0022-3018/79/1672-0105$02.00/0
THE JOURNAL OF NERVOUS AND MENTAL DISEASE

Vol. 167, No. 2
Printed in U.S.A.

Do Psychiatric Patients Fit Their Diagnoses?

Patterns of Symptomatology as Described with the Biplot

JOHN S. STRAUSS, M.D.,[1] K. RUBEN GABRIEL, PHD.,[2] RONALD F. KOKES, PH.D.,[3] BARRY A. RITZLER, PH.D.,[4] AUTUMN VANORD, M.A.,[1] AND ELAINE TARANA[1]

Systems of psychiatric diagnosis have been regularly criticized for their low reliability and their inability to fit accurately the kinds of patients coming for treatment. To explore the reasons for these problems, this study utilizes a new method, the biplot, for defining groups of similar patients and the relationships of these groups to key symptom clusters. Using this technique to analyze data from a representative sample of first admissions for psychiatric disorder, results showed: a) symptom clusters representing the classical diagnostic categories, mania, schizophrenia, neurotic depression, and psychotic depression, were readily identified; b) however, only a few patients were clustered near these traditional syndromes.

These findings suggest that although syndromes do exist that fit traditional diagnostic categories, the vast majority of patients fall between these syndromes, having characteristics from several of them. For most patients, forcing the diagnostician to choose among the categories requires an arbitrary decision that may contribute to dissatisfaction in the diagnostician who recognizes how misleading the diagnosis can be.

Diagnostic systems in psychiatry have been criticized for their limitations in reliability, validity, and applicability. These systems have generally been developed from accumulated clinical experience, but without the aid of carefully standardized patient evaluation procedures, optimal sampling techniques, statistical analyses of patient characteristics, or evaluation of the reliability and applicability of diagnostic categories across psychiatric centers. The foundations upon which our diagnostic systems rest, therefore, may have incorporated considerable distortion and vagueness because they were derived without the aid of improved research strategies. As the difficult task of improving diagnostic systems has progressed, the increasing clarity and reliability of diagnostic categories (2, 4, 21) have caused another important problem to emerge. Large numbers of patients coming for treatment cannot be fitted neatly into the improved diagnostic categories (1).

One possible explanation for this inadequate coverage of patients by diagnostic categories is that all diagnostic systems are inherently inadequate, since there might always be large numbers of borderline or unusual cases which cannot be categorically labeled. However, a review of the history of diagnostic systems (14) suggests an alternative possibility. There may be an unnecessarily large incongruity between the present classification system used in psychiatry and patients as they are. Current systems of psychiatric diagnosis actually may not be appropriately designed to accommodate the descriptive data observed by the clinician. Recently, attempts to understand this problem have utilized multivariate statistical techniques to determine whether patients and diagnostically important symptoms group themselves into the standard diagnostic categories. However, these attempts have encountered several difficulties. Occasionally, reliable methods for collecting patient data have not been used. More frequently, it has not been possible to obtain data from a representative sample of patients coming for treatment. However, most crucial, limitations in the available multivariate statistical techniques have not permitted the necessary simultaneous evaluation of symptom groups, patient clusters, and the relationships among these groups and clusters. For example, factor analysis (3, 12, 15, 26) has been particularly useful in providing information on how patient characteristics may be grouped into closely correlated sets, but does not by itself provide tools for simultaneously classifying patients. Another valuable technique for grouping data, cluster analysis, has produced some interesting definitions of patient groups (9, 18, 23). However, the validity of these groups is problem-

[1] Yale Psychiatric Institute, Box 12A, Yale Station, New Haven, Connecticut 06520. Send reprint requests to Dr. Strauss.
This research was supported in part by National Institute of Mental Health NIMH Grants MH25466-02 and MH00006-02.
[2] Department of Biostatistics at the University of Rochester, Rochester, New York.
[3] Department of Psychology at the University of Rochester, Rochester, New York.
[4] Department of Psychology, University of Southern Mississippi, Hattiesburg.

atic since various clustering techniques often provide different groupings with the same data. Indeed, clustering algorithms always produce clusters, whether the patient population actually falls into distinct groups or forms a continuum on the characteristics used to define its members. Finally, cluster analyses do not generally describe the characteristics typical of each patient cluster nor do they determine whether some patients who are classified into different clusters are actually more similar to each other than they are to most other members of their clusters.

Another important problem with these multivariate techniques is that they attempt to operate "objectively" on the data and provide an exact mathematical solution to the problem of grouping patient characteristics and clustering patients. A more promising approach would be to use statistical analysis interacting flexibly with psychiatric insight. For this purpose, descriptive statistical techniques can be used to bring out patterns in the data and to allow the investigator to formulate syndromes and clusters which accord *both* with the data *and* with his psychiatric understanding. Such techniques are "objective" in the way they display the data, but they also allow the investigator to employ his clinical and research experience for picking out those aspects of the statistical analysis which he believes to be important.

A multivariate descriptive technique that provides these advantages for evaluating the groups formed by symptom and patient data has recently been described (5). This technique, known as the biplot (see Appendix), is based on principal components analysis. Analyses generated from the biplot can be graphed to display patient characteristics, such as symptoms, for example, as arrows from a common origin and the patients as points on the same plane with the arrows (see Figure 1). The lengths of the arrows represent the standard deviations of the symptom measurements. The angles between arrows represent the correlations between the symptoms. The distances between points represent dissimilarities of patients on their symptom measures, *i.e.*, close points represent patients with similar symptoms; distant points, patients who differ on at least one symptom. Sheaves of arrows which go in much the same direction would represent sets of highly correlated symptom measures that could qualify as syndromes. Clusters of points would represent groups of patients with similar symptoms. The syndromes typical of each group of patients are represented by the arrows going through the particular group of points. The mathematical explanation of this technique has been described in detail elsewhere (5–7).

To interpret the biplot, it is important to be guided both by the biplot configurations and by total knowledge about the patients. For a valid use of statistics, selection of sheaves and clusters must not violate the data as displayed. For a psychiatrically meaningful analysis, the symptom sets and patient groups selected must also reflect psychiatric judgment.

Procedure

Clinically, symptoms account for the great majority of variance in arriving at psychiatric diagnoses (2). In this study, therefore, symptom data were used to test the structure of psychiatric classification. To evaluate the correspondence between empirically derived symptom groups, clusters of psychiatric patients as defined by their symptoms, and diagnostic categories, symptom ratings from two different patient groups were analyzed with the biplot technique. For these analyses, symptoms were defined in terms of dimensions (23) derived by combining rating data from items on the structured interview schedule described below. Dimensions were scored: 0 = symptom absent; .25 = symptom possibly present; .50 = symptom definitely present but not continuous or severe; .75 = symptom definitely present and moderately continuous and/or severe; 1.0 = symptom present and continuous and/or severe.

Archetypal patients. The first data studied consisted of symptom ratings on 100 "archetypal patients" (23). These patients were actually fictitious subjects on whom ratings had been made on 25 dimensions of symptomatology (Table 1) to fit stereotyped concepts of five psychiatric diagnostic categories. These diagnostic categories were: neurotic depression, psychotic

TABLE 1
Symptom Dimensions

Disorientation (Disorient.)	Suspicious (Suspic.)
Somatic Complaints (Somat.)	Delusions of Reference (Del. Ref.)
Withdrawal (Withdra.)	Delusions of Grandeur (Del. Grand.)
Retarded Actions (Ret. Mo.)	Delusions, Religious and Sexual (Oth. Del.)
Depression (Depress.)	Insight, Lack of (Low Insite)
Elation (Elate)	Flatness of Affect (Flat)
Anxiety (Anx.)	Incongruous Affect (Incong.)
Obsessions (Obsess.)	Labile Affect (Labile)
Auditory Hallucinations (Aud. Hall.)	Retarded Speech (Ret. Sp.)
Visual Hallucinations (Vis. Hall.)	Incomprehensibility (Incomp.)
Other Hallucinations—tactile olfactory, gustatory (Oth. Hall.)	Bizarre Behavior (Bizar.)
Depressive Delusions (Dep. Del.)	Depersonalization (Depers.)
Delusions of Passivity (Del. Pass.)	

depression, paranoid schizophrenia, simple schizophrenia, and mania. This group of patients (rated on a larger number of symptom dimensions) had been useful previously to evaluate cluster analysis techniques for patient classification (23). Use of such a population made it possible to determine how the various techniques would cluster a population of known group structure.

Real patients. The second set of data used for this investigation were symptom ratings from a representative sample of all patients who were first admissions for functional psychiatric disorder from two catchment areas (11, 25). Data on these patients had been collected in the course of the First Admission Study (FAS), an investigation of diagnostic and prognostic characteristics in first psychiatric admissions. In that study, a representative sample of all of such patients between the ages of 15 and 55 who came from two demographically heterogeneous catchment areas and were admitted for their first psychiatric hospitalization during a 1-year period had been evaluated with standardized interview procedures. Among these procedures was an interview to describe symptoms and signs (the PAI) derived from the Present State Examination (27). The reliability of the symptom and sign ratings on these interviews has been demonstrated. Complete information for the FAS was collected on a total of 217 patients from the two catchment areas, which was a representative sample of the total population of first admissions from those areas.

For this report, the same 25 symptom dimensions used with the archetypal patients were calculated from the data of the 217 real patients. Separate analyses of data from each of the two catchment areas were carried out to provide cross-validation by checking the results from one catchment area against the other. This was thought to be important in view of the frequent failures in replicating results from multivariate statistical analyses (13, 23).

Results

To evaluate whether the biplot would reproduce the groups of patients and symptoms built into the archetypal patients, the biplot computer program was used to analyze the symptom data from these patients first. Results are shown in Figure 1. The first two principal components, those used for the biplot, accounted for 45 per cent of the symptom rating variance. A third principal component, for purposes of simplicity not included in the following description, accounted for an additional 13 per cent of the variance. Further principal components were far less significant and none of the components beyond the first two appreciably affected the findings to be described.

Symptoms formed general groups on the plot representing archetypal syndromes. Symptom groups representing manics and paranoid schizophrenics were especially distinctive. Some syndromes fell between two groups of archetypal patients, apparently having been conceived as common to both. The symptoms defining psychotic depressives and neurotic depressives, while still identifiable, tended to coalesce somewhat. The tendency of these two depressive syndromes to merge has been described by other investigations (12, 23). The patients, a sample of whom are represented by dots, clustered neatly, with 20 patients in each of five locations. Interestingly, the simple schizophrenics are found near the center of the plot, suggesting the less distinctive nature of their symptomatology as we had apparently conceived of it.

After these encouraging results, the next step was to analyze the data from the real patients using the biplot. The results of this analysis with patients from one of the two catchment areas are shown in Figure 2. The first two principal components accounted for 35 per cent of the symptom rating variance. The symptoms arranged themselves for the real patients generally as they had with the archetypal patients. Those defining mania, schizophrenia, psychotic depression, and neurotic depression each generally grouped together on the biplot.

There were notable differences, however, between the distribution patterns of real patients and those of the archetypal patients. Although a few real patients were located on the periphery of the biplot in positions similar to those of the archetypal patients, most real patients fell in the left-hand portion of the plot, a portion representing diffuse and relatively low levels of symptomatology. Thus, the great majority of real patients did not fit neatly into archetypal categories, but clustered amorphously together in an area representing low symptom scores.

Those relatively few patients who were located around the periphery of the plot and fit the archetypal syndromes clinically also met the symptom characteristics of the diagnostic labels given to those syndromes. For example, patient AK was a man in his 40s whose major symptoms and signs were a frenetic hyperactivity, and the delusions that he had a special relationship to God and received much attention from important people. Patient CE had auditory and visual hallucinations and delusions of being controlled by an outside force. Patient XX was diagnosed as borderline by the research psychiatrist who interviewed her and on the biplot is actually on the border between the sheaf of neurotic depressive symptoms and the sheaf of schizophrenic symptoms. Patients AV and DF are two women with severe neurotic depressive symptoms.

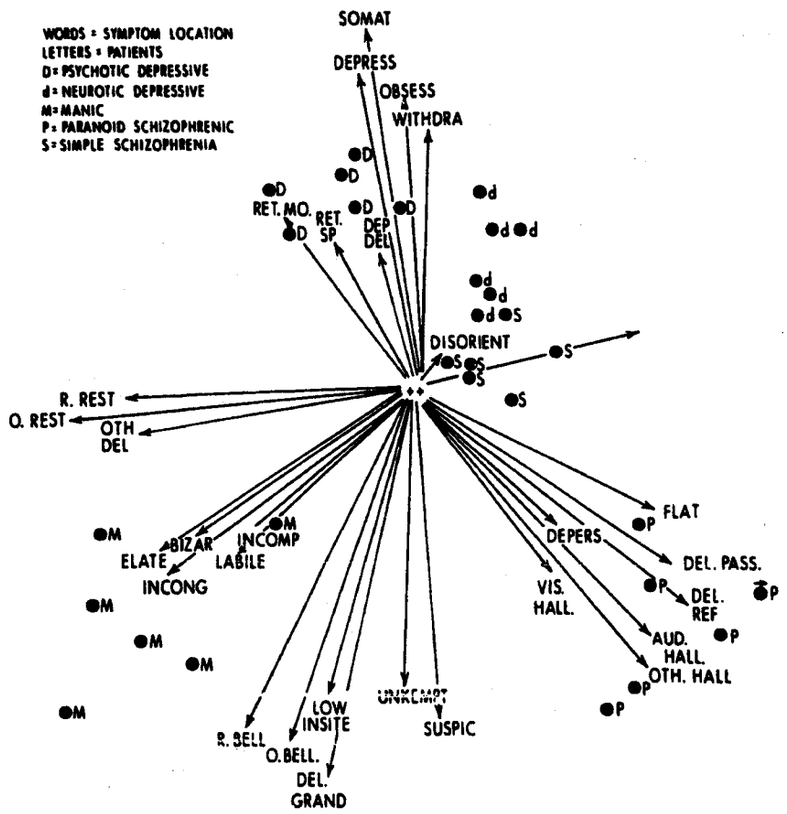

FIG. 1. Biplot of archetypal patients.

Interestingly, many patients around the biplot mean (the double + sign), *e.g.*, patient AF, were diagnosed as personality disorders. This area in the biplot reflects less severe and distinctive symptomatology, suggesting the tendency to apply the diagnosis of personality disorder to such patients. To identify where "patients" with no symptoms would be plotted, the biplot with the original patient group was run again, adding four fictitious patients with ratings of 0 on all symptom dimensions. Those patients (NN) clustered together on the left side of the biplot.

A second difference between the biplots of the real patients and the archetypal patients was that the symptoms of real patients distributed themselves in the pattern of a circle flattened on one side, but the symptoms of archetypal patients formed a full circle around the center mean point. This indicates that for the archetypal patients, each of whom was rated positively on a few symptoms and as having no symptoms in many other areas, many symptoms correlated negatively with other symptoms. With real patients, however, except for the manic and depressive polarity, most symptoms were positively correlated with each other, suggesting that each real patient had many more symptoms than we had conceptualized in the archetypal patients and that the presence of some symptoms suggested that the patient was likely to have many others as well. Few symptoms were mutually exclusive.

To evaluate whether these findings could be replicated, the biplot was run on a random sample of patients from the second catchment area. The same patterns of symptoms, failure of patients to form discrete groups and diffuse clumping of patients toward the less severe symptom part of the plot, were found.

To determine whether a grouping of patients who were readmissions rather than first admissions might better fit the archetypal pattern, the biplot was run on a set of symptom data from 88 readmitted patients from the Washington Center of the International Pilot Study of Schizophrenia—a sample described more completely elsewhere (24, 27). The biplot groupings formed by these patients were similar in some respects to the biplot patterns of the first admissions and were not more like the archetypal patient patterns.

In all three samples of real patients, the symptom

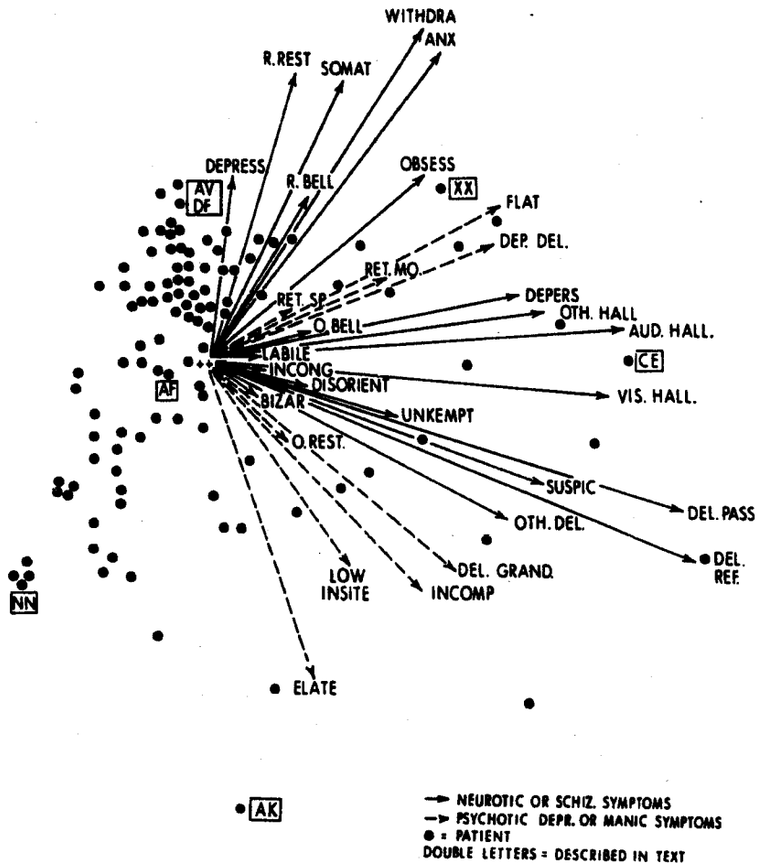

FIG. 2. Biplot of real patients.

sheaf reflecting manic disorders (low insight, elated mood) and the sheaf of neurotic depression symptoms (depression, anxiety, somatic complaints, and obsessions) were particularly distinct and at opposite poles from each other. Between these two extremes, schizophrenic symptoms were grouped together in all three samples, schizophrenic symptoms overlapping somewhat with the manic and the psychotic depressive symptoms. Psychotic depressive syndromes were also found grouped in each of the three patient samples and overlapped somewhat with the schizophrenic and the neurotic depressive symptom sheaves. Grandiose delusions, bizarre behavior, incomprehensibility, and religious and sexual delusions overlapped with the schizophrenic and manic symptom sheaves. Depersonalization—an important symptom in borderline patients (10)—overlapped with both schizophrenic and neurotic depressive syndromes.

In the first analysis described in this report, archetypal patients and their symptoms were plotted. In the next biplot, groups of real patients and their symptoms were analyzed. However, with the biplot technique, it is also feasible to plot the distribution of one group of subjects with respect to a second group by using coordinates of variables generated by the second group. It was possible, therefore, to see how the real patients would be distributed by plotting them according to the way their symptom ratings would locate them on the plot generated from the archetypal syndromes. In this way, the real patients were distributed according to plotting "rules" determined by the archetypal syndromes. Results are shown in Figure 3. For simplicity, the points representing the real patients receiving the five most common clinical diagnoses given by the research team are represented by concentration ellipses. For these, the center is at the point of average symptom ratings for the group and the width of the ellipse in any direction reflects the standard deviations of the ratings in that direction. Thus, a group's variability is greatest on symptoms whose arrows are in the direction of the major axis of its ellipse, and least on symptoms in the direction of the

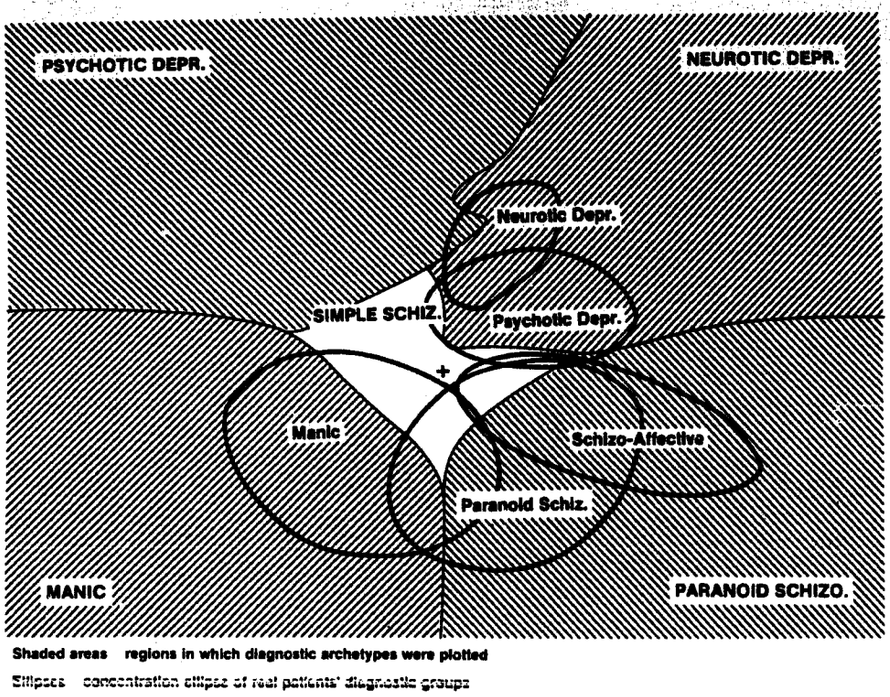

FIG. 3. Archetypal patients and real patients.

minor axis. The plot demonstrates that the diagnostic groups of real patients were generally located in the same segment of the biplot as the archetypal patients with that diagnosis. However, there is considerably more overlap among the various real patient diagnostic groups than there was among the archetypal groups.

Comparing the ellipses of real and archetypal patients for each of four diagnostic groups (Figure 4) shows that there are systematic differences for all diagnostic categories except neurotic depressives. For the paranoid schizophrenics, the psychotic depressives, and the manics, real patients were found to be nearer the mean (at the center of the biplot) than the archetypes. This indicates that real patients had less severe and less distinctive symptoms. For all four diagnostic groups, the ellipse of the real patients is larger than that of the archetypal patients, indicating that the real patients had more heterogeneous symptomatology.

In a separate series of analyses, biplots were generated from symptom data of 100 never-hospitalized ambulatory patients evaluated in the study as a comparison group to the first admissions. The biplot results from the ambulatory patients showed again that the distribution of symptoms formed the archetypal syndromes, but the distribution of patients, as might have been expected with an ambulatory sample, showed the majority to be located scattered evenly in the center area of the plot. This indicated the failure of ambulatory patients, even more than inpatients, to cluster neatly into the diagnostic groups represented by the archetypal syndromes.

Comment

The biplot classified the archetypal patients and their symptoms into clusters and syndromes corresponding to the groups that had been constructed. This indicates the value of this procedure for the kinds of data we were using, and the ability of the biplot to retrieve clusters of patients and symptom groupings if such clusters and groupings do exist. Use of the biplot with real patients indicates that basic syndromes, similar to those built into the archetypal data, also occur among real patients.

In contrast to the results with the archetypal patients, only a few real patients were located in the area corresponding to the archetypal syndromes. The great majority of the real patients either fell between these syndromes or formed a rather homogeneous mass on the biplot in the area representing less differentiated syndromes and less severe symptomatology.

These findings suggest that major traditional psychiatric diagnostic syndromes can be retrieved by using multivariate analysis of symptoms manifest in a

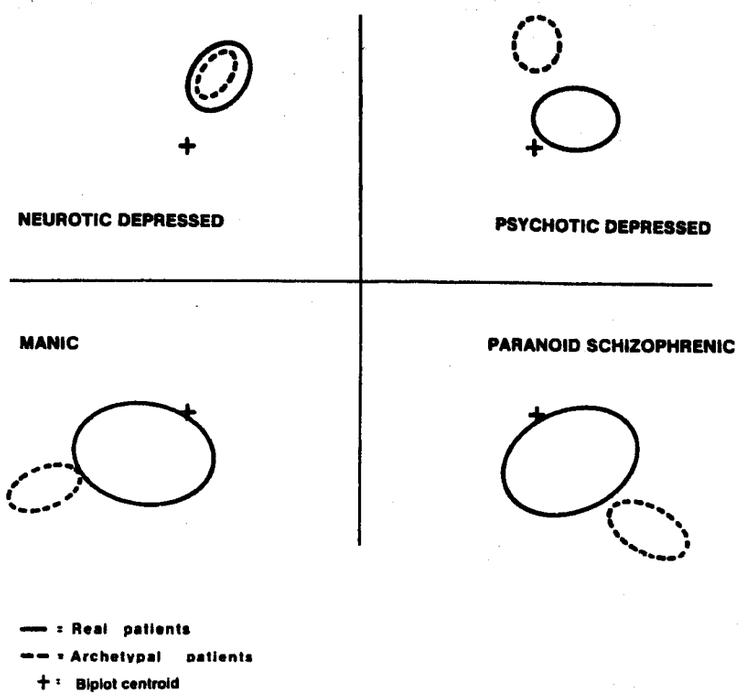

FIG. 4. Comparative biplot locations of real and archetypal patients. The equation for the concentration ellipse of a diagnostic group D is

$$(g' - \bar{g}_D')(G_D' G_D)^{-1}(g - \bar{g}_D) = n_D,$$

where n_D is the number of individuals in that group, g_D their average biplot coordinates, and $(G_D' G_D)$ their sum of squares and products from the averages. For a more complete description of the biplot, see references 13 and 14. For further applications see reference 15. Computer programs for biplotting are available from the second author at the Division of Biostatistics, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642.

representative sample of patients. Thus, these diagnostic syndromes can be derived by statistical procedures as well as by clinical practice from which they originated. However, in contrast to the symptoms, the great majority of first admission patients do not group according to these categories.

The implications of these findings for clinical practice and research are considerable. The diagnostician dealing with real patients is apparently forced to place patients into discrete diagnostic categories based on relatively severe levels of symptomatology, even though a great number of patients actually fall between the categories or exhibit low symptom levels where the categories themselves have relatively little meaning. This lack of correspondence between real patient and diagnostic categories suggests one important source of the dissatisfaction expressed by clinicians and investigators with the present approaches to psychiatric classification. Much of this dissatisfaction may arise from the procrustean structure of the diagnostic system that defines discrete types of patients in contrast to the actual nature of real patients who are distributed along gradations of symptomatology.

These findings are particularly important, since they were derived from a representative sample of first admission patients. Subjects in the First Admission Study were not a select group of a special kind of patients from a particular type of hospital, but were representative of all first admissions for functional psychiatric disorder from two catchment areas. These are the patients that clinicians are called upon to treat and on whom investigators are expected to provide information. Yet both clinician and investigator are led to categorize these patients in ways that simply do not fit the actual nature of the population.

One question that arises is whether these patients appear more homogeneous than they otherwise would because of a leveling effect on symptoms caused by medications. This is possible, but seems unlikely. The symptom data on which these patients were rated covered the month prior to hospitalization as well as the observed behavior at the time of the interview.

The worst level for each symptom over this time period was the level that was rated. Since the patients were first admissions, many of them had not received treatment prior to hospitalization, and those who were treated usually received only limited care (8).

Another question is whether the addition of data on psychiatric history and levels of social function would demonstrate more discrete clusters in the patients. We carried out the biplot analysis to evaluate this possibility and found no increase in discrete clusters. In any case, diagnostic problems from assuming high percentages of discrete syndromes would remain.

Another issue is whether the use of principal components analysis together with the relatively small percentage of variance accounted for by the first two principal components distorts the distribution of syndromes and patients, forcing them into a diffuse mass. If this were so, the clear group structure of the archetypal patients should also have been lost with the biplot. That, of course, did not occur. Further, inspection of the biplot's higher order principal components showed that at most, the third may have accounted for significant additional variance. This third component did not, however, alter the distribution findings reported here. The characteristics of the patients, not limitations of the biplot method, seem likely to be the basis for the distribution patterns found in this study.

Together these findings suggest that a major problem with our current diagnostic typology is that it is based on archetypes which, although occurring in nature, do not provide sufficient coverage of the patient population to provide an adequate diagnostic framework for the great majority of patients coming to treatment. This interpretation is supported by studies carried out by Overall and co-workers (16, 17). These investigators, using a set of archetypal patients rated on a symptom scale different from that used in this study, found that it was possible to have psychiatric raters recognize archetypal symptom profiles designed to fit major diagnostic categories. However, when the profiles of real patients were compared with the archetypal profiles, much less correspondence was found. The authors explained this partly by apologizing for the shortcomings of their profile model and partly by noting the tendency of psychiatrists to exaggerate distinctive characteristics in the process of reaching a diagnosis. We would suggest that our findings indicate that the issue is not so much a deficiency in rating scales or the tendencies of psychiatrists, but a question of the appropriateness of the current diagnostic system for representing real patients accurately. It is important, therefore, to see whether another type of diagnostic system could be developed that more accurately reflects symptom patterns of real patients needing treatment.

What should such a system involve? One logical, although radical answer would be a multidimensional or graphical diagnostic approach where real patients might be plotted in terms of their location on a two-, three-, or four-coordinate system. The multiaxial approach described elsewhere (22) might provide such a framework. Another alternative would be to decide that current diagnostic concepts are valid for a relatively small percentage of patients but need to be modified or new categories need to be added to fit the majority.

At the very least, the findings reported here suggest that considerable exploration of alternatives to the current diagnostic system is needed. Such exploration must consider symptom and patient distribution characteristics, along with clinical criteria for validating psychiatric diagnoses, and the psychological-cognitive needs of those using diagnostic systems (20). By paying attention to all three requirements, it may be possible to develop an approach for psychiatric diagnosis that is more adequate than our current typological system.

References

1. Blashfield, R. K., and Draguns, J. G. Evaluative criteria for psychiatric classification. J. Abnorm. Psychol., *85:* 140–150, 1976.
2. Carpenter, W. T., Jr., Strauss, J. S., and Bartko, J. J. A flexible system for the identification of schizophrenia: A report from the International Pilot Study of Schizophrenia. Science, *182:* 1275–1278, 1973.
3. Eysenck, S. B. G., and Eysenck, H. J. The measurement of psychoticism: A study of factor stability and reliability. Br. J. Soc. Clin. Psychol., *7:* 286–294, 1968.
4. Feighner, J. P., Robins, E., Guze, S. B., *et al.* Diagnostic criteria for use in psychiatric research. Arch. Gen. Psychiatry, *26:* 57–63, 1972.
5. Gabriel, K. R. The biplot-graphical display of matrices with applications to principal component analysis. Biometrika, *58:* 453–467, 1971.
6. Gabriel, K. R., Hill, M., Law-Yone, H. A multivariate statistical method for regionalization. J. Reg. Sci., *14:* 89–106, 1974.
7. Gabriel, K. R., Rave, G., and Weber, E. Graphische Darstellung ven Matrizen durch das Biplot. EDV Med. Biol., *7:* 1–15, 1976.
8. Gift, T., and Strauss, J. S. Schizophrenia: The Limited Use of Preadmission Treatment. Presented at the Annual Meeting, American Psychiatric Association, Toronto, Canada, 1977.
9. Grinker, R. R., Werble, B., and Bryce, R. E. *The Borderline Syndrome.* Basic Books, New York, 1968.
10. Gunderson, J., Carpenter, W. T., Jr., Strauss, J. S., and Scott, W. Borderline patients—A comparative study. Am. J. Psychiatry, *132:* 1257–1264, 1976.
11. Kokes, R. F., Fremouw, W., and Strauss, J. S. Lost subjects: Source of bias in clinical research? Arch. Gen. Psychiatry, *34:* 1363–1365, 1977.
12. Lorr, M. *Explorations in Typing Psychotics.* Pergamon Press, Oxford, 1966.
13. Maxwell, A. E. Multivariate statistical methods and classification problems. Br. J. Psychiatry, *119:* 121–128, 1971.
14. Menninger, K. *The Vital Balance.* Viking Press, New York, 1963.
15. Overall J. E. Dimensions of manifest depression. J. Psychiatr. Res., *1:* 239–245, 1962.
16. Overall, J. E., and Gorham, D. R. The brief psychiatric rating scale. Psychol. Rep., *10:* 799–812.

17. Overall, J. E., and Woodward, J. A. Conceptual validity of a phenomenological classification of psychiatric patients. J. Psychiatr. Res., *12:* 214–230, 1975.
18. Paykel, E. S. Classification of depressed patients: A cluster analysis derived grouping. Br. J. Psychiatry, *118:* 275–288, 1971.
19. Paykel, E. S., Prusoff, B., and Klerman, G. The endogenous-neurotic continuum in depression: Rater independence and factor distributions. J. Psychiatr. Res., *8:* 73–90, 1971.
20. Raven, P. H., Berlin, B., and Breedlove, D. E. The origins of taxonomy. Science, *174:* 1210–1213, 1971.
21. Spitzer, R. T., Endicott, J., and Robins, E. Clinical criteria for psychiatric diagnosis and *DSM-III*. Am. J. Psychiatry, *132:* 1187–1192, 1975.
22. Strauss, J. S. A comprehensive approach to psychiatric diagnosis. Am. J. Psychiatry, *132:* 1193–1197, 1975.
23. Strauss, J. S., Bartko, J. J., and Carpenter, W. T., Jr. The use of clustering techniques for the classification of psychiatric patients. Br. J. Psychiatry, *122:* 531–540, 1973.
24. Strauss, J. S., and Carpenter, W. T., Jr. Characteristic symptoms and outcome in schizophrenia. Arch. Gen. Psychiatry, *30:* 429–434, 1974.
25. Strauss, J. S., Kokes, R. F., Ritzler, B. A., *et al.* Patterns of disorder in first admission psychiatric patients. J. Nerv. Ment. Dis., *166:* 611–623, 1978.
26. Wittenborn, J. R. The dimensions of psychosis. J. Nerv. Ment. Dis., *134:* 117–128, 1962.
27. World Health Organization. *The International Pilot Study of Schizophrenia,* Vol. 1. World Health Organization Press, Geneva, 1973.

Appendix 1: Biplot Construction and Plotting

Biplot Construction

Observations $z_{i,v}$ for n individuals (i = 1, ···, n) on m variables (v = 1, ···, m) are adjusted for means $z_{\cdot v}$ and deviations $y_{i,v} = z_{i,v} - z_{\cdot v}$ arrayed in (n × m) matrix Y. Matrix product Y′Y is analyzed by standard routines for largest and second largest eigen-values λ_1^2 and λ_2^2 and the corresponding normalized eigen-vectors q_1^2 and q_2^2, respectively.

The biplot is constructed from the following matrices:

$$H_{(m\times 2)} = (\lambda_1 q_1, \lambda_2 q_2)$$
$$F_{(m\times 2)} = (\lambda_1^{-1} q_1, \lambda_2^{-1} q_2)$$
$$G_{(n\times 2)} = Y F$$

Row g_i' (i = 1, ···, n) of G is plotted as the biplot point for individual i and row h_v of H as the vertex of the arrow from the biplot origin which represents variable v.

For Plotting Extra Points and Concentration Ellipses

To project m variate observations

$$z_{n+j} = (z_{n+j,1}, \cdots, z_{n+j,v}, \cdots, z_{n+j,m})$$

for further individuals j beyond the original n, obtain the vector of means of the original observations

$$z.' = (z_{\cdot 1}, \cdots, z_{\cdot v}, \cdots, z_{\cdot m}),$$

compute

$$g_{n+j} = (z_{n+j} - z.')F$$

and plot the points g_{n+j} on the biplot.