Biometric Assessment of Mental Patients

Joseph Zubin, Ph.D.

Perhaps it is necessary at the very start to draw attention to the nature of the biometric approach before launching into the problem of assessment. The biometric approach is to be distinguished, on the one hand, from the taxonomic, and on the other, from the statistical approach although it shares common ground with both of these. Biometrics is the science which applies measurement to living organisms—including, of course, man. Taxonomy is concerned primarily with classification of organisms, objects, or entities into categories that are useful and meaningful such as species, genera, diseases, etc. Statistics, of course, is a tool for ordering data and drawing inferences from the obtained orders. Biometrics provides measurements which can be used by taxonomists and which can be analyzed and synthesized by statisticians, but the substantive contribution of biometrics ranges beyond these two applications. It forms an objective basis for aetiological investigations, prognosis, and discovering of new taxonomic categories, as well as the investigation of the relationships within and between organisms in the obtained measures.

While taxonomists utilize biometric data, the nature of such measurement is usually on a primitive level (either nominal or at most ordinal scaling). The distinction between biometrics and taxonomy can be put in the following terms: The biometrician collects data in order to establish diagnoses while the taxonomist collects diagnoses as data by which to justify the categories he has chosen. Thus the biometrician provides data for assessing patients while the taxonomist provides data for assessing diagnosticians. Both methods are valuable, but the distinction must be drawn in order to understand why the techniques and methods may differ in accordance with the goal. Thus while the taxonomist or diagnostician, biostatistician, and biometrician each has a place at this table, their functions and purposes are not identical by any means and bearing these differences in mind, it will be easier to understand the contributions each will make.

Perhaps the following contrast provided by Karl Jaspers will make the distinction even more trenchant. In this passage he refers to the clinical evaluation of body types à la Kretschmer, but it is equally applicable to other clinical evaluations. (Jaspers, 1963)

"The biometric methods give us more than figures and correlations. They foster clarity in all fields in which biometric variations can be established. Moreover, through the application of these methods we have concrete experiences which we would never have had without them.
though they may once more be lost to sight in the purely statistical results.

"...As we proceed we acquire concepts and ways of looking at things which in fact are unanswerable questions. They are the key to wide possibilities but they bring us no definite knowledge. Our mind is kept open for the true whole in its present concrete form as an experience which constantly gains in depth. It is as if we could tangibly grasp it but it evades our clutch while it never opposes our efforts to penetrate yet further and advance."

Thus, while accentuating the need for an evaluation of the totality of the patient, Jaspers nevertheless points out the need for specific measurement as a means of clarifying the total picture.

Why the opposition to measurement persisted and still persists today is one of the paradoxes of science. Perhaps one of the reasons is the tradition, firmly entrenched, that only physicians can adequately study medical problems. This is tantamount to saying that ichthyology should be turned over to fishermen, sociology must be studied only by social workers (Shryock) or astronomy by sea captains (Comte). Fortunately, the present conference is not hampered by such considerations.

Is the biometric method new in psychopathology? Far from it. Pinel was one of the first to use statistical methods (1809–20), and Kraepelin in 1890 had this to say about biometrics itself:

"As soon as our methodology has sufficiently proved itself through experience with healthy individuals, it would be possible to approach the actual ultimate goal of these efforts, the investigation of sick personality, especially of the inborn pathological disposition. In an investigation of many individuals we will always find some who deviate profoundly from the behavior of the vast majority in one or another aspect. If this deviation appears to be damaging to the mental life, and if it reaches a certain degree—which admittedly can only be arbitrarily determined—then we tend to regard it as an illness. Experience teaches us that persons with pathological traits of this kind are, on the whole, in greater danger of a general mental disturbance than those personalities (natures) whose characteristics are in the middle range. We therefore have first of all to investigate whether it is possible by means of psychological tests to determine individual deviations which cannot be recognized by ordinary observations. If that succeeds, we would be in the position, through the quantitative determinations at our disposal, to establish the borderline between health and disease much more precisely and more validly than has been possible so far (1896, p. 77)."

Having differentiated between the clinical and the biometric approach, let us now examine the present status of diagnoses from the point of view of both methods.

There is an uneasy feeling abroad that there is something wrong with our diagnostic system. Whether this is indeed the case or not is difficult to determine but that there has been a loss of interest in the making of diagnoses and a consequent drop in their accuracy and value is undeniably true. The cause of this loss of respect for the diagnostic system has had many sources. Among them are, first of all, the long period during which diagnoses were purely of academic interest because only one treatment — namely, custodial care — was available; second, the repeated demonstration of the lack of reliability of the diagnoses; and third, the low relationships that were discovered with regard to its validity as measured by the course of illness, outcome, and relation to treatment. It is disheartening to note that as the armamentarium of new therapies expanded, and the need for better classification and diagnosis increased as a consequence, the acumen and ability of the diagnosticians seemed to drop.

To counteract this general trend which has led in the extreme to the denial of the
need for diagnosis and to the claim that mental illness is a myth, a new approach to the evaluation of the patient's behavior is essential. While the current methods and techniques in the hands of gifted clinicians yield insight into the behavior of the patient sufficient for clinical treatment, they are unsatisfactory for research purposes. The biometric approach takes as its point of departure the utilization of objective methods in the assessment of patient behavior without regard to current diagnoses. The chief purpose is to document the overt behavior of the patient as well as his mental content in order to provide objective data on the basis of which classification and effectiveness of treatment can be made.

Classification is an essential step in the evaluation of treatment, since evaluation can be made on a group basis only, and the provision of homogeneous groups is the sine qua non of good classification. To this end the entire spectrum of behavior is examined and use is made of the entire range of disciplines from the physiological to the anthropological in obtaining information about the characteristics of the patient. By means of such a wide-angled lens the entire behavior of patients can be viewed and methods for classifying individuals can be obtained on the basis of a variety of behaviors, each of which has some implication for assessment.

The attempt at sidestepping diagnosis and developing techniques for the objective assessment of behavior without reference to current diagnosis has certain dangers. It tends to ignore the current grouping of patients and the classifications which lead to the well known diagnostic categories in the mental disorders. However it also has certain values. It permits the observer to view the behavior without bias and permits the data to be grouped independently. If the current classification categories are educed, well and good. If new classification categories emerge, it would be a gain for science.

The biometric assessment of mental patients involves objective techniques and methods that are less dependent on the subjective and intuitive skills of the clinicians than are the usual methods of clinical evaluation. Where subjective methods have to be resorted to, an attempt is made to make them explicit and public so that others can replicate the obtained data wherever possible. Where no such replication is possible, appeal is made to a consensus on the part of trained observers. Thus, even if we return to the current nosology, we will do so on the basis of biometric data and not on the basis of the subjective, intuitive, albeit insightful clinical approach.

However, before assessment can begin, we must bear in mind some type of framework against which the assessment is to be made. The number of characteristics, traits, and behaviors which man is capable of is infinite and the assessment must selectively be based on some theoretical framework or scientific model. The scientific models which have proved useful in the field of psychopathology are the following: (1) The social-cultural model, (2) the developmental model, (3) the conditioning or learning model, (4) the genetic model, (5) the internal environment model, and (6) the neurophysiological or brain function model. Until recently, the first three models—social-cultural, developmental, and learning, were the most prominent. In recent years, genetics, internal-environment, and brain-function models have become more popular.

Before introducing the aetiological models it might be well to point out that at the present time the most generally accepted approach to psychopathology is the naturalistic approach of description of behavior. This is always the forerunner of more specific models, since without a knowledge of the phenomena no aetiology can be understood. We have had several hundred years of description by astute observers and much knowledge has been
gathered. On the basis of these observations, a rather loosely organized structure has been developed, in which observation and aetiological inferences have been intertwined. It is time now to undertake a critical evaluation of the tenability of the structure and of the validity of the inferences.

The social-cultural model is built on the assumption that all mankind is vulnerable to mental disorders and that given sufficient deprivation, stress-producing loads, or other alterations in our environment, our behavior will be altered to the point where our ability to continue living normally as independent individuals in society is endangered. The evidence for social-cultural-environmental pressures as aetiological agents comes largely from studies of socioeconomic status, isolation, educational and social deprivation and social-cultural uprooting in immigration or migration or rapid acculturation. Even the most sanguine environmentalist will not be satisfied with merely pointing to the above mentioned factors as "causal" agents, but will try to determine just how these malignant factors bring about their deleterious effect. While the story is far from told, there is already sufficient evidence to at least question whether these factors cause mental disorders.

A more thorough review of the evidence which has been presented elsewhere (Zubin, 1963) leads to the belief that social-cultural forces may elicit a mental disorder or even occlude it, but cannot cause it unaided. However, much more research is required to transform this belief into fact.

Despite this conclusion, the fact remains that the detection, diagnosis, and even rehabilitation of the mentally ill today rests on a social-cultural framework, deviations from which identify the mentally ill, and return to which constitutes the basis for improvement.

In summarizing the social-cultural model, we might point out that the current revolution in management in psychopathology with regard to hopefulness of treatment, reduction of patient population, rehabilitation, etc., is to a considerable extent a social-cultural change involving change of attitude on the part of patient, family and therapist. Hence, though social-cultural forces may not be so important in aetiology of some disorders they are of great importance in detection, treatment, and rehabilitation.

The developmental model for aetiology is built on the assumption that mental disease develops as a result of some specific deprivation or interference during a critical period in development when the specific deficit or interference is crucial. Evidence for this model is afforded by the recent investigation of Pasamanick (1961) and his collaborators on the role of intrauterine events on the continuum of "reproductive casualty." They postulate that certain untoward events, such as intercurrent illness, toxemia, and other interference with the foetus during the first 9 months of life will produce mental and physical disability ranging from stillbirth, through live births with epilepsy, cerebral palsy, mental deficiency. Finally, even those who appear unscathed at first may not escape entirely but develop such lesser ailments as reading disability. One of the most exciting events in the developmental area has been the investigation of the impact of early experience on subsequent personality and its deviations. While the evidence from human infants remained controversial, the evidence from animal studies seemed at first to yield data which appeared too closely in keeping with Freud's hunches to give much comfort to those who had refused to accept his clinical surmises. Gentling of animals in their infancy, following the credo of tender love and care, actually produced less emotionally unstable adults, but to the great surprise of most investigators, shocking the infants was equally effective and most recently, Theodore Schaefer et al. at Columbia (1962) has found that merely lowering
the temperature a few degrees is equally effective. Thompson's (1962) demonstration that emotional mother rats give rise to emotional pups, Harlow's (1962) demonstration that monkeys raised on surrogate mothers and not permitted to play with their peers tend to develop poorly in the psychosexual sphere, and Melzack's deprived dogs (1957) are still other triumphs for the developmental model as a possible causal factor in emotional disorders.

The conditioning or learning model postulates that the source of the deviant behavior of the mental patient is to be sought in his reinforcement history. An example of this theory is Bateson's (1956) double-bind model in which the mother's ambivalence in her relationship to her offspring evokes ambivalent behavior and other types of deviation in him which we recognize as schizophrenia. While Bateson's double-bind model has aroused considerable interest in psychodynamic circles, it has thus far defied experimental testing of any of its hypotheses. Several more experimentally founded models have been provided by psychologists. Thus, Sarnoff Mednick (1958) bases his approach on the evidence that the early or acute schizophrenic conditions more quickly, and shows greater stimulus generalization (less steep gradients). These are related to the higher level of arousal which is attributed to early schizophrenia.

With regard to the genetic model, it should be recognized that strictly speaking there are probably no exclusively genetic or environmental disorders. All disorders are both genetically based and environmentally elicited. Without the hereditary-environmental interaction, no disease, in fact no development at all, would be possible. What then is meant by a hereditary disease and by an environmental disease? PKU is a hereditary disease in our particular social-cultural-physical environment because of the presence of phenylalanine in our diet. Had our diet been free of this substance the phenylketonurics in our society would never develop mental deficiency and, in fact, PKU would never have been discovered. On the other hand, if only the poor, or only the mountaineers had developed this illness, and the error of metabolism were unknown, we would have regarded this illness as environmentally produced.

Thus, a disease is regarded as definitely hereditary if we already have evidence of the presence of the specific genetic component and have not yet discovered the specific pattern of environmental components required for eliciting it. On the other hand, a disease is regarded as environmental if we have found the specific pattern of environmental components but have not yet discovered the pertinent genetic factors. In diseases where both the hereditary and environmental aetiology is known the question never arises. In the end, however, all diseases will probably be found to require both a genetic as well as environmental component. Apparently, heredity is no more the cause of an illness than the automobile is the cause of an automobile accident. The interaction between specific hereditary and environmental factors required for the emergence of an illness must be sought if we are to detect the vulnerability of an individual before the illness overcomes him.

Genetics may be viewed as a biochemical mechanism in which the genes serve as precursors for the production of certain enzymes whose absence prevents the organism from prospering. There is, therefore, considerable hope that an investigation of the internal environment of the body may reveal the particular metabolic deficiency or excess which characterizes the patient. A particular error of metabolism may, of course, be inherited or acquired. A considerable amount of effort has been spent in the attempt to relate schizophrenia to metabolic error. Certain fractions of schizophrenic blood have produced metabolic changes and changes in such behavior
as rope climbing in rats as well as transitory changes in the psychomotor behavior of normal human subjects. Presumably similar fractions from the blood of normals do not produce such changes.

Perhaps the most exciting new development is the discovery of the ability to manipulate behavior directly through implanted electrodes. This has given new significance to the work of the neurosurgeons who carried out psychosurgery in the late forties and early fifties. The attempt to detect deviations in the neurophysiological substrate of mental patients through the means of electrodes implanted in various portions of the brain, by recording evoked potential or by introducing current through these electrodes, and the recording of evoked potentials from the intact brain through the scalp, are opening up new possibilities for the detection of the neurophysiological anomalies correlated with deviant behavior.

The final scientific model is the epidemiological model, which seems to be a supermodel including each of the above models as partial factors in the explanation of the mental disorders but requiring careful field studies to determine the relative role of each of them. Mental disorder is conceived as the end result of a series of probabilistic events, each of which must occur in interaction with others to produce the disorder, although the threshold value for each factor may differ from person to person and from one disorder to another. Thus, two people may have inherited the same predisposition, but because of differential stress, nutritional or deprivational factors, will not both develop the illness. The virtue of epidemiology is that it takes in all possible factors ranging from radiation, paranatal existence, genetics, to social-cultural environment, etc. Thus, the epidemiological model both permits and requires the weighing of each of the submodels in the total picture of causation; the difficulties of assessing their relative importance, and of devising studies which will not overlook some of the factors, are too well known to need reemphasizing here.

While many of the social-cultural, developmental, learning, genetic, biochemical and neurophysiological claims remain somewhat in doubt, they nevertheless lend credence to the possibility that the mental disorders will be found to be characterized by either the deviations which are now postulated by these models, or by deviations of the same general scope that have not yet been postulated. Whether they are the cause or the effect of the disorder remains to be seen, but the testing of the hypotheses generated by these models depends to a large extent on the detection of some type of deviant behavior which characterizes the patient.

One question that is often raised regarding the behavior of the mentally deviant individual is the question of its continuity vs. discontinuity with normal behavior. This is an age-old issue which has never been dealt with conclusively. But it is interesting to inquire where the discontinuity, if it exists, is to be sought: in the patient or in the observer—i.e., in the behavior of the patient or in the behavior of the community with regard to the patient? If the discontinuity exists in the community so that, once a threshold of deviation is passed, the individual in question is regarded as abnormal, it becomes clear that we do not have to postulate a discontinuity in the patient's behavior. In fact the individual immediately adjacent to him on the continuum, who differs by only a small degree from him, may still be regarded as normal. Hence, it would be a vain hope to find universal characteristics of mental disorder, since they are detected only by deviation from local norms.

On the other hand, if discontinuities exist in the patient's behavior or in his characteristics (biochemical, metabolic, etc.) a search for such discontinuities would be rewarding. Another possibility to consider is that the discontinuity in the community's behavior or in the patient's behavior may

358
rest not on a single dimension but on patterns of several dimensions. In that case, no single dimension may exhibit any sharp discontinuity. Instead, the patterning of the dimensions may identify the disorder even when all the individual dimensions remain continuous.

Thus, an individual who has an IQ of 70 may still be within the normal range. One who has a social quotient of only 70 may also be within the normal range. But given the combination of IQ 70 and social quotient of 70, the individual may be regarded as retarded by a given community. Here too, it would be a vain hope ever to find universal patterns of mental disorder, since they can be detected only as deviations from local normative patterns. On the other hand, if discontinuities exist in the patient’s behavior or in the patterning of his behavior, a search for such discontinuities ought to be rewarding.

Each of the six models described above demands different approaches to testing its hypotheses. Even the first model, that of naturalistic observation, involves some types of assumptions and special techniques.

The techniques which have been developed are of three types. First is the approach which deals with the culture dependent methods of detecting and diagnosing mental deviation. This can serve the naturalistic and the social cultural model. The second approach is the culture fair approach which can serve the developmental and learning and conditioning models. The third is the culture free approach which can serve the genetic, internal environment and neurophysiological models.

The techniques for measuring the culture dependent, culture fair and culture free techniques can be classified in a Mendelejeff-like table which relates stimulus to response as shown in table 1.

Table 1 demonstrates the stimulus and response classes of which functional analysis may be fruitful in studying psychopathology, under load and nonload conditions. The actual representatives of these classes—
TABLE 1.—Examples of Measurable Activities as Functions of Stimulus Variables

<table>
<thead>
<tr>
<th>Level of observed behavior</th>
<th>Idling state</th>
<th>Energy variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>None</td>
<td>Appropriate energy</td>
</tr>
<tr>
<td></td>
<td>Variable</td>
<td>Function</td>
</tr>
<tr>
<td>Sensory</td>
<td>Background noise; cortical gray.</td>
<td>Light of graded intensity.</td>
</tr>
<tr>
<td>Physiological</td>
<td>BMR; basal ECG; basal PGR.</td>
<td>Increase in carbon dioxide concentration.</td>
</tr>
</tbody>
</table>

sible to detect regularities which can be shown to be related to the sensory component (modality), the conceptual component (the guess), and to the emotional component (effect of being right or wrong). By contrasting patients and normals in this experiment the relative deviation in the sensory, conceptual, and emotional components can be established (Sutton, Braren, and Zubin, 1965). While we have obtained data only on normals with this technique so far, the problem of extending it to patients is simply a matter of time.

But it is rarely possible to investigate all behavioral levels for all stimulus classes at once. While no single family of techniques is adequate to explore all areas of functioning in an individual or group of the mentally ill, we may profitably return to an examination of what we have called the culture-dependent, culture-fair, and culture-free indicators to provide examples of the techniques that are being used to good effect.

Culture-Dependent Techniques

The culture-dependent techniques are based on the assumption that mental disorders reveal themselves through deviations from social-cultural norms. The initial detection of mental disorders is usually made by laymen—the patient himself, his family, friends, or public officials. Since the identifying behavior is usually unexpected or deviant, it is no wonder that different cultures will not recognize as mentally disordered the same kinds of behaviors. This may be the reason why such great differences are observed between cultures with regard to the incidence of various forms of disorder. Once the deviation is observed and the
<table>
<thead>
<tr>
<th>Level of observed behavior</th>
<th>Stimulus variables</th>
<th>Signal variables</th>
<th>Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Configurations</td>
<td>Signs</td>
<td>Symbols</td>
</tr>
<tr>
<td></td>
<td>Variable</td>
<td>Function</td>
<td>Variable</td>
</tr>
<tr>
<td>Conceptual…….</td>
<td>Aircraft forms</td>
<td>Recognition</td>
<td>Practical</td>
</tr>
<tr>
<td></td>
<td>or silhouettes</td>
<td>of identity</td>
<td>trouble shooting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>of forms.</td>
<td>test.</td>
</tr>
<tr>
<td>Psychomotor…….</td>
<td>Star-shaped</td>
<td>Mirror tracing.</td>
<td>Classical</td>
</tr>
<tr>
<td></td>
<td>maze.</td>
<td></td>
<td>delayed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>response</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>stimuli</td>
</tr>
<tr>
<td>Perceptual…….</td>
<td>Rotating</td>
<td>Subjective</td>
<td>Usual visual</td>
</tr>
<tr>
<td></td>
<td>Benham Disc.</td>
<td>color experience.</td>
<td>alternatives</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>in animal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>discrimination</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>experiment.</td>
</tr>
<tr>
<td>Sensory………</td>
<td>Patterned light</td>
<td>Visual</td>
<td>Infant’s faint</td>
</tr>
<tr>
<td></td>
<td>stimuli.</td>
<td>threshold.</td>
<td>ery.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiological……</td>
<td>Photic driving.</td>
<td>Change in EEG</td>
<td>Bell ringing in</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pattern.</td>
<td>Pavlovian</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>conditioning.</td>
</tr>
</tbody>
</table>

The suspected patient is brought for diagnosis, the clinician begins to interview the patient to determine whether he is in fact mentally ill, how severe the condition is and what can be done about alleviating it. We shall consider here only the diagnostic procedure. What does it consist of?

Here again, the chief concern of the diagnostician is to determine whether the behavior exhibited by the patient is the type you would expect from a normal individual. We probe for feelings of hostility, depression, aggression, anxiety, apathy, affect, delusions, hallucinations, but our probing is simply to determine whether the patient deviates from the expected in these kinds of behaviors. The best technique thus far developed for this probing is the interview, despite its subjectivity and unstandardized character. However, it need not remain as unstandardized and unreliable as it is now. The introduction of standardized interviewing methods of the type provided in the “Mental Status Interview and Inventory,” is a first step in providing a biometric approach to assessment (Spitzer et al., 1964).

The primary design for objectifying and standardizing the interview and observational approach is best exemplified by the animal observational studies reported by Hebb and Thompson (1954) on the recognition of social attitudes in chimpanzees. Noting that the seasoned caretakers refuse more often than not to say what the motivation behind an animal’s act might be unless they have had considerable experience with it, Hebb and Thompson examined carefully just how the caretakers and staff members eventually develop confidence in their understanding and prediction of a given animal’s behavior.
order category (e.g., aggression not preceded by quasi-aggression)."

The significance of the change from second-order to third-order categories is demonstrated dramatically in the relation of male versus female animals' behavior to the staff.

No comparison was made of the first-order categories because of their disparate nature and uneven distribution. A comparison of the second-order categories indicated that although the two sexes were equally friendly, the males excelled in quasi-aggression and in aggression. However, when the third-order categories were compared, it was found that the pattern of quasi-aggression followed by aggression occurred only in males, while seductive, friendly behavior followed by aggression occurred only in the females.

It is clear that mere atomistic enumeration of behavior is difficult to integrate into meaningful observations. However, second and third order categories and even higher categories where three or more events are included in a pattern promises to be useful. The parallel method followed in the evaluation of observed behaviors during structured interviews is to first note the molecules of behavior in items, then group the items into second-order categories or scales and finally find the pattern of scales which represent the profiles of the various groups of patients.

Reliability and Validity of Scaling

It is quite apparent that each member of the clinical team, the clinical psychiatrist, clinical psychologist, ward nurse, attendant, social worker, occupational and physical therapist, etc., makes some unique contribution to the understanding of the patient's pathology. Because of the different observational situations afforded each member, a different technique is required for eliciting and recording the observations in each clinical context. The development of such techniques was begun by our Biometric Re-
search Group and the following techniques are now available in various stages of completion: (1) The Ward Behavior Inventory, (2) Mental Status Schedule, (3) Structured Clinical Interview, (4) Children's Behavior Inventory, and (5) Social Adaptation Inventory and Anamnesis.

The Ward Behavior Inventory (Burdock et al., 1960)

The first result of our efforts was the provision of an instrument for recording the behavior of the patient as observed by nurses, attendants, or special observers on the ward. Material relevant to ward behavior was culled from the literature and in 1953 a Hospital Adjustment Rating Scale was introduced for use in our prognostic studies at the Brooklyn State Hospital. This became the starting point for the development of the Ward Behavior Inventory.

This instrument capitalizes on the observational opportunities afforded nurses and ward attendants in everyday contact with the patients. The WBI consists of 188 items rated true or not true. All the items describe observable behaviors as seen during a specified interval of a single day. Neither retrospective material nor dynamic inferences are included. Total scores reflect global pathology and provide evidence of change in response to treatment. The Ward Behavior Inventory has proved itself reliable when raters are properly trained and motivated (Burdock, Hakemel, Hardesty, and Zubin, 1960). Reliability coefficients ranged from 0.40 for untrained raters to 0.84 for well-trained and well-motivated observers. Followup studies indicate that the global pathology score correlates at a low but significant level with outcome in terms of length of time out of hospital. WBI scores of 107 patients followed for 1 year had a correlation of 0.23 with an Outcome Index (1961). Several drug studies have proved the instrument sensitive to behavioral change. Among these are the NIMH 9-hospital collaborative study of phenothiazine therapy conducted by the Psychopharmacology Service Center (Cole, J., 1964) and an unpublished study of the efficacy of a psychic energizer for seniles.

Future work with the WBI will focus on analysis of patterns of items and identification of clusters of patients possessing similar symptomatology. In order to obtain more objective data than those made available by the overburdened ward nurse, specially trained research nurses are to be included in our research team. The new data to be used in the construction of subscales will be based on a suitably sampled population of patients drawn from the Washington Heights area in New York City.

The Mental Status Schedule (Spitzer et al., 1964)

In 1943 I assisted Dr. Nolan D. C. Lewis in the revision of the third edition of the "Outlines of Psychiatric Examination," the classic handbook for guiding the new resident in his mental examination of patients. A mimeographed sheet was provided for checking the pertinent items regarding the psychopathology of the patient. This was used for a short time and was included in the case history, but the rapid changes brought on by the end of World War II soon overwhelmed the training of residents and the use of the Mental Status Schedule fell into discard.

Beginning in 1960, a new attempt at standardizing the interviewing of psychiatric patients was initiated. This culminated in the development of two instruments, the Mental Status Schedule and the Structured Clinical Interview.

The Mental Status Schedule consists of an interview schedule for the mental status examination and a matching inventory of 248 dichotomous items descriptive of small units of pathological behavior. The schedule contains questions arranged in a definite sequence designed to provide for followup of incomplete responses. Most of the ques-
tions are open-ended so as to encourage the patient to reveal his own mentation. Properly administered, the interview has the feel of the clinical evaluation. However, unlike the usual clinical interview, the provision of a specific schedule of questions, a fixed order of presentation, and a uniform coverage of the same areas of psychopathology with each patient makes it more likely that the differences observed will be due to actual differences among patients rather than to different interviewing procedures. This technique has yielded reliabilities of the order of 0.90 or more when groups of patients were evaluated independently by three psychiatrists. Moreover, it has distinguished significantly between the amount of psychopathology shown by inpatients, clinic outpatients, and former inpatients on followup.

Recently, scoring systems for two types of subscales have been developed.

Clinical subscales.—Items from the MSS were grouped into 20 scales which seemed to reflect the clinical dimensions most frequently used in descriptive psychiatry. These subscales are being revised and improved by analyzing the data from a study in which clinicians made global clinical judgments on the 20 dimensions after they had interviewed each of a series of patients with the MSS. The contribution of each item in a scale to the clinical judgment of that dimension was estimated in order to determine the weight which the clinician gave to the item and thus to test whether its placement in the subscale was correct. Since some of the subscales have been revised, we plan to retest their reliability and validity.

Factor subscales.—The 248 MSS items have been grouped into 138 clusters of from 1 to 4 items each in order to have a more manageable number of variables for factor analysis. These clusters were formed according to the following principles: (1) items were combined which reflected different aspects of the same behavior or feelings; (2) items reflecting an observed behavior were not combined with items reflecting reported behavior; and (3) items associated with different diagnostic categories were not combined. The correlation matrix for the 138 clusters of items has been obtained from a sample of 529 patients, and the process of extracting factors has begun.

Of late we have devised a series of programs for biometrics research for computing for each patient or for a group of patients (1) the total score (number of items judged true), (2) the 20 clinical subscale values (above), (3) the correlational matrix for these 20 clinical subscales, (4) the grouping and scoring of the 138 clusters (above), and (5) the individual item frequencies. These programs are available for use by investigators employing the MSS, and a computer program is being developed for the scoring of the factor subscales.

Audiotape recordings of MSS interviews have been made for training interviewers and for estimating their reliability and bias. A method for estimating accuracy has been developed, whereby a number of experienced clinicians have come to a consensus as to which items are true, false, or possible for a number of teaching tapes. A rater listens to and scores a number of recordings, and later notes for each item whether he had a “hit” (he and the experts both agree that the item is either true or false), an “add” (he rated it true, experts rate it false), or a “miss” (he rated it false, experts rate it true). The degree of inaccuracy and the direction and sources of bias are thus indicated to the rater.

Structured Clinical Interview (Burdock, 1962)

The purpose of this instrument is to focus on the general social and psychological adjustment of the interviewee. It consists of open-ended questions that are quite neutral in tone and content but which nevertheless permit the interviewee to express psychopathological ideation and
behavior. For this reason it is most suitable as an initial instrument for community surveys as well as for admission services in hospitals and clinics.

If this instrument yields a score which makes the interviewer suspect that he is dealing with a possible psychiatric case, more probing instruments like the Mental Status Interview can be applied to determine further the nature and extent of the psychopathology. If the interviewee fails to show any psychopathology on this instrument he can be tentatively regarded as normal. Interviewees who give evidence of psychopathology in this neutral instrument tend to have severer illness than those who require probing to elicit their psychopathology.

The instrument consists of an interview schedule of open-ended questions and 178 inventory items which are to be marked “true” or “not true” on the basis of the interviewee’s answers. The reliability of the total score, which represents the severity of psychopathology exhibited by the interviewee, has been studied by comparing the results of two or more simultaneous but independent observers. Intraclass correlations among observers have ranged from 0.83 to 0.92 for groups of patients ranging in size from 26 to 83. With score on the Ward Behavior Inventory as a criterion for validity, the coefficients for the SCI of 0.30 to 0.68 were found, while the coefficient for an unstructured interview with the same criterion was only 0.22. The following subscales have been developed: Anger-hostility, conceptual dysfunction, fear and worry, incongruous behavior, incongruous ideation, lethargy-dejection, perceptual dysfunction, physical complaints, self-depreciation, and sexual deviance.

Future work with the SCI will be in the direction of developing norms based on suitable samples drawn from the general population and from patients.

Social Adaptation Schedule (in preparation)

This technique follows the form of the Mental Status Schedule in that the patient is examined by means of a structured interview and observations are recorded on an inventory of dichotomous items descriptive of small units of pathological behavior. However, it differs from the examination of the mental status in that the focus is not on symptomatology but on disturbance of function in the context of social adaptation. The patient is examined for evidence of disturbed functioning in the following areas of adaptation: (1) Use of leisure time, (2) friendship patterns and involvement in social activities, (3) work adjustment, (4) sexual and marital adjustment, (5) school or vocational training, and (6) level of aspiration. The technique has been designed so that it can be used to supplement the examination of the mental status or can be used on a separate occasion as an independent instrument.8

With regard to the anamnesis, we have not yet developed an instrument for this area. It is important, however, to realize that obtaining information on the entire spectrum of a person’s developmental history is a research task of the first magnitude. In the experimental approach to this problem it is necessary to develop techniques which will focus on particular critical periods in the history of the individual. It is proposed to begin with a focused interview dealing with the question of the status of the patient a year before he came for help and, by focussing on this particular period, to arrive at more objective information.

A second focus might be the adolescent friendship patterns of the patient, since in some current research we have found this phase of development important in prognosis, especially with regard to type of onset. A third focus might bear on early

---

8 The Mental Status Schedule has recently been combined with the Social Adaptation Schedule to constitute a more comprehensive instrument, known as the Psychiatric Status Schedule.
childhood, especially based on history obtained from the family. In all of these aspects of the anamnesis, at least two members of the family ought to be interviewed for checking purposes. While this may sound impractical for everyday use, it should not be beyond the range of a research project. In fact, the Katz Adjustment Scales (Katz, M., and Lyerly, S. B., 1963) are based upon interviews with relatives of patients. Hopefully, the more intensive methods used in this research can form the basis for more practical approaches in the future.

The Children’s Behavior Inventory
(Burdock & Hardesty, 1964)

This inventory is being used in the collection of observations of children hospitalized for physical illness as well as for collection of observations in children’s services of psychiatric facilities. Three successive 2-hour observations are made on an observation day by specially trained psychologist observers. This technique allows for determination of frequency of pathological behaviors over the observation day. Comparisons of the hospitalized normals with the psychiatric patients will lead to development of normative criteria.

An example of the value of such systematic interviews is demonstrated by the following experiment (Katz et al., 1966). One of our structured interviews was videotaped so that it could be presented to groups of psychiatrists for their clinical judgment and diagnosis.

![Diagram of symptom profiles](image-url)

**Mean symptom profiles of a psychiatric patient as viewed by two groups of clinicians who differed on diagnosis**

(Reproduced by permission from Katz, M.M., Cole, J.O., & Lowry, Henri A. Studies of the diagnostic process. Unpublished manuscript.)
In addition to the overall diagnosis the psychiatrists were asked to fill out ratings on an inventory for such factors as excitement, paranoid projection, anxious intro- 

punitiveness, perceptual distortion, motor disturbances, hostile belligerence, apathy and retardation, grandiose expansiveness and thinking disorganization (Lorr, 1953). The psychiatrists were all seasoned veterans of psychiatry; nevertheless, of the 35 participants, 14 diagnosed the patient as neurotic and 21 as psychotic. An examination of their ratings, revealed however, that the groups differed significantly only in one respect—the rating on apathy. Those who rated the patient high on apathy diagnosed him as psychotic, while those who rated him low on apathy diagnosed him as neurotic.

Culture-Fair Techniques

In addition to direct interviewing with the structured instruments, certain focussed aspects of interviewing procedures have been utilized for more objective evaluation of specific patient characteristics such as flatness of affect, intelligibility of speech, and similar traits. While the structured interviews may elicit material that is largely culture dependent, the techniques to be discussed now might belong in the culture- 

fair category. By this we mean techniques that, though imbedded in a given culture, nevertheless have equivalents in other cultures. Examples of such culture-fair behaviors are response to greeting, bereave- 

ment, reinforcement, or praise, etc. All cultures exhibit behavior in response to such situations, but the type of behavior is colored by local norms.

In the course of our studies on the experimental analysis of the interview we have found that the rate of emission of self-referred affect statements can be lawfully modified in acute schizophrenics by the asking of questions which serve as discriminative stimuli for eliciting affect (described in Zubin, 1958), and by reinforcers such as verbal expression of agreement—e.g., mhm, yeah, etc. (described in Zubin, 1958; Salzinger and Pisoni, 1958, 1960, 1961). We were able to show that the process of conditioning is effective both in the beginning or middle of the interview, that a minimum number of reinforcements is necessary for conditioning to take place, and that rate of extinction varies as a function of rate of acquisition (Salzinger and Pisoni, 1961).

Rate of emission of affect statements was also successfully modified in normal subjects, and a comparison of normal and acute schizophrenic subjects showed the schizophrenics to extinguish faster than the normals (Salzinger and Pisoni, 1960).

Similar interviews with chronic schizophrenics led us to conclude that they can be more accurately described as having a low rate of responding in general, rather than as having "flatness of affect" in particular. Their low operant levels for self-referred affect made it impossible to measure their reactivity to reinforcement (Salzinger and Portnoy, 1964).

Most recently we have devised a technique which makes interviewer questions unnecessary. Under these conditions we have obtained monologs—extended uninterrupted samples of continuous speech from acute schizophrenics. We found that, depending upon the specific reinforcement contingency in effect, we could either increase general speech rate or increase the rate of a specific response class like self-referred affect (Salzinger, Portnoy, and Feldman, 1964a).

Comparisons of the rate of repetition of words in schizophrenic and normal speech samples revealed that schizophrenics tended to use the same words more frequently than normals for passages of 900 words in length (Hammer and Salzinger, 1964).

Furthermore, rate of conditioning of schizophrenic patients (i.e., the degree to which a patient becomes influenced by the interviewer) appears to be a prognostic indicator. Those patients who condition
better are more likely to leave the hospital at the end of a 180-day period (Salzinger and Portnay, 1964).

Communicability

In order to measure the degree of communicability of schizophrenic speech, 200-word samples from the above-mentioned interviews and monologs were selected and subjected to the cloze procedure (the cloze procedure consists of systematically deleting words from continuous speech passages and having other subjects guess the words which have been so deleted). Schizophrenics' and normals' samples were compared and we found that normals communicate better than schizophrenics (a larger number of correct words were guessed). Furthermore, differentiation between the patient and normal was better for the second 100 words of the sample than for the first 100 words (Salzinger, Portnay, and Feldman, 1964b).

We have recently replicated the above study using the same samples but a different group of predictors and separating the 100-word samples (each subject got all samples, but the 200 words were not presented consecutively). The results were almost identical with those obtained in the earlier study. Despite the separation of the two 100-word samples the second 100 words differentiated better the schizophrenic and normal than the first 100 words, and speech from schizophrenics continued to be less communicable.

We have also added two other techniques for assessing the communicability of schizophrenic speech. The first we have referred to as the "method of reconstruction." It consists of having subjects sort speech segments of equal length into the order they think they were emitted by the speaker. Comparisons of the schizophrenic and normal samples revealed that subjects had greater success in restoring the original order of the normal than of the schizophrenic samples (Salzinger, Portnay, and Feldman, 1966).

The second procedure we have referred to as the "method of unitization." It consists of having subjects divide unpunctuated speech samples into grammatical sentences and cross out words they feel do not belong. The results indicated that schizophrenic speech is characterized by having a larger number of intrusions (i.e., words crossed out) than normal speech samples. Number of units and length of unit did not differentiate normals from schizophrenics.

We have also related the cloze procedure and reconstruction method results to outcome of illness and found that the better a patient communicates the shorter his stay in the hospital, and the more difficulty the subjects found in ordering the speech samples of patients the longer these patients remained in the hospital.

The results of these more recent studies together with the results of the earlier ones have led us to propose the notion that the basic deficit in the schizophrenic patient's behavior consists of his being controlled primarily by stimuli immediate in space or in time. It is for this reason that he has difficulty in sorting, shows less object constancy (at least in the more deteriorated stages), that he extinguishes faster after conditioning, and that in language he shows "loose association" and poor ability to communicate—since language communication requires that the speaker react to long-range stimuli and not simply to short-range associations (Salzinger et al., 1966).

Culture-Free Indicators

Though no measures can be said to be completely culture free, the way in which culture might affect such measures as pupillary response to light stimuli is indirect, unlike the direct way in which culture influences primarily conceptual measures like vocabulary. The major way in which culture will tend to influence the
culture-free or -fair tests is likely to lie not in the function under measurement but in the subject's approach to the testing situation—e.g., in the subject's understanding of the purpose of the test, in the degree of fear experienced, in his motivation, attention, and cooperation, etc.—in other words, specifically in those variables which also tend to contaminate comparisons of schizophrenics and normals even when they come from the same cultural background. But it is these very contaminations that the experimental culture-fair or -free techniques have been constructed to minimize.4

One of the best established findings with regard to schizophrenia is that schizophrenics are retarded in reaction time as compared to normals. Since speed of response is to a large extent motivationally determined, it could readily be argued that the difference between schizophrenics and normals with regard to reaction time is not specific to schizophrenia. In order to eliminate this source of irrelevant difference, motivation must be either measured, permitted to vary at random with regard to the experimental conditions, or kept at a high level.

Since measurement of motivation is rather difficult, we chose first to sample from situations in which motivation was uncontrolled. The subject was instructed to lift his finger off a brass plate just as soon as he saw the light or heard the sound which served as stimuli. There were two light stimuli, red and green, and two sound stimuli, high and low tones. The stimuli were presented in random order. Our chief concern was not with the primary reaction times but with the influence on reaction time of shifting from one stimulus to another. There were 3 types of sequences possible: (1) No shift, the second stimulus being identical with the first; (2) shift within the modality (red to green or vice versa, high to low tones or vice versa)—ipsimodal; and (3) shift between modalities—crossmodal (Sutton, S., and Zubin, J., 1965).

In each case, the reaction time to the second member of the pair of stimuli is recorded. We found that normals are faster on all accounts, and that reactive schizophrenics are faster than process schizophrenics. The influence of overall speed was eliminated by performing covariance analyses comparing groups on crossmodal stimuli and on ipsimodal stimuli with performance on the no-shift category held constant. Both of these analyses discriminated patients from normals on reaction to sound, but not on reactions to light. Analyses comparing process and reactive patients were not significant.

The experiment has been repeated with a number of variations in a number of populations. In summary, it may be said that there is a consistent tendency across several studies for schizophrenic males to be disproportionately impaired on reactions to sound stimuli if the stimulus in the previous trial is a light. These findings hold in both acute and chronic patients. This impairment is over and above the differences in overall level of response between groups. Comparable analyses for reactions to light do not yield differences between patients and normals.

In a comparable study done by Benton et al. (1962), on brain damaged subjects, at Iowa, there was an exactly opposite finding. The shift from light to sound did not discriminate brain-damaged subjects from normals, but the shift from sound to light did. This intriguing reversal between brain-damaged subjects and schizophrenics has not to my knowledge been followed up.

In interpreting these data one must make some assumptions about the state of readiness of the subject. It is reasonable to assume that, in all reaction time experiments, some aspect of readiness of the subject is

---

4 Such designs include obtaining complete functions rather than isolated points of measurement for each individual, comparisons of slopes of functions rather than level, the measurement of functions under idling state conditions and under load conditions, and finally, the use of the range of variation within a population as a basis for assessing observed deviation.
fluctuating from moment to moment. For these experiments, there must also be some such process, although how one prepares to observe with the eye rather than with the ear, or vice versa, is less apparent. In this connection, one might speculate on the possible role of descending inhibitory pathways which are present in most if not all sensory modalities. Shorter reaction time may then be taken as evidence that the stimulus occurred at a point in time corresponding to maximal readiness, or that the stimulus is one for which maximal readiness has been made. When the stimuli do not fall in with these expectancies, there is retardation or lengthening of reaction time. As has been shown, this state of readiness is strongly influenced by the character of the preceding trial—viz, whether the stimuli are in the same or a different sensory modality.

With respect to the differences between schizophrenic patients and normals, the studies strongly suggest that the state of readiness of the patient is disproportionately affected by events which are recent in time. In other words the patient is over-prepared by his last trial and reaction time is lengthened when the unprepared-for stimulus appears. In the random program, since all sequences are equally probable, best performance would occur if the subject assumed (correctly) that the next stimulus may be of any of the four used. Therefore preparation for any one type of stimulus is maladaptive.

Our further studies with normals have made it apparent that in these experiments there are two effects rather than one at work, one of which may be called the probability effect, in which reaction time to the uncertain or less probable stimulus is lengthened, and the other of which may be called the sensory effect. The sensory effect is revealed by the fact that manipulation of the probability of cross-modal and ipsimodal stimulus does not produce identical results—there is something special about cross-modal stimuli which makes them peculiarly vulnerable to the state of certitude of the subject. We have not yet decided whether the differences between patients and normals are due to one of these effects, or to both. These observations have, incidentally, been supported in our evoked potential studies. In these we have found dramatic amplitude increases in the evoked potential to sensory stimuli when the subject did not know in advance which of two stimuli would be presented. However, the program in which the cross-modal stimulus is uncertain is the one that has produced the larger amplitude changes, presumably as a result of these subjects’ greater degree of incertitude (Sutton et al., 1964).

Time does not permit a complete summary of our results in pupillography, nor certain other areas where cultural explanations of the observed differences are hard to find. In an initial study we found that the dark-adapted size of the pupil of acute schizophrenics, but not of normals, is smaller than that of comparable normals (Hakem, Sutton, and Zubin, 1964). This held true significantly for males but does not reach significance in females. Both the acute and chronic schizophrenics reach the point of maximum contraction upon stimulation by light sooner than normals. However, there were no differences between patients and normals in the extent of contraction. We concluded following the interpretation of pupillary diameter offered by Lowenstein and Loewenfeld (1950) that the acute patients show a deficiency in the inhibitory control of the Westphal-Edinger nucleus by higher centers. In order to demonstrate this deficiency in chronic schizophrenics, we have to subject the pupil to a stimulus, or put it under some load.

However, our more recent work has seemed to show almost opposite results, in the sense that normals and schizophrenics (both acute and chronic), showed similar initial pupil diameters after standard dark-adaptation but both groups of schizophrenics showed contraction amplitudes
markedly different from the normals'. In our previous experiments, the light stimulus was of one-second duration, while more recently we used briefer stimuli averaging 30 to 40 msecs. A series of parametric studies will obviously be needed to resolve the apparent conflict. The nature of the patient sample must be more completely specified as to history, treatment, accuracy of diagnosis, etc., and a possible curvilinear light-stimulation function investigated. It may be that measurements close to visual threshold may yield differentiating results quite different from those obtained under supraliminal loads.

Instead of sampling or measuring motivation we can take another tack. If we find functions in which patients are more sensitive to the stimulus input than normals, we might deduce that motivation must be at a high pitch. We thought that the response to delayed auditory feedback of speech might be such a function, since schizophrenics might be bothered less by the delay and could, therefore, perform better. Unfortunately, this is not the case; schizophrenics have shown greater interference than normals. We have, however, found several functions in which schizophrenics excel. Although these results are still tentative, they do encourage the hope that more results of this type may be expected. The Bunsen-Roscoe or Bloch's law seems to hold for normals when reaction time is used as the response to the varying parameters of intensity and duration during a critical period of 10 to 20 msecs. Schizophrenics apparently have not heard of this law and fail to obey it, or have a much shorter critical duration than normal.

Similarly in dichotic auditory stimulation experiments, when one ear receives either a more intense or a slightly earlier stimulus than the other ear, the localization of the sound will deviate from the center in the direction of the ear receiving the more intense or earlier click. The intensity and time differences can be matched so that the subject cannot discriminate the displace-

ment from center brought about by the two techniques. Preliminary results indicate that values cannot be found at which the schizophrenic will not discriminate the ΔI stimuli (intensity differences) from the Δt stimuli (time differences). In other words, the schizophrenic is detecting differences which the normal cannot detect. In several other ways, it becomes apparent that greater sensitivity to, or lack of integration of, input is characteristic of the schizophrenic. It is very difficult to find cultural explanations for such differences.

One of the more exciting techniques for the study of differences between patients and normals is the use of evoked potentials by Shagass and Schwartz at Iowa (1961; 1962) to measure cortical excitability and responsiveness. These workers recorded evoked potentials to two discrete, electrical stimuli presented to the wrist and separated by brief periods of time. From these electrophysiological data they plotted the resulting recovery functions for the effects of the first stimulus upon the amplitude of the evoked potential to the second stimulus. Some types of mental patients—depressives—were found to require a longer temporal separation between pulses than normal subjects for the amplitude of the second evoked potential to return to its base line, suggesting that they display less cortical excitability and responsiveness than normals.

With the aid of objective measures of both varieties—the culture-bound as well as the culture-free—a new approach can be made to the development of more homogeneous diagnostic groupings. At the present time, much of the research in psychopathology suffers from the great heterogeneity which characterizes today's diagnostic classifications. Even when there is full agreement on the part of a diagnostic team regarding such diagnoses as schizophrenia or neurosis, the individual patients in each category are far from a homogeneous group even with respect to the variables which presumably characterize each category. Bleuler
has pointed to this fact by substituting for Kraepelin's "dementia praecox," the term "group of schizophrenias." The heterogeneity in the category of neurosis is too well known to require documentation.

Herenfore, the technique applied to the statistical classification of patients has been factor analysis, either direct or inverse, of the correlations obtained from a sample. Such techniques are elegant and quite suitable if the underlying population from which the sample is drawn has a multivariate normal distribution. Were this the case, there would be no need to search for better classifications, because to begin with, our population would already be homogeneous. However, if we start with an assumption of homogeneity and wind up with a conclusion of heterogeneity, the only justifiable procedure is to reject factor analysis as inapplicable.

For this reason, once heterogeneity is established, we discard factor analysis and base our new analysis on the test profile of each patient expressed in standard scores. The usual method of correlating profiles (inverse factor analysis) takes into consideration only the shape of the profiles and neglects entirely the distance between profiles. For this reason, we deal with two measures of distance: discrepancy in shape and discrepancy in level, and group together all patients who show the greatest similarity with respect to these two measures. In this way, under certain assumptions (normality of multivariate distributions as well as zero correlation within each subgroup or some similar assumptions) a fractionation of the sample into homogeneous subgroups becomes possible.

REFERENCES:


Summary Presented at Conference:

My presentation can be condensed into the following four points: (1) The contrast between the biometric and the clinical approach of diagnosis; (2) the evaluation of the current status of diagnosis; (3) the preparation of instruments to improve present-day diagnosis; and (4) organization of the data resulting from such an attempt.

The clinical approach is too well known to require description. The essential element of a good clinical diagnosis is that it constitutes a creative act which belongs to the realm of discovery, using Reichenbach's term, rather than to the realm of verification. Such verification as is in-
dulged in consists of impressionistic testing of successive hypotheses that arise in the course of the clinical examination.

The biometric approach takes as its point of departure the objective measurement of the behavior of the patient, applies taxonomic principles for classifying such behavior in profiles of the systems, and finally appeals to statistical evaluation for determining the reliability and validity of a diagnostic label placed on the patient, as well as his prognosis and the outcome of followup.

Both approaches are essential, since, without the clinician there would be nothing to measure, and without measurements the clinician would soon become a prey to whatever wind of fashion blows.

One might say that the clinician provides the hunches and the inspiration. He is the one who storms the beachhead, while the biometrician may be likened to the foot soldier who secures the beachhead and organizes the territory, bringing law and order into its interior.

In short, without the clinician there would be no problem to solve; without the biometrician there would be no reliable solution.

There is one more difference. The final diagnosis—at least at the present stage of our development—must perforce be made by the clinician who, like the judge in court, despite the paucity of evidence, must arrive at a decision. That is why a study of the decision process itself, as Meehl and as Ham mond propose, is so essential.

Biometrics is essentially nomothetic, not concerned with the diagnosis of an individual case; but it can buttress the idiographic approach very nicely.

At the present time our diagnostic schema consists of a mixture of descriptive terms observed by the clinician and etiologic considerations arising from theoretical structures. The symptoms which the patient complains of, and the signs of deviant behavior which the clinician observes, are embroidered into a structure of postulated causes interwoven with dynamics, and the resulting structure is one that is so complex that the reliability and validity of the end results are very difficult to establish.

The symptoms and signs which constitute the present-day warp and woof of diagnostic procedure have developed over a long period of time, and have been primarily derived from the observations of patients who were hospitalized. Only now we are freeing ourselves from too great attention to "hospitalism," which was an essential feature of institutionalized mental patients, to observe behaviors which occur in the family and in the community as the mental patient begins to be seen more frequently outside of the hospital.

Because of the varying composition of the population under observation it becomes quite clear that one must understand the social-cultural milieu of the individual before the detection and the diagnosis of illness, which represents deviation and expectancy, can be made on some sound grounds.

Regarding etiology of the illness, a variety of scientific models have been provided for structuring the elements that go into possible causation. Among these are the social-cultural model, the developmental model, the learning and conditioning model, the genetic model, the enviromental model, the neurophysiological model. It would take too long to go into the details of each of these; but, for the most part, it is quite clear that for any illness each of these models plays a varying role, and that we need a supermodel, such as is afforded by the epidemiological model, for weighting the different models and their proper contribution to the causation of an illness.

While waiting for the doctor—that is, for science to provide keener tools for evaluating these models—we have to resort to the best methods we now have. Tools are being developed along the following three basic lines: culture-dependent methods, culture-
fair methods, and culture-free methods. A culture-dependent method tries to determine to what extent a person's behavior deviates from the framework of the social-cultural norm. The best method now available for such purposes is the interview.

Instead of the usual free-wheeling clinical interview we have introduced systematic structured interviews yielding total scores relating to overall psychopathology, with subscales related to the individual dimensions that contribute to the total score.

These interviews cover the general area of the mental status of the patient and his performance in the following areas of life: travel, eating, personal hygiene, living arrangements, drug and alcohol usage, sex, illegal acts, leisure time, interpersonal relations, and self-injury. The role that an individual occupies in society as a mate, breadwinner, housekeeper, parent, or student, are also covered. This is included in the psychiatric status schedule which is being developed in biometric research by Dr. Spitzer and his colleagues, including Mr. Fleiss, who is here tonight.

We are also developing a psychiatric history schedule, which is designed to enable an investigator systematically to record those aspects of the subject's history considered of importance in diagnosis and prognosis. Information is obtained about the development of the present illness, nature and number of past psychiatric episodes, academic and occupational performance, adolescent and adult social and sexual adjustment, and severity and temporal characteristics of a wide range of pathological signs and syndromes.

The culture-fair method attempts to deal with those behaviors which, though rooted in a particular culture, nevertheless can be translated from one culture to another. For example, such behavior as bereavement has equivalence in almost all cultures. Greeting behavior is another example.

The only culture-fair technique that we have time to describe at the moment is the technique for assessing communicability of speech. In this method the cloze technique is used for detecting whether a given person's speech is intelligible; for example, when every fifth word is eliminated is it still possible to understand what he is trying to say? The method involves the filling in of the gaps with the word thought to be originally present, and the proportion of correctly filled gaps is a measure of the intelligibility. This technique distinguishes between schizophrenics and normals and has prognostic value.

In the area of culture-free indicators, the responses that are elicited occur during the first 1,000 milliseconds following stimulation. They occur so quickly that we think culture doesn't have a chance to get its licks in.

The primary difficulty that one faces in this area is that of ensuring the motivation of the individual, since it is well known that mentally ill people are not as well motivated as normals. Hence a technique which shows the mental patient to be poor in performance may simply reflect his lack of motivation.

There are three ways of dealing with this problem. First, measuring motivation directly, which turns out to be very difficult, as you know; second, randomizing motivation; and, third, ensuring that motivation is on a high level. The last two methods have been attempted. And, as an example of the third method—ensuring a high level of motivation—we have used techniques such as dichotic auditory stimulation, in which the patient can distinguish sensory inputs which the normal fails to distinguish. You present two sounds to both ears simultaneously and ask the person to indicate whether the sound is located in the midplane. If one sound reaches the ear a little earlier, or is more intense, the apparent sound localization will move in the direc-
tion of the earlier ear, or of the more intense stimulus.

We can map around the head an area where discrepancies in time and in intensity are equivalent: you can't tell the difference, if you are a normal; but the patient can tell the difference.

Another technique which we use in this general area is that of pupillography, which we can't go into in greater detail. There are several other techniques of this type, and the further discovery that the evoked potential can serve to differentiate between patients and normals gives us a method which may be relatively free from cultural influence.

We come now to the last point: organization of the data resulting from such research. Here is perhaps the primary problem facing us in this conference.

An important thing to remember is that the organization of data must always follow some goal, as Dr. Kramer sagaciously insists. The goals which we would like to set for the analysis of the data are: first, the evaluation of the severity of psychopathology present, and second, its prognosis.

For the time being we are satisfied with using concurrent validation as a measure of severity. By having a trained clinician estimate the severity of the illness we can utilize his judgment as a first approximation in the iterative process of evaluating the scores we obtain for total psychopathology as well as on the subscales. By a process of successive approximations we can perhaps eventually obtain objective measures which may be better than the clinical evaluation, in the same way that Binet's tests were an improvement on teachers' judgment of intelligence, even though the latter served as the initial criterion for intelligence.

With regard to prognosis, which is essentially concerned with predictive validity, we plan to use followup of patients to see how well they eventually turn out, and, in this way, to develop methods for evaluating the culture-dependent, culture-fair, and the culture-free techniques.

Whether we need to turn to typological approaches, or be satisfied with dimensional approaches, will remain an open question until the data are collected and some of the statistical tests which are necessary for the detection of discontinuities in the data become available.

The presence of discontinuities is a very important matter. One often wonders whether the discontinuity inheres in the behavior of the patient or in the behavior of society. It may depend on the attitudes of the community, or on the examiner, rather than on the patient's behavior.

Furthermore, discontinuity may exist, not in a given variable but in interrelations between several variables. However, only the development of better methods or criteria for classification can determine what path to take in the future.

I have great faith—perhaps unfounded—that the master mathematician of the future can find transformations that will eliminate discontinuities, typologies, and what-not. But the fact that at the present state of our knowledge some typological analyses seem to be better than dimensional analyses, as several of the empirical papers have shown, is sufficient warrant for their usefulness.

Furthermore, disease entities, if they exist, are more "gemütlich" to typological than to dimensional approaches. In other words, the dimensional approach seems to go contrary to the kind of things that clinicians, in search for individual similarities, often tend to look for.

So, in summary, by providing more objective tools, better classification can ensue. Since the clinician is more concerned with individual similarities than with individual differences, typology may be a useful technique to use in assisting clinical diagnosis.