Brit. J. Psychiat. (1970), **117**, 241–50

The Classification of Depressive Illnesses

By H. J. EYSENCK

As is well known, there has been a good deal of discussion about the unitary or binary nature of depression. Mapother (1926) and Lewis (1934) made a strong case for the unitary view on clinical grounds, Curran (1937) concurring. Kendell (1968) has reviewed the history of this argument; it is marred by confusion which has persisted through recent attempts to use statistical techniques of factor analysis and discriminant function analysis in an effort to find a more objective and empirical solution. It is the purpose of this brief note to draw attention to this confusion, to show how it has affected arguments of both adherents and opponents of the binary position, and to argue that the data are in fact in sufficient agreement to make possible a valid answer to both problems. It is suggested that the apparent disagreement between workers such as those of the Newcastle group (Kiloh and Garside, 1963; Carney, Roth and Garside, 1965) and the London (Maudsley) group (Kendell, 1968) is in fact quite irrelevant and is based on a misunderstanding of the statistical properties of factors, a misunderstanding apparently introduced in one of the first studies of this kind to be concerned with the problem of the classification of depressive illness, that by Hamilton and White (1959).

Much of the discussion referred to has been carried on in terms of the simple issue between unitarians and binarians—is it true, as Lewis (1938) said, that 'all the tables and classifications in terms of symptoms are nothing more than attempts to distinguish between acute and chronic, mild and severe; and where two categories only are presented, the one—manic depressive—gives the characteristics of acute, severe depression, and the other of mild, chronic depression'? This question can be translated into factor analytic logic quite simply and clearly: if Lewis is right, then intercorrelations between items thought on clinical grounds to be relevant to both the putative types of depression should give rise to a matrix of intercorrelations of rank one, i.e. should result in a *single*, general factor having *positive* loadings throughout (assuming that items are worded in such a way that a positive mark is given for the depressed response, and a negative, or zero mark for the non-depressed response). It is not fatal to this hypothesis that minor, relatively unimportant factors (in terms of their contribution to the total variance) should be found, as long as this major need for a single, large factor with entirely positive loadings was satisfied, and no other important factors, corresponding in their pattern of loading to the endogenous and reactive types of depression, were observed. If Roth were to be supported, again the factor analytic outcome can be set down which alone would carry conviction. In an unrotated solution we would expect a single general factor with positive loadings throughout, and in addition an equally large, or larger, bipolar factor, with positive loadings for those items characteristic of one type of depression, and negative loadings for those items characteristic of the other type. Rotation to simple structure, whether orthogonal (Varimax) or oblique (Promax) would lead to two independent factors having positive loading patterns, the one for endogenous items and the other for reactive items. In the case of Promax it would be feasible that even a slight and entirely random departure from orthogonality of these two factors (not smaller than corresponding to a correlation of $\pm$ ·02) would enable the programme to end up with a single factor, but this would have positive *and* negative loadings falling into a pattern corresponding to the two types of depressive illness.

It is not impossible that empirical data, collected from patients suffering from depression of one kind or the other, or differing in severity

of depression, might give an ambivalent and ambiguous answer. Thus factors of one kind or the other might be on the border-line of significance; there might be arguments about the respective size of the contributions different factors made to the variance; or the actual pattern of the items and their loadings might only partly correspond to the theoretical pattern expected on the binary hypothesis. In actual fact these difficulties have not arisen; all the protagonists have found a clear-cut positive answer to the question put—the unitary hypothesis is wrong, and the binary hypothesis is supported very strongly indeed. Kendell has made some comments suggesting that factors so extracted are subject to interpretation, and are in any case only statistical condensations of existing data; but this argument is not strictly applicable to this particular case. As Eysenck (1953) has pointed out, we must discriminate very carefully between two different uses of factor analysis, and the standing of the factors extracted: there are *factors suggesting a hypothesis*, and *factors supporting or disproving a hypothesis*. Kendell's objections apply only to the former; they are not relevant to the latter. (To be quite fair he does not in fact argue that his results are merely statistical artefacts, but the point is worth making because some readers might be misled into confusing these two uses of factor analysis.)

If there is any such universal agreement, why is it that Kendell can present such a strongly critical review of the Newcastle data and arguments (we shall deal with his criticisms in detail later), and arrive at what is at first sight a conclusion entirely opposed to theirs, namely that 'depressive illnesses are best regarded as a single continuum extending between the traditional neurotic and psychotic stereotypes'? To answer this question brings us to the second of the two problems mentioned above, and one not properly appreciated by the two warring groups. This problem is one which has been given much attention, both theoretically and empirically, by the writer, but which has not been considered very much by psychiatrists in recent years (although in passing, both Kiloh and Garside and Kendell do refer to it). Psychiatric diagnoses may be regarded from two points of view: either they are regarded as medical disease entities, i.e. categorical systems of classification, or they are regarded as points of intersection on a dimensional framework (Eysenck, 1969). An alternative way of expressing this is to say that we may view diagnoses as either *qualitatively* or *quantitatively* distinguished from each other. Malaria is qualitatively distinguished from typhoid, cancer from tuberculosis; these diagnoses are categorical classifications representing separate disease entities. Within each there may be differences in severity, responsiveness to drugs, and many other characteristics, but it is meaningless to think of them as lying on some continuum. When we consider such psychiatric diagnoses as psychopathy, hysteria, anxiety state, or obsessional-compulsive disorder, we may regard these in very much the same way, i.e. as specific, separate and qualitatively different disease entities, or we can regard them as resultants or combinations of a number of different dimensional categories. Thus an anxiety state may just be a person with high scores on neuroticism and introversion, a psychopath one with high scores on neuroticism and extraversion, and so on. A psychiatric diagnosis would thus refer, not to a distinct disease entity, but rather to a specific region in multi-factor space; patients in that region would have certain *resemblances* by virtue of their similar position on the relevant dimensions, but these would shade insensibly into other diagnostic categories as we change the position on one or other, or both, of these dimensions.

The dimensional approach accounts for many of the difficulties which have beset psychiatric efforts to construct a system of diagnosis, and to operate it; the arbitrary nature of the system in use, and the unreliability of the diagnoses made, even by experts of high standing, speak eloquently against the underlying hypothesis of categorical allocation, and in favour of some dimensional system. A great deal of work has been published from our laboratory dealing with the statistical methodology appropriate to the investigation of the categorical *v.* dimensional types of hypotheses (Eysenck, 1950), and much empirical material is available to demonstrate the superiority of the latter over

the former, both with respect to neuroticism and also to psychoticism (Eysenck, 1952a, 1952b, 1955, 1960, 1964; Eysenck and Claridge, 1962; Eysenck, Granger and Brengelmann, 1957; S. B. G. Eysenck, 1956; Devadasan, 1964). Using Eysenck's method of criterion analysis, as well as discriminant function analysis, it was found that both neuroticism and psychoticism were continuous variables, from the most neurotic or psychotic person through intermediate degrees of abnormality right down to the most normal; it was further found that these dimensions were independent of each other, as well as from extraversion/introversion. (Cattell and Scheier, 1961, have also demonstrated this independence in their own work.) Recent work with personality questionnaires has extended this work on psychoticism to normal groups, both adult (Eysenck and Eysenck, 1968a, 1968b, 1969) and children (Eysenck and Eysenck, 1969); the older work was carried out with laboratory tests of various kinds.

The relevance of this digression will become obvious when we turn to a second argument which has been presented in an effort to test the unitarian hypothesis. This argument relies on the unimodal or bimodal distribution of scores of unselected groups of depressive patients, and was first put forward by Hamilton and White (1959). Having shown by factor analysis that their data gave rise to two factors corresponding to endogenous and reactive depression (which they prefer to label retarded and agitated), they go on to say: 'Accepting the finding that the two types of depression differ, there are then two theories about the difference. One theory states that the two types are extremes of a single population, and the other states that they are two distinct, but overlapping populations. In statistical terms, this would be put in the form that the first theory would suggest a unimodal distribution of scores, and the second a bimodal distribution.' It is suggested that this argument is mistaken, and that its adoption by both the Newcastle and the London school is responsible for the confusion which has arisen. To clarify this confusion, let us consider the two sets of facts and theories which we have been discussing. In the first place we have the argument between unitarians and binarians; as pointed out above, this has been conclusively decided in favour of the binarians. In the second place, we have the argument between categorical and dimensional systems of diagnosis; as pointed out above, this has given rise to data strongly favouring the latter. But the argument from distribution (bimodal *v.* unimodal) is relevant, not to the first controversy, but to the second; it is this confusion which has bedevilled the discussion for so long.

The position may be stated as follows. There are four possible theories: (1) Unitarian and categorical. (2) Unitarian and dimensional. (3) Binary and categorical. (4) Binary and dimensional.* Mapother and Lewis would seem to hold the first of these positions; Freud and Kendell the second; the Newcastle group the third; and the present writer the fourth. Factor analysis is relevant to the decision between (1 and 2) *v.* (3 and 4); the results have decisively favoured 3 and 4.† We have thus a

* The possibility exists, of course, that in a binary model one of the two types of depression might be of the categorical, the other of the dimensional type. In view of the evidence quoted this does not seem likely, and in any case this would not alter in any important way the discussion given. Thus in strict logic we ought to add two more combinations to those given in the text.

† 'Decisively' is perhaps too strong a term, in view of the occasional negative findings, such as those of McGonaghy, Joffe and Murphy (1967). Nevertheless, it remains true to say that the very large population studied by Kendell, coming as it does from the stronghold of the unitarian doctrine and evaluated by registrars trained in this particular approach, agreed very closely in the nature of the factors produced with that studied by the Newcastle workers, and that Hamilton's original work also gave closely similar results. No thorough review of all the work done in the field has been attempted here, in view of the fact that several of the writers named give good summaries of prior work, but the evaluation given above represents an estimate of all the reported studies, not only of those quoted. It might of course still be objected, not without reason, that statistical processes, however advanced, cannot differentiate between 'true' and 'biased' information; the 'factors' extracted may simply mirror subjective biases of the raters. This seems unlikely in view of the contrasting biases of the two groups under consideration, but the most convincing evidence to this effect comes, of course, from objective correlates discovered for the two factors or diagnoses—work on sedation thresholds (Shagass and Jones, 1958), salivation rates (Strongin and Hinsie, 1939), crying (Davis *et al.*, 1969), patterns of salivary flow (Palmai *et al.*, 1967), differential response to ECT

two dimensional space generated by the two dimensions discovered by the investigators quoted, and the question arises as to the expected distribution of the cases. On a continuity (dimensional) model we would expect the distribution to be a normal bivariate distribution, or at least approximating to such a normal distribution; difficulties of measurement and problems arising with the metric to be used may distort this surface to a considerable extent, but without disguising its essential nature. On a non-continuity (categorical) model we would expect to find the majority of cases clustering around the two major axes, i.e. to be *either* endogenous or reactive; the space in the quadrants would be largely empty, and so would the axes below the origin (because there would be found the people who did not have either of these categorical disease entities). But as long as we concerned ourselves only with a clinic population we would not need to bother about non-clinic, non-depressed people; we would have one group of endogenous patients, differing from each other in severity of depression, and another group of reactive patients, differing from each other in severity of depression, with only a few patients in the intervening space (i.e. those unfortunates who were suffering from two qualitatively different depressions simultaneously — they would be very small in number because their proportion in the population would be the product of the separate frequencies, both small, with which these two disease entities occurred in the population). Curiously neither Hamilton and White nor the later writers have in fact plotted the bivariate surface, or carried out any calculations relevant to this problem; what they have done instead has been something quite different. Furthermore, they have done this without putting forward any argument to defend their rather unorthodox procedures. Fortunately there is a statistical relationship between the plots given by them of patient score distributions, and the hypothetical bivariate distributions discussed above, so that certain preliminary decisions become possible. What these various authors have done, in one form or another, is that having demonstrated that there are two independent axes, factors or causal principles involved in the distribution of the depressive symptoms they then *collapse* these continua into one and plot distributions along this one, single dimension, arguing that in some way this is relevant to the problem of unitary or binary nature of depression. It may be useful to illustrate what has been done by an example. Suppose that an investigator took a number of measurements of the lengths of different bones in a given subject population, and that he also took daily measures of their wine, spirit, and beer consumption. On intercorrelating and factor analysing these measures, he would find two independent factors, height and amount of alcohol consumed. Not content with this, he might argue that if these were really independent factors, then the distribution of cases from one to the other should be bimodal. He would then plot an axis running from the height (+)/alcohol consumption (+) end to the height (—)/alcohol consumption (—) end, and give each person a score, weighted by the loading on the two factors. Thus at the one extreme of this distribution would be tall drinkers; at the other end small abstainers. In between would be medium-sized persons drinking in moderation, tall abstainers and small drunkards; the distribution of scores, given that the bivariate distribution was normal or roughly so, would be unimodal! Lo and behold, our investigator has proved that height and drinking habits formed one continuum!

A more realistic, and less hypothetical example may be taken from the author's work on sociability. Social shyness was put forward

---

and antidepressive drugs (Kiloh, Ball and Garside, 1962), and the like. Kendell argues that 'the fact that the two forms have significantly different response rates to a therapeutic agent, or significantly different means for a physiological characteristic such as salivation rate or sedation threshold, no more proves that they are different diseases than would a demonstration that tall men weigh significantly more than short men prove the two to be different species. Only the demonstration of a bimodal distribution curve for some property of an unselected sample of the whole population is adequate to distinguish the two possibilities.' This quotation illustrates better than anything else the confusion between the two problems of unitary *v.* binary and categorical *v.* dimensional nature of depressive illness to which this note draws attention.

by Guilford as a unitary factor; correlating scores on the questionnaire so named with introversion and neuroticism produced positive and moderately high correlations in each case. The writer (Eysenck, 1956) put forward a two-dimensional hypothesis, suggesting that some behaviours characteristic of social shyness were related to *introversion* (introverts don't much care for being with other people), while other behaviours characteristic of social shyness were related to neuroticism (neurotic people are afraid of other people). When individual items from the Guilford S scale were correlated with measures of introversion and neuroticism they split neatly into some which correlated with introversion, but not with neuroticism, and others which correlated with neuroticism, but not with introversion. Thus social shyness behaves very much like depression, breaking up into two orthogonal, unrelated entities. It would now be possible to draw a continuum from introverted to neurotic social shyness (i.e. at an angle of 45° to E and N), and plot the scores of our sample on this continuum. It is unnecessary to do this, as it is known that the distribution of scores on E and N forms a close approximation to a normal bivariate surface; it follows mathematically that the distribution of scores on this new, hypothetical continuum of 'sociability' must be unimodal and normal. This simply proves that introversion and neuroticism are continuous variables, not categorical types; it does not affect our decision about the unitarian or binary nature of sociability.

The point of these examples is simply that having determined the existence of two independent factors as giving rise to the observed inter-correlations of the symptoms, investigators cannot then turn round and construct a single continuum running from one factor to the other. This is strictly meaningless, and so is the nature of the distribution of scores on this continuum, at least as far as a check on the unitarian *v.* binary nature of the surface is concerned. Having decided this question once and for all by factor analysis, it is not permissible to go back and take another bite out of the cherry by assuming a continuum (from endogenous to reactive) which the previous analysis has ruled out of court. No doubt figures can be plotted along such a continuum, but these are as meaningless as those combining height and amount of alcohol consumed; they do not throw any light on the issue in question. The only argument which is statistically admissible is as follows: If the bivariate surface shows a normal distribution, then any arbitrary axis through the origin, at whatever angle to the main factor axes it is constructed, will also permit the construction upon it of a unimodal and in fact normal distribution of scores—however meaningless these scores may be in psychological terms. But, as we have seen, a normal bivariate distribution is only to be expected if we have a dimensional model; a categorical model would not generate such a distribution. Consequently, a normal distribution, such as that found by Kendell (1968), is suggestive of a dimensional model for both endogenous and reactive depressions; this is the only issue to which his plotted distributions are relevant.* The somewhat bimodal distributions of the Newcastle workers may be interpreted as supporting, though rather weakly,

* The distributions reported by Kendell, e.g. in his Fig. 17, are normal only in a relatively non-informative sense. He himself says of the plot of 53 Maudsley patients that 'it appears to be unimodal'; no statistical test is reported. To the eye the distribution does not appear to be particularly unimodal; the cleavage is as obvious as that in certain pre-war English films censored for that reason by the American Hays Office. A statistical test comparing the observed distribution with a hypothetical normal one having the same mean and variance just fails to disprove the null hypothesis (corrected chi square = 6·45), but this is probably not the proper comparison; the real question is whether the observed distribution is significantly different from a bimodal one such as that reported by the Newcastle workers. Such a comparison is difficult to make because the variances of the two distributions differ quite markedly; this difference renders meaningless Kendell's point that 'a high proportion of the scores fall in the trough between the two peaks of Carney's distribution curve'—this does not disprove the bimodal hypothesis, unless variance differences are adjusted. When such an adjustment is carried out, it can be seen that Kendell's own distribution is closer to a bimodal one (i.e. to the Newcastle distribution) than it is to a normal one. The number of cases is too small to make detailed reporting of the calculations worth while, but Kendell's argument is certainly untenable and lacks the needed statistical support; visual inspection is not sufficient to establish a point which he considered vital for his argument, although, as we have argued, it is really irrelevant and concerned with an entirely different argument.

a categorical model, or they may be the result, as Kendell hints, of halo factors deriving from theoretical preconceptions. It does not seem that arguments from distributions are particularly strong, for a variety of reasons. Even when the dimension along which distribution of scores is plotted is properly established, and relevant, the *shape* of the distribution is subject to many influences which seriously weaken any attempt to use the observed score values as estimates of the true values (Eysenck, 1960). The metric used is arbitrary, and almost certainly not of the needed cardinal kind; unreliability introduces error variance which is normally distributed and may distort the 'true' distribution (particularly in relation to psychiatric diagnoses and judgements, as here, where according to Kendell their reliability is only about 50 per cent of the total variance, leaving 50 per cent as simple error variance); halo and other influences, including preconceived opinions, may determine the judgements which are being aggregated.

But even more detrimental to any evidential value of such distributions is the unsolved problem of selection. Kendell used successive admissions, whereas the Newcastle workers seem to have selected (in part) purposely clear cases of either endogenous or reactive depression (at least what they say leaves this possibility open; their words are not entirely clear on this point). But even successive admissions are biased by a great variety of factors which influence individual decisions by consultants involved in this procedure; admissions certainly do not represent fairly all applicants. And all applicants are not a proper sample of all persons who might be considered to be suffering from depression. Thus the final distribution of scores, along this meaningless continuum, will depend very strongly on selection procedures (the nature of which is largely unknown) used on samples themselves self-selected (or G.P. selected) on principles equally unknown, from a universe entirely unknown! It does not require much knowledge of statistics and sampling procedure to see that this method is unlikely to give us a population very representative of anything, other than itself.

An interesting example of how selection at various levels may give rise to quite unreal associations between symptoms is given by Elwood and Wood (1967), from their work on anaemia. As they point out, it is believed that low levels of haemoglobin in the circulating blood are causally related to certain symptoms, such as irritability, palpitations, dizziness, breathlessness, fatigue and headache. They argue that these symptoms are fairly widespread and commonly ascribed to a very wide variety of other conditions, such as high blood pressure and psychoneurosis; 'it could be, therefore, that these symptoms occur in a large proportion of persons in the community who consult a doctor. A proportion of these may be referred to hospital, where, among other investigations, an estimation of haemoglobin level is made. The basis on which a doctor makes a decision whether or not to refer a patient to hospital is not clearly known, but it could well be that pallor, which may be caused by anaemia, is an important determinant. In any case a relatively high proportion of those referred to hospital will show a haemoglobin level below the range generally accepted as "normal", because even in a random sample of women from the community about one-quarter to one-third are known to have such levels. In this way a group of subjects with symptoms and low haemoglobin level could be selected, without its being realized that any association between symptoms and anaemia which is detected in the group is coincidental. Furthermore, whatever the origin of the symptoms, one can expect patients to report benefit from whatever therapeutic measure is advised, whether this is blood transfusion or a simple iron tonic.' When correlations were worked out in an unselected, representative group between the six symptoms mentioned above and haemoglobin level, these were half positive and half negative, and wholly insignificant! (Correlations were also run between six symptoms and neuroticism; all were positive and all were significant.) This study shows how selection can produce quite unreal relationships, and warns that widely spread beliefs of an erroneous nature can be built upon such artificial associations. (It remains possible, of course, that *extremely* low levels of haemoglobin could give rise to symptoms of the kind

mentioned; there is no implication in these figures that the regression is a linear one.)

How does this example relate to our problem? Before taking up this point we may, with advantage, look at a possible objection to our argument about the artificial nature of the endogenous-reactive continuum suggested by Kendell. He might be willing to admit that this continuum does not throw much light on the unitarian *v.* binary argument, but would almost certainly argue that from the practical point of view scores on this continuum are useful in allocating patients to different types of therapy. As he says, 'a patient's position on that continuum . . . provides more information about symptomatology, treatment and prognosis than does assignment to a traditional diagnostic theory'. This is true, of course, but it is relevant to the categorical *v.* dimensional argument, not the unitarian *v.* binary one; indeed, this is precisely the advantage I have always claimed for dimensional systems of diagnosis as opposed to the categorical, 'medical disease entity' ones. But the alternative to Kendell's continuum is not a categorical model; the proper alternative would be a dimensional model which does not contradict the findings of all the investigators named, as does Kendell's continuum. In other words, where we have two separate and independent continua, one of endogenous and one of reactive depression, there we must characterize each patient by *two* scores, stating his position on both continua! Consider Fig. 1. The two major axes represent endogenous and reactive depression, and on both the patient's position is marked in terms of an arbitrary 10-point scale. The oblique line represents Kendell's continuum, going from E (positive scores show predominance of endogenous symptoms) to R (negative scores show predominance of reactive symptoms). In the two-dimensional representation each of the hypothetical patients (A, B, C, D) is represented by two numbers, i.e. A would have a score of 9 on endogenous and 2 on reactive depression; B would have scores of 3 and 8; C of 3 and 3; and D of 7 and 7. The scores of these patients on the Kendell continuum are indicated by the projections on to that continuum. These scores discriminate well enough between A and

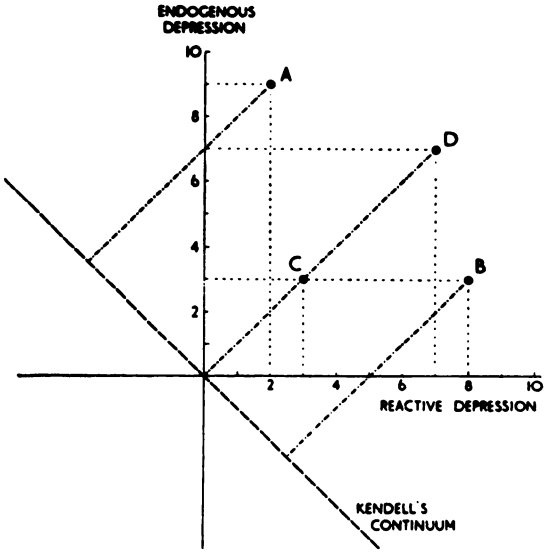

FIG. 1.—Diagram indicating the factorial resolution of symptoms associated with depressive illnesses in terms of two factors (endogenous and reactive). The diagram also shows Kendell's proposed continuum, and the position of four hypothetical patients on the two factorial continua, and on Kendell's continua.

B; their failure to represent all the important variables becomes apparent when we compare C and D. Both have the same score on the Kendell continuum, although they are widely different on both of the original continua! This simply highlights the obvious point that two dimensions cannot be collapsed into one, or projected onto one, without loss of information; such information would be contained in an axis drawn through the origin at right angles to the Kendell continuum, and would represent severity of illness; it would lie along the line connecting D, C, and the origin. We must conclude that, while for some very limited practical purposes Kendell's continuum might be sufficient, it has little theoretical value and even from the applied point of view is severely limited and inferior to a proper two-dimensional representation of the patient's position in the two-dimensional surface generated by the main factors of endogenous and reactive depression.

We can now return to our anaemia example of the power of selection to distort distributions and even to generate completely artificial 'illnesses'. It seems likely that selection for

hospitalization is more in terms of severity (i.e. the axis in Fig. 1. which would be orthogonal to the Kendell continuum) than in terms of endogenous or reactive; thus it would matter whether a patient was in D's position or C's, but not so much whether he was in A's or B's. This might easily lead to a predominance of patients *combining* symptoms of endogenous and reactive depression in hospital, thus distorting any kind of distribution of scores based on Kendell's continuum and disguising any bi-modality that might actually exist. Conversely, it is possible that the Newcastle workers may have selected patients for inclusion on the basis of extreme scores on Kendell's continuum, thus affecting the shape of the distribution in the opposite direction. Only a sample drawn at random from a representative population, like that used by Elwood and Wood, can give us information on the distribution of scores—although even information gained from such a sample is still subject to several of the criticisms mentioned earlier. Fortunately correlations (and consequently factors) are much less affected by even quite serious deficiencies in sampling, and it is significant that although the various authors cited previously obtained different distributions of scores, they all obtained very similar patterns of inter-correlations.

Kendell claims in favour of his 'single continuum' motion that 'it preserves the traditional stereotypes as the two poles of the continuum and acknowledges that the differences between them are genuine and not simple questions of severity and chronicity. On the other hand it recognizes the impracticability of drawing any clear boundary between them'. These two sentences put very clearly the fundamental fallacy in Kendell's thinking. He writes as if there were only one two-dimensional surface, containing the two 'traditional stereotypes', and leading to difficulties in 'drawing any clear boundary between them'. But there are two such surfaces. The first is given in his Fig. 11 (p. 44); it contains the two factor plot of item-loadings on endogenous and reactive depression. There is a clear cut boundary between them, and this could be made even clearer by dropping items from each factor having sizeable loadings on the other. The second surface is that given in our Fig. 1; it contains the factor scores of *persons* and here indeed there are no clear-cut boundaries between these patients. But this second surface, and the difficulties it presents to psychiatric diagnosis, relate to the categorical *v.* dimensional argument; all this is irrelevant to the unitarian *v.* binary argument, which finds an answer in relation to the first (factor loadings) surface. Factor-space and person-space are two different conceptions, and should not be used interchangeably. Extraversion is entirely different from neuroticism; this does not mean that persons do not exist who are both extraverted and neurotic. Similarly, the existence of two independent and separate factors of depression does not preclude the existence of patients suffering both from endogenous and reactive depression, and showing symptoms of both. Kendell's solution was along the right lines, but it did not go far enough. If the binary view is right, and if the dimensional view is right, then the only proper solution to the diagnostic problem is to give each patient *two* scores, representing his or her position on these two dimensions.* Nothing less will do, either in doing justice to the theoretical model, or in affording the optimum guidance to treatment and prediction of outcome. Using the Kendell or the Newcastle continuum has been shown to be clearly superior to using categorical diagnosis, but predictions and success of treatment were still clearly sub-optimal; it is suggested that by working with two continua the usefulness of diagnosis will be increased even more.

*From the point of view of information, it could be argued that two scores based on Kendell's continuum and on the 'severity of illness' continuum orthogonal to it would transmit as much information as do the two scores based on the reactive and endogenous continua. Statistically this is true; the main point of this note is that two continua, not one, are needed to incorporate all the available information. Any arbitrary rotation of the two axes shown in Fig. 1 would give continual scores on which would transmit identical information—as would a set of polar co-ordinates. From the point of view of psychological meaningfulness it is suggested that some positions of these axes are clearly superior to others, and in particular that that shown in Fig. 1 results in scores having maximum meaning. This point, however, unlike the question of the number of dimensions needed to describe the symptoms rated, is subject to argument and does not admit of an objective answer.

Summary

It is argued that the dispute between the Newcastle group of workers and the London school about the classification of depressive illnesses in terms of one or two dimensions is based on a misunderstanding of the issues involved, and the logic of the statistical methods used by them. There are two, not one, problems involved, relating (a) to the unitary or binary nature of depression, and (b) to the categorical or dimensional nature of these illnesses. Factor analysis is relevant to (a), and conclusively favours the binary view; distribution of scores is relevant to (b), and cannot throw any light on the binary-unitary problem. Kendell's compromise solution of a single continuum running from reactive to endogenous depression is shown to be inadequate statistically, and irrelevant psychologically. A general solution to the diagnostic problem of depressive illnesses is presented.

References

Carney, M. W. P., Roth, M., and Garside, R. F. (1965). 'The diagnosis of depressive syndromes and the prediction of ECT response.' *Brit. J. Psychiat.*, **111**, 659–74.

Cattell, R. B., and Scheier, I. N. (1961). *The Meaning and Measurement of Neuroticism and Anxiety*. New York: Ronald Press.

Curran, D. (1937). 'The differentiation of neuroses and manic-depressive psychosis.' *J. ment. Sci.*, **83**, 156–74.

Davis, D., Lamberti, J., and Ajans, Z. A. (1969). 'Crying and depression.' *Brit. J. Psychiat.*, **115**, 597–8.

Devadasan, K. (1964). 'Cross-cultural validity of twelve clinical diagnostic tests.' *J. Indian Acad. appl. Psychol.*, **1**, 55–7.

Elwood, P. C., and Wood, M. M. (1967). 'A bloody shibboleth.' *New Scientist*, 28 December, 756–8.

Eysenck, H. J. (1950). 'Criterion analysis: an application of the hypothetico-deductive method of factor analysis.' *Psychol. Rev.*, **57**, 38–53.

——(1952a). 'Schizothymia-cyclothymia as a dimension of personality.' *J. Personal.*, **20**, 345–84.

——(1952b). *The Scientific Study of Personality*. London: Routledge & Kegan Paul.

—— (1953). 'The biological basis of factor analysis.' *American Psychologist*, **8**, 105–14.

—— (1955). 'Psychiatric diagnosis as a psychological and statistical problem.' *Psychol. Rep.*, **1**, 3–17.

—— (1956). 'The questionnaire measurement of neuroticism and extraversion.' *Rev. di. Psicol.*, **50**, 328–33.

—— (1960). 'Classification and the problem of diagnosis', in *Handbook of Abnormal Psychology* (ed. H. J. Eysenck). London: Pitman.

—— (1960). *The Structure of Human Personality*. London: Methuen.

—— (1964). 'Principles and methods of personality description, classification and diagnosis.' *Brit. J. Psychol.*, **55**, 284–94.

——(1969). 'A dimensional system of psychodiagnostics', in *New Approaches to Personality Classification and Psychodiagnosis* (ed. A. R. Mahrer). New York: Columbia University Press.

—— and Claridge, G. (1962). 'The position of hysterics and dysthymics in a two-dimensional framework of personality description.' *J. abnorm. soc. Psychol.*, **64**, 46–55.

—— and Eysenck, S. B. G. (1968a). 'A factorial study of psychoticism as a dimension of personality.' *Multivar. Beh. Res.*, Special Issue, 15–31.

—— Granger, G., and Brengelmann, H. (1957). *Perceptual Processes and Mental Illness*. Oxford University Press.

Eysenck, S. B. G. (1956). 'Neurosis and psychosis: an experimental analysis.' *J. ment. Sci.*, **102**, 517–29.

—— and Eysenck, H. J. (1968b). 'The measurement of psychoticism: a study of factor stability and reliability.' *Brit. J. soc. clin. Psychol.*, **7**, 286–94.

—— —— (1969). 'Scores on three personality variables as a function of age, sex and social class.' *Brit. J. soc. clin. Psychol.*, **8**, 69–76.

—— —— (1969). 'Psychotism' in children: a new personality variable.' *Res. in Educ.*, **1**, 21–37.

Hamilton, M., and White, J. M. (1959). 'Clinical syndromes in depressive states.' *J. ment. Sci.*, **105**, 985–998.

Kendell, R. E. (1968). *The Classification of Depressive Illness*. Oxford: University Press.

Kiloh, L. G., Ball, J. R. B., and Garside, R. F. (1962). 'Prognostic factors in treatment of depressive states with imipramine.' *Brit. med. J.*, *i*, 1225–7.

—— and Garside, R. F. (1963). 'The independence of neurotic depression and endogenous depression.' *Brit. J. Psychiat.*, **109**, 451–63.

Lewis, A. J. (1934). 'Melancholia: a clinical survey of depressive states.' *J. ment. Sci.*, **80**, 277–378.

—— (1938). 'States of depression: their clinical and aetiological differentiation.' *Brit. med. J.*, *ii*, 875–8.

McConaghy, N., Joffe, J. D., and Murphy, B. (1967). 'The independence of neurotic and endogenous depression.' *Brit. J. Psychiat.*, **113**, 479–84.

Mapother, E. (1926). Opening paper of discussion on manic-depressive psychosis. *Brit. med. J.*, *ii*, 872–6.

Palmai, G., Blackwell, B., Maxwell, A. E., and Morgenstern, F. (1967). 'Patterns of salivary flow in depressive illnesses during treatment.' *Brit. J. Psychiat.*, **113**, 1297–308.

SHAGASS, C., and JONES, A. L. (1958). 'A neurophysiological test for psychiatric diagnosis: results in 750 patients.' *Amer. J. Psychiat.*, 114, 1002–9.

STRONGIN, E. I., and HINSIE, L. E. (1939). 'A method for differentiating manic-depressives from other depressives by means of parotid secretions.' *Psychiat. Quart.*, 13, 697–704.

H. J. Eysenck, *Professor of Psychology, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London S.E.5*

(*Received 15 July 1969*)