Correspondence

DEPRESSIVE ILLNESSES IN LATE LIFE

DEAR SIR,

In their letter (*Journal*, May 1973) Drs. Kendell and Post report that they have obtained a trimodal combined distribution of the scores of 271 patients on the Newcastle diagnostic index (Carney *et al.*, 1965). They claim, however, that their trimodal distribution is devoid of meaning and a statistical artefact.

To take first the meaning of the observations: when Kendell and Post's combined distribution, which is reproduced below, is examined, it is apparent that there is a dip at score 5 (the frequency at this score is only 22, whereas both adjoining frequencies are 31).

Scores of Maudsley cases on the Newcastle Scale

Score on Newcastle Scale	Kendell	Hemsi and McClure	Post	Post, Cawley and White-head	Total
+13	0	0	1	0	1
+12	0	1	0	0	1
+11	1	0	0	0	1
+10	2	1	3	1	7
+ 9	0	1	2	3	6
+ 8	1	9	6	4	20
+ 7	7	7	9	3	26
+ 6	7	7	13	4	31
+ 5	2	8	6	6	22
+ 4	6	12	8	5	31
+ 3	5	9	10	7	31
+ 2	6	2	9	5	22
+ 1	7	4	6	3	20
0	4	6	11	7	28
− 1	4	5	6	1	16
− 2	1	4	1	0	6
− 3	0	1	1	0	2
Total	53	77	92	49	271

In the distribution of scores published by Carney *et al.* (1965) the dip occurs at a closely similar point including the score of 5. Post's (1972) individual data also show an inflection at 5, and according to his Fig. 2 this can be regarded as the line of separation between his psychotic and neurotic patients. The dip in Kendell's (1968) individual distribution occurs at precisely the same score.

Drs. Kendell and Post can hardly be regarded as having been biased in favour of bimodality. Yet the distributions they have both, independently, found conform, in relation to the dip at score 5, to that reported by Carney *et al.* (1965), which led them (Carney *et al.*) to conclude that there were at least two distinct groups of depressed patients.

The meaning of the second dip in Kendell and Post's combined distribution becomes equally plain when the findings are related to other observations. It occurs at scores 2 and 1 with the third highest mode at score 0. The probable significance of this third mode is made clear by Post's own findings. The distribution (Post, 1972, Fig. 2, group EF) of his 36 'neurotic depressive' subjects (1, 1, 5, 10, 3, 5, 6, 2, 2, 1) is bimodal, but does not quite reach significance (p = ·076). The frequency of 3 corresponds to a score of 1. That part of the distribution to the left of score 2 (20 patients) must consist, in the main, of the 20 patients in Post's group F, because this group has a mean score of 0·65, whereas the 16 patients in Post's group E have a mean of 2.00.

The clinical features of group F are described by Post (1972, p. 397) as follows: 'These patients came to require in-patient care for states characterized by poor sleep and appetite, loss of interest in outside matters, and also by non-delusional preoccupations with unpleasant internal physical sensations, tensions and anxiety'. They had 'neither overt depression nor depressive thought content'. In other words, they were anxious rather than depressed.

Several factor-analytic and related studies have isolated an 'anxiety' factor relatively independent of depressive features. In particular, in the Newcastle diagnostic scale employed by Kendell and Post, anxiety has the largest negative weight. As their findings conform to a pattern discernible in a number of other enquiries they can hardly be regarded as clinically meaningless.

The three modes of the significantly non-normal distribution that emerge from their combined data not only provide some independent corroboration for their view that there are at least two relatively distinct forms of depressive illness; the data are also

in accord with the evidence that favours the classification of affective disorders into two broad groups—namely, anxiety states and depressive disorders (Roth *et al.*, 1972; Gurney *et al.*, 1972). Far from signifying nothing, as Drs. Kendell and Post claim, their trimodal distribution has considerable clinical, practical and scientific interest.

We shall now consider the arguments advanced by Kendell and Post to suggest that their distribution is a statistical artefact which has arisen 'either because the data from which the weightings of the scale were originally derived did not satisfy the requirements, basic to discriminant function analysis, that the two criterion populations should have multivariate normal distributions and equal variance/co-variance matrices; or because the item distributions in our data were markedly skewed'.

To consider their last point first: skewness in the original items does not produce a multi-modal distribution of their sum. On the contrary, it follows from the central limit theorem (Cramer, 1955, pp. 114–6) that when independent or only slightly correlated items such as those of the Newcastle index (whether skewed or not) are added, the distribution of their sum is unimodal (Guilford, 1956, p. 452).

So far as normality and equal variance/co-variance matrices are concerned, it is true that these conditions must be met if *optimum* differentiating weights are to be found, as Winkel and Juhl (1971) have pointed out. But if these conditions are not met, it only means that the weights will not be optimally differentiating; it does not mean that the weighted sum will be multimodal.

The view that the findings are not susceptible of interpretation because the weights in question were derived so as to distinguish between two groups only, is equally devoid of substance. As Moran (1966, p. 1168) has pointed out—'It is worth emphasizing that the evidence for such heterogeneity rests on the bimodality of the frequency distribution of the scores and *not on the manner in which the latter has been found*' (our italics). Moran's conclusion obviously applies equally to trimodality. Whatever the derivation, the distribution of the sum of 18 weighted clinical features cannot be dismissed as meaningless.

It is puzzling that the authors see no contradiction between their assumed failure to substantiate the existence of two or more distinct 'forms' of depressive illness and Kendell's recent reaffirmation (Kendell, 1972), of his view that there is not one kind of depressive illness only but rather that 'there are important and fundamental differences between different depressions that are not simply differences in severity or chronicity' (Kendell, 1968).

To assert that there is not one *kind* of depression only amounts to saying that there are two or more. Such distinct forms may not be mutually exclusive—indeed they are clearly not so in this case. But that they are independent, or negatively correlated, syndromes has been repeatedly demonstrated in studies of depressive illness. If Drs. Kendell and Post mean that they have found no evidence for two distinct *groups* of depressed patients, this conclusion is warranted only if, as they claim, their trimodal distribution is a meaningless artefact. As we have indicated, this conclusion does not seem justified. Drs. Kendell and Post should not dispute the validity of their own results merely because their distribution of scores contains a serendipitous third mode.

This controversy has some interesting implications for the objectivity and reliability of clinical observations in psychiatry. In Kendell's monograph (1968) evidence was adduced to suggest that previous findings may have been markedly influenced by the observers' preconceptions; the sharp distinction between endogenous and neurotic depression found by some workers could have largely stemmed from bias and observer error. But the data published by Kendell and Post (1973) tend to controvert this view. They suggest, on the contrary, that systematic clinical observation by experienced psychiatrists may help to establish the facts of the case, despite halo effects from preconceptions and former enquiries. The recent findings of these authors serve to substantiate views about the classification of depressive disorders which they have consistently called in question.

MARTIN ROTH.
R. F. GARSIDE.

Department of Psychological Medicine,
University of Newcastle upon Tyne,
Royal Victoria Infirmary,
Queen Victoria Road,
Newcastle upon Tyne NE1 4LP.

REFERENCES

CARNEY, M. W. P., ROTH, M., and GARSIDE, R. F. (1965). 'The diagnosis of depressive syndromes and the prediction of ECT response.' *British Journal of Psychiatry*, 111, 659–74.

CRAMER, H. (1955). *The Elements of Probability Theory.* New York: Wiley.

GUILFORD, J. P. (1965). *Fundamental Statistics in Psychology and Education.* New York: McGraw-Hill.

GURNEY, CLAIR, ROTH, M., GARSIDE, R. F., KERR, T. A., and SCHAPIRA, K. (1972). 'The relationship between anxiety states and depressive illnesses.' *British Journal of Psychiatry*, 121, 162–6.

KENDELL, R. E. (1968). *The Classification of Depressive Illnesses.* London: Oxford University Press.

—— (1972). *British Journal of Psychiatry*, 121, 575.

KENDELL, R. E. and POST, F. (1973). *British Journal of Psychiatry*, **122**, 615–7.

MORAN, P. A. P. (1966). 'The establishment of a psychotic syndrome.' *British Journal of Psychiatry*, **112**, 1165–71.

POST, F. (1972). 'The management and nature of depressive illnesses in late life: a follow-through study.' *British Journal of Psychiatry*, **121**, 393–404.

ROTH, M., GURNEY, CLAIR, GARSIDE, R. F., and KERR, T. A. (1972). 'The relationship between anxiety states and depressive illnesses.' *British Journal of Psychiatry*, **121**, 145–61.

WINKEL, P., and JUHL, E. (1971). 'Assumption in linear discriminant analysis.' *Lancet*, **ii**, 435–6.

REPLICATION STUDIES OF ANXIETY

DEAR SIR,

There are many who feel that the paucity of replication studies is a consistent failing in current psychiatric work. It was therefore with considerable pleasure that we turned to the paper on 'Neurotic and Thyrotoxic Anxiety: Clinical, Psychological and Physiological Measurements by Greer *et al.* (1973). The pleasurable anticipation was, however, speedily transformed to dismay on realizing that while departing in the crucial technical details from the Lader and Wing study (1966) they nevertheless drew conclusions as if their own experiments were a faithful replication of that study.

To particularize, spontaneous fluctuations were measured in resistance units (Sternbach, 1960), wthout an explanation being proffered. Lader and Wing are explicit in insisting on the use of log conductance units as being more appropriate; and, depending as they do upon basal levels on which the changes of spontaneous fluctuations are superimposed, these would result in quite different figures.

The most important modifications, which Greer *et al.*, appear to dismiss as being of small relevance, relate to the auditory stimuli: these numbered 15 instead of 20, were given over a period of 15 minutes instead of 20 minutes, varied in duration from 5–15 seconds instead of all being of 1 second duration, and were randomized at intervals from 20–120 seconds instead of 45–80 seconds. These are major modifications, as numerous reports in the literature associating habituation rate with number, frequency and duration of application of stimuli will testify (Malmo *et al.*, 1948).

At another level of criticism one might argue: that a 20-second interval between stimuli is insufficient to allow for spontaneous fluctuations to emerge, especially in individuals in whom the recovery arm of the response curve is unduly prolonged; that no information is given whether the women were in the follicular or luteal phase of the menstrual cycle—yet this is known to be important

(McKinnon, 1954; McKinnon and Harrison, 1961); that there was no mention of standardization of the time of day at which recordings were made, which might be thought important because skin conductance is subject to circadian rythmn, being greatest at 2.00 p.m. (Christie and Venables, 1972; Venables and Martin, 1967). The implication by the authors that they have avoided 'methodological defects' is open to question.

At St. Bartholomew's, a faithful replication of the Lader and Wing experiment revealed a high correlation between a psychological rating, the Morbid Anxiety Inventory (Salkind, 1972), and the 'D' score (composite physiological activity); the recording environment was carefully controlled, all recordings were made under exactly the same conditions at the same time of day, and the female subjects were only recorded during the follicular phase of the menstrual cycle. It is hoped to report this study in full but we were able to confirm other substantial correlations between skin conductance variables and the psychological rating.

It appears that in the absence of carefully controlled conditions all that emerged from the Greer *et al.* study was yet another demonstration of the variability of skin conductance measurement.

M. R. SALKIND.
J. HARRISON.

Department of Psychological Medicine,
The Royal Hospital of Saint Bartholomew,
London EC1A 7BE.

REFERENCES

CHRISTIE, M. J., and VENABLES, P. H. (1972). 'Diurnal variation in "basal skin potential level".' (In preparation.)

GREER, S., RAMSEY, I., and BAGLEY, C. (1973). 'Neurotic and thyrotoxic anxiety: clinical, psychological and physiological measurements.' *Brit. J. Psychiat.*, **122**, 549–54.

LADER, M. H., and WING, L. (1966). *Physiological Measures, Sedative Drugs and Morbid Anxiety.* London: Oxford University Press.

MALMO, R. B., SHAGASS, C., DAVIS, J. F., CLEGHORN, R. A., GRAHEM, B. F., and GOODMAN, A. J. (1948). 'Standardised pain stimulation as controlled stress in physiological studies of psychoneurosis.' *Science*, **108**, 509–11.

McKINNON, P. C. B. (1954). 'Variations in the number of active digital sweat glands during the human menstrual cycle.' *Journ. Obstet. Gynaecol.*, **61**, 390–93.

—— and HARRISON, J. (1961). 'The influence of hormones associated with the pituitary, adrenal and sexual cycle activity and palmar sweating.' *Journ. Endocrinol.*, **23**, 217–25.

SALKIND, M. R. (1972). 'Anxiety neurosis in general practice.' *Postgrad. med. Journ.*, Sept., Supplement, 34–41.