certain "factors" are extracted whose validity does not depend on any outside criterion, but is derived from the whole "gestalt" of the original matrix. In this way, Spearman originally proved that there is one general factor, "g", which is common to all cognitive tests, in various proportions; in this way, we have learned since that there are a number of more specialized groups of abilities, common to some but not all tests, called verbal, arithmetical, visuo-spatial, æsthetic, mechanical and perceptual ability.

Superficially, the fact that we have two kinds of validity seems to face us with an impasse. The "intelligence" measured by the common-sense type of person who relies on external validity is not acceptable to the statistically minded psychologist, who is apt to raise awkward questions about the external criterion and its validity; the "g" measured by the statistical psychologist is said (by the non-statistician) to be a mere statistical artefact, without any real psychological meaning. Fortunately, there is a strong tendency for both types of validation to give the same answer to the question—which is the best test of intelligence? When we compare the correlations of a number of tests with an external criterion, we see that some tests correlate highly, others show low correlations. When we intercorrelate these same tests, and factor-analyse the resulting matrix, we find that some tests have high correlations with the resulting "g" factor, while others have low correlations. Now in such a situation it is usually found that the test that correlates highly with the outside criterion will also be the test that has a high factor saturation, and the test that shows a low correlation with the outside criterion also has a low factor saturation. In that way, we can identify our statistical factor, "g", with the psychological dimension, intelligence, and justly claim that we have succeeded in identifying and measuring this particular aspect of personality.

These considerations suggested that a similar process in the orectic field might lead to similar results; in other words, it appeared that the use of factorial analysis

in connexion with temperament, character, and constitution might lead to the isolation, and finally to the measurement, of the main dimensions in those fields. Working on that basis, a team of psychologists and psychiatrists spent some four years at Mill Hill Emergency Hospital, and at the Maudsley Hospital, carrying out large-scale factorial and experimental studies into those problems; preliminary results have been published in a series of papers (*see* References: Eysenck, with Furneaux, Halstead, Himmelweit, Rees, and Yap). More recently, a complete account of this work has been published in book form (Eysenck, 1946). Only the main conclusions can be presented here.

Using factorial techniques as "internal validity" criteria, and psychiatric judgments, personal history data, Service career, and similar data as "external validity" criteria, we found that just as there is one general factor dominant in the cognitive field, so there are discernible strong general factors in the fields of character and temperament as well, using our definitions of these terms as presented in the first paragraph of this paper.

In the field of character, or "conation", there appeared a very powerful factor which might be called neuroticism, maladjustment, neurotic constitution, lack of integration, or lack of will-power; this factor clearly distinguished (1) the person referred to a neuropsychiatric hospital from the person not so referred; (2) the neurotic patient judged by the psychiatrist to be severely ill from the patient judged to be less severely ill; (3) the "normal" person outside the hospital whose adjustment to Army, factory, or life was considered faulty from the person whose adjustment was more successful.

These findings were checked and counterchecked using a great variety of objective, reliable tests. The most successful of these tests was the "body sway" test of primary suggestibility, which was described in a paper read to this Society two years ago (Eysenck, 1943*a*). Another successful test was an objective version of the Rorschach ink-blot test, in which the subject has to select responses from a printed list, instead of providing original responses. A third type of test which gave adequate differentiation was a test of persistence, requiring the subject to maintain a certain posture involving fatigue of various muscles for as long as possible. Questionnaires, although less objective than the other tests in our battery, tended to give excellent results when given to suitable groups.

Other differences between persons situated towards the "neurotic" end of the dimension under discussion, and persons situated towards the "normal" end included the following: neurotics tended to be slightly less intelligent, although the difference was too small to be of any practical importance; they tended to have a markedly more leptomorphic body-build; they tended to score very badly on tests of dark-vision; they had a slow "personal tempo", low fluency, extreme perseveration (as opposed to lack in persistence) poor effort response, and a very uneven and irregular work-curve. On all these points, as well as on many others, there was good correspondence between external and internal criteria.

Whenever attempts were made to investigate the distribution of this factor, we found a roughly normal curve of distribution; this would appear to indicate that this factor may have a constitutional basis, and may be subject to multifactorial inheritance, like intelligence. Investigation of this possibility is urgently needed before any definite claims can be made as to the genotypical reality of our findings; factor analysis cannot give any but descriptive, phenotypical results, the possible causation of which must be investigated by other methods more directly concerned with the specific problem under discussion.

It is possible that the use of the term "neuroticism" in denoting this factor may be open to criticism, and indeed the multiplicity of meanings associated with

psychiatric terms such as "neurosis" makes them of doubtful value if applied to more specific findings with operational connotations. However, no other term could be found to obviate these difficulties, and in any case problems of nomenclature are of semantic rather than of scientific interest.

In the field of temperament, or "affection", there appeared a factor which seemed to be very similar in nature to Jung's extravert-introvert dichotomy. Among the neurotic patients investigated, this distinction pointed to a dimension ranging from one extreme, containing the more hysterical types of symptoms and reactions, to the other extreme, containing anxiety, depression, obsessional and other "dysthymic" reactions. These two groups of symptoms, the hysteric and the "psychasthenic" or, as we prefer to call it, the affective or dysthymic, are of course the prototypes of Jung's famous dichotomy, and thus our results strongly support his claims. It should be noted, however, that we regard extraversion and introversion not as types in the sense that a given person is either an extravert or an introvert; we merely regard these two concepts as the extreme ends of a normal curve of distribution, with the majority of people falling somewhere in between the extremes. Support for this view of a normal distribution of this factor comes from actually plotted distributions, all of which resemble the normal curve so familiar from work on intelligence tests.

Experimental measurement along this dimension is based on a variety of findings, involving a number of tests from various modalities. To begin with, there is strong evidence that the introverts tend to be leptomorphic in body-build, while the extraverts tend to be euryomorphic. Jones and Richter have shown that effort response is poorer in the introverts, and choline esterase secretion more pronounced. Intellectually, the introverts tend to score more highly; when equated for intelligence, introverts tend to give markedly higher scores on vocabulary tests than do extraverts. Introverts tend to carry out various tasks slowly, but accurately; extraverts tend to do them quickly and inaccurately. Introverts tend to have a high level of aspiration, to be very rigid in their demands of themselves, and to underrate their own performances; extraverts tend to have very low levels of aspiration, to lack the rigidity of the introvert, and overrate their own performances. These tendencies are exaggerated in the neurotic extravert (i.e. the hysteric) and in the neurotic introvert (i.e. the dysthymic); normal subjects are intermediate between these extremes.

Extraverts also tend to be differentiated from introverts by their æsthetic preferences, in terms of their sense of humour (both with respect to how amusing they find various types of material, and which type of material they find most amusing; on the whole, introverts find all types of material less amusing than do extraverts, but in the field of sexual humour in particular this difference is much more pronounced than elsewhere), and in terms of their æsthetic creativity, as shown for instance in the Lowenfeld Mosaic Test, where introverts tend to make compact, extraverts scattered designs. Questionnaires also show marked differences between introverts and extraverts, as do various other tests not enumerated.

It is important to clarify the relation of our introvert-extravert factor to Kretschmer's cyclothyme-schizothyme typology. It is our considered opinion, based upon a variety of experimental evidence, that these two typologies cannot justifiably be considered identical, and that in reality there is no relation between them whatsoever. In terms of experimental work, this conclusion follows from our demonstration that tests which had been shown by others to differentiate between cyclothymes and schizothymes did not discriminate at all between extraverts and introverts. The tests used included a variety of colour-form tests, reversal-of-perspective tests, and measurements of body-build. As already explained, we found that introverts (dysthymics, i.e. persons suffering from anxiety and reactive depression) tended to be leptomorphic in body-build, which would align them with Kretschmer's schizo-

thymes, while our extraverts (hysterics) tended to be euryomorphic in body-build. But in Kretschmer's system, hysteria tends to go with the schizothyme type, and depression with the cyclothyme type; consequently it does not appear that any identification of these two systems is possible. On colour-form tests, on which great differences usually become apparent between cyclothymes and schizothymes—the former being more colour-reactive, the latter more form-reactive—and on the reversal-of-perspective tests, in which also great differences usually appear—the cyclothymes giving fewer reversals than the schizothymes—no differences appeared between our extravert and introvert samples.

Our findings, while largely based on various neurotic groups, are also applicable, in our view, to normal groups; whenever we have had an opportunity of testing normal samples of the population we have found precisely the same differentiation in our test results as we found in our neurotic groups. As the total number of cases studied is quite considerable, approaching twenty thousand, we believe that our conclusions, while no doubt subject to many detailed criticisms and improvements, are perhaps a useful first approximation to the goal we set ourselves, viz. the isolation and measurement of the major dimensions of personality. This hope is strengthened by the fact that whenever it was found possible to check certain partial findings against the results obtained by other workers in the field, using quite different types of subjects (students, children, factory workers, &c.), considerable agreement was evident. Similarly, our main conclusions on the factorial side are not in contradiction to the results obtained in certain pioneer studies by Burt, Webb and others. A full discussion of the relation of our work to that of others will be found in "Dimensions of Personality" (Eysenck, 1947).

The picture of personality which emerges from these studies bears some similarity to a globe, or sphere. Any point on that sphere can be identified by reference to three axes constructed at right angles to each other, and penetrating the sphere. One of these axes is the familiar cognitive axis, labelled "g" or intelligence; the second axis is labelled "neuroticism", and may be identified with the conative side of personality; the third axis is labelled "introversion-extraversion", and corresponds to the affective side of personality. In actual fact, these three axes are not quite orthogonal, but show a slight degree of obliqueness; in other words, the angles separating them are not right angles, but only approach this particular structure. For instance, intelligence and neuroticism show a negative correlation of  $-0.30$ , corresponding to an angle deviating by 3 degrees from a right angle. A similar correlation is found between introversion and intelligence, again slightly tilting the two axes towards each other. But by and large these departures from orthogonality are too small to influence the general picture to any significant extent, and for practical purposes we may assume independence among our dimensions.

If we want to extend our personality sphere still further, so as to take in the physical aspect as well, we must add a fourth axis, at right angles to the other three; this addition of course necessitates a four-dimensional representation, and makes it impossible to visualize the resulting structure. Again slight departures from orthogonality must be admitted; leptomorph body-build correlating to the extent of  $0.30$  approximately with "neuroticism", and to the same extent with "introversion". It is not possible to say whether such a four-dimensional picture is adequate, or whether other dimensions will be required; only further research can answer this question.

The main usefulness of these studies, assuming that our general conclusions be accepted, will probably lie in two fields. In the first place, practical work in industrial psychology, vocational guidance, occupational selection, in clinical psychology, and in educational psychology, may benefit from the possibility of carrying out objective

measurement along the two dimensions indicated. In the second place, the provision of batteries of tests for such measurement will make possible a concentrated attack on such theoretical problems of outstanding importance as the influence of nature and nurture in the genesis of conative and affective traits, or the relation of personality variables to a great variety of psychological phenomena. Only by such further experimentation can the claims made here be proved or refuted, and our knowledge in the field of personality measurement be extended beyond the present narrow boundaries which have prevented its growth for too long.

## REFERENCES

- BURT, C. (1940) *The Factors of the Mind*. London.
- CAMPBELL, N. R. (1928) *An Account of the Principles of Measurement and Calculation*. London.
- , (1938) Symposium: Measurement and Its Importance for Philosophy, *Proc. Arist. Soc. Suppl.*, London, **17**, 121.
- EYSENCK, H. J. (1943a) *Proc. R. Soc. Med.*, **36**, 349.
- (1943b) *Lancet*, (ii), 362.
- (1943c) *J. Neurol. Psychiat.*, **6**, 22.
- (1944a), *Brit. J. med. Psychol.*, **20**, 100.
- (1944b) *Amer. J. Psychol.*, **57**, 406.
- (1944c) *J. ment. Sci.*, **90**, 851.
- (1945a) *Brit. J. Psychol.*, **35**, 70.
- (1945b) *Psychol. Bull.*, **42**, 659.
- (1947) *Dimensions of Personality*. London.
- , and FURNEAUX, W. D. (1945) *J. exp. Psychol.*, **35**, 485.
- , and HALSTEAD, H. (1945) *Amer. J. Psychiat.*, **102**, 174.
- , and HIMMELWEIT, H. T. (1946) *J. gen. Psychol.*, **35**, 59.
- , and REES, W. L. (1945) *J. ment. Sci.*, **91**, 301.
- , and YAP, P. M. (1944) *J. ment. Sci.*, **90**, 595.
- GUILFORD, J. P. (1937) *Psychometric Methods*. New York.
- GULLIKSEN, H. (1946) *Psychol. Rev.*, **35**, 199.
- HENDERSON, D. K., and GILLESPIE, R. D. (1943) *A Text Book of Psychiatry*. Oxford.
- HIMMELWEIT, H. T. (1946) *Brit. J. Psychol.*, **36**, 132.
- (1945) *J. Person.*, **14**, 93.
- REES, W. L. (1945a) *J. Neurol., Neurosurg., Psychiat.*, **8**, 34.
- , and EYSENCK, H. J. (1945b) *J. ment. Sci.*, **91**, 8.
- SCATES, D. E. (1937) *Psychometrika*, **2**, 27.
- SPEARMAN, C. (1927) *Abilities of Man*. London.
- THOMAS, L. G. (1942) *Psychol. Monogr.*, **54**, No. 3.
- THOMSON, G. H. (1939) *The Factorial Analysis of Human Ability*. London.
- THURSTONE, L. L. (1935) *Vectors of Mind*. Chicago.
- WARREN, H. C. (1934) *Dictionary of Psychology*. Boston.
- YOUNG, G., and HOUSEHOLDER, A. S. (1938) *Psychometrika*, **3**, 19.
- , — (1941) *Psychometrika*, **6**, 331.