Chapter 4

Clinical, Phenomenological, and Biometric Assessment of Psychopathology
With Special Reference to Diagnosis

Joseph Zubin, Ph. D.*

The three preceding chapters have dealt with the problem of definition of mental disorders. I will touch but lightly on this question and will instead deal most directly with the second part of the topic of this symposium—measurement. Since biometrics is the science dealing with the measurement of living organisms, I will consider the biometric approach to detection, diagnosis, and classification of the mental disorders. Biometrics, however, is not the only approach to this problem. There are at least two more: the clinical and the phenomenological approaches. These must necessarily precede measurement, since they yield the initial observational data in psychopathology.

The clinical or bedside approach has as its chief goal the cure or alleviation of illness. Insofar as diagnosis helps in the process of therapeutic intervention, the recognition of the particular category (diagnosis) to which the disease belongs is an important aspect of clinical work. When diagnosis is of little or no help in therapeutic intervention, it tends to be denigrated. The most characteristic aspect of a good diagnosis is that it is basically a creative act, belonging to the realm of discovery, using Reichenbach's term, rather than to the realm of verification. As the clinician attends to the unfolding of the patient's illness, and as he elicits further information by direct or indirect probing, he formulates some hypothesis regarding the nature of the patient's difficulties. This hypothesis may be altered as he proceeds to infer certain consequences and goes on to test them against additional information. He continues in this manner until, like a judge in court, he arrives at a diagnosis, even though he sometimes feels the evidence is not all in. At the present time, there is no better way to integrate the data available on a patient than to permit the diagnostician to weigh the available evidence in his intuitive manner. Hopefully, computers may some day help or even take over, but their day has not yet arrived.

Recently, following the lead of cardiology and other specialties, William Smith of Cornell proposed the utilization of Bayes' Theorem in computerizing diagnosis. After obtaining from a group of 14 clinicians the probabilities with which each one of a series of 41 well-known symptoms of psy-

* Professor of psychology, Columbia University, and Chief, Biometrics Research, New York State Department of Mental Hygiene, New York, N.Y.
chopathology occur in each of 38 specific mental disorders, he determined first that the degree of agreement on the stereotypes for each disorder was rather high in these 14 clinicians; that is, 71 percent for the specific symptoms and 86 percent for the specific diagnosis. He then chose 30 patients to be seen by a clinician who evaluated the presence or absence as well as the intensity of the 48 symptoms. The clinician's diagnoses were compared with the probabilities given by the computer for these 30 cases when the conditional probability for each symptom was fed into the program. The results indicated that the disorder with the highest probability obtained from the computer agreed with the clinician's diagnosis in 87 percent of the cases. If the two top probabilities are taken together, the agreement reaches 97 percent.

The phenomenological approach is primarily aimed at the description of the symptoms and signs which the patient exhibits in the same way that a naturalist would describe the behavior of some species. The description is in terms of the subjective feelings which the patient expresses; that is, the symptoms, and in terms of the outward manifestations which the examiner can observe; that is, the signs. This naturalistic approach is relatively free of any systematic framework, though some phenomenologists prefer to describe their observations against the framework of psychoanalysis or existentialism, or against some other frame of reference that permits the description of the behavior to be related to certain theoretical foci. It should be borne in mind that phenomenology, strictly speaking, deals with the immediate experience of the observer; that is, with his private world. This is especially important for the phenomenological psychopathologist because he is trying to fathom the immediate experience of the patient by means of his own immediate experience. Thus, one set of private events (the patient's) is inferred from another set of private events (the observer's). The phenomenologist insists that raw private experience, prior to translation, is the beginning of all science. However, before it becomes acceptable as science, it must leave the world of private events and enter the public domain, where it can be subjected to external reliability and validity checks or at least to consensus. Insofar as this is possible, we can regard it as scientific data. Insofar as it not yet possible, we must still regard it as private, artistic, and prescientific. In the field of mental disorders, most of the data are still of the phenomenological variety, and to make them public, objective, reliable and valid is the province of biometrics.

The biometric approach takes as its point of departure the objective observation and measurement of the behavior in profiles or in some other system of description, and finally utilizes statistical techniques for determining the reliability and validity of the diagnostic label placed upon the patient as well as the degree of agreement between prognosis and outcome on follow-up. Because all objective data are grist for the biometrician's mill, he is in a better position to include measures which may reflect etiology and structural aspects, as well as the description of current symptoms.

It is quite clear that the three methods are actually different aspects of the search for knowledge regarding psychopathology. However, without the
clinician there would be no questions to answer, without the phenomenologist there would be nothing to measure, and without the biometrician there would be no solution to the questions that arise.

It is of interest to inquire why the biometric method has come into prominence today. The best answer one can give is that the Zeitgeist has demanded it. While diagnosis in psychopathology has come a long way from its pre-Kraepelinian days, it has not kept up with the needs of the times. Up until the 1930's, diagnosis was in the doldrums because it did not serve the clinician's main purpose; namely, choice of therapy. Before 1930 there was no need for choosing since there was only one major kind of therapy available; that is, custodial care. From 1930 to 1950, more therapies became available, but diagnosis was of little help in choosing. Whatever skill in diagnosis was available fell into disuse except for statistical purposes. Today, there are a tremendous number of therapies from which a clinician has to select for a given patient, and without proper assessment of the patient's behavior, there would be no way of evaluating the effectiveness of the therapy. Current diagnoses do not provide such evaluations of behavior. In fact, current diagnoses in the mental disorders may suffer the same fate as befell diagnostic work in the field of bacterial diseases. With the discovery of antibiotics, entire fields of practice, like ear-nose-and-throat, have practically disappeared. The only problem left is to determine whether the disease is viral or bacterial and, if bacterial, which antibiotic to apply. In TB also, the group of sanitarium specialists as well as the search for diagnostic subcategories has fallen by the wayside because of the introduction of streptomycin and isoniazid. Similarly, the typing of pneumococci has fallen into discard with the advent of antibiotics. In the mental disorders, with the coming of the drugs and other types of therapy, the diagnosis is also not nearly as important as the specific symptoms; for example, depression, which the patient presents in his acute state, and once the target symptoms are determined, suitable drugs or other techniques for their elimination are applied. The actual nature of the disease the patient is suffering from has become secondary. Unlike the contagious diseases, however, in whose case the antibiotics actually "cure" the illness, drugs, psychotherapy or somatotherapy do not claim to "cure" but merely contain or mitigate the illness. Further, we are on the threshold of a wide expansion in the kinds of individuals who will be coming for help to our clinics and various other facilities as a result of the community mental health movement and medicare. The question whether an individual coming for help is truly a "case" or not and what kind of a case he presents and what is to be done for him will tax our present diagnostic facilities and diagnostic methodology tremendously, since it is ill equipped to deal with the new flood of patients who will come.

Let me quote you from the case record of a famous patient: "* * * wearisome nights are appointed to me. When I lie down I say, When shall I arise, and the night be gone? And I am full of tossings to and fro unto the dawning of the day. My flesh is clothed with worms and clods of dust, my skin is broken, and become loathsome * * * When I say, My bed shall comfort me, my couch shall ease my complaint; Then thou searest me
with dreams, and terrifiest me through visions: So that my soul chooseth strangling and death rather than my life. I would not live always: let me alone; for my days are vanity."

This is the phenomenology of the disease as reflected in the words of the sufferer himself. The symptoms are ably described by him, and the signs of his illness are equally well observed by him. The clinician listening to this patient’s complaints would emerge with some kind of a clinical label, and the Book of Job has it, the illness is diagnosed in the following terms: "Satan * * * smote Job with sore boils from the sole of his foot unto his crown." What would be the biometrician's function in this case? It may appear mundane in contrast with the literary quality of the poet, but he would presumably try to measure the depth of the anguish by a variety of items which would scale the intensity of the feeling of depression, of suffering, of anxiety, obtain ratings from Job’s companions and his wife regarding the depth of his previous state of emotionality, and wind up with a variety of measures attempting to triangulate the various characteristics that the illness presents, including perhaps the number of boils on the surface of the skin and the extent of the attack. This might be repeated at monthly intervals to see how Job’s condition changes from month to month.

Disease is usually defined in terms of etiology, structure, and symptomatology. In most mental disorders etiology is unknown, the structure of the organs of the patient as far as we know is unaffected, and symptomatology is the only available basis for the definition. Mental diseases whose etiology and structural defects become known are usually lost to psychopathology. Thus, disorders, like general paresis, pellagra with psychosis, epilepsy, even PKU are now largely in the hands of other disciplines. Only diseases of unknown origin remain in the field of psychopathology. Furthermore, there is also the question of whether mental disorders are in fact diseases or merely reaction patterns. In order to bypass these philosophical dilemmas and still have a working definition, we shall define mental disorders (and at the same time any kind of a disorder, mental or physical) as follows: Mental disorders are of two kinds: diseases and defects. A disease is a progressive condition which, unless attended to, will result in premature death, extreme reduction of efficiency and happiness (severe inner distress) or both of these. A defect is a more or less stationary condition which leads to severe reduction of efficiency but not necessarily to premature death.

Someone has likened psychopathology to a landscape, the clinician to a taxi passenger, the phenomenologist to a tourist, and the biometrician to a surveyor. The clinician is not concerned with the landscape and its beauties as much as he is with arriving at his destination; that is, the diagnosis, by the shortest route. The phenomenologist takes no short cuts because he is most interested in viewing all the sights and in satisfying his curiosity about the nature of the terrain. The biometrician is less concerned with beauty than with a scientific survey of the entire landscape by means of his instruments. While each of these three disciplines can exist in its own right, it is clear that for experimental investigations and for the comparative evaluation of individuals as well as groups, the biometric method offers many advantages.
History and Accomplishments of Clinical Phenomenological Method

A historical survey of the clinical method in psychopathology is beyond our present purpose. As a result of such a survey (Zubin, 1965; Zubin and Fleiss, 1965), three dimensions emerged as underlying the current diagnostic structure: (1) Organic-functional, (2) hereditary-environmental, (3) multiple versus unitary causation, (4) disease versus reaction pattern, and (5) acute versus chronic. Each of these dimensions has made itself felt in our current diagnostic systems, and many of our successes as well as failures inhere in the clashing contrasts of these dimensions. Whether these dimensions represent continuities or dichotomies is still a moot point. The current interest in typology stems from the assumption of discontinuity in these dimensions or in their interrelationships. It has often been said that there is no more continuity between normality and schizophrenia than there is between pregnancy and the nonpregnant state. In other words, there is no such thing as slightly pregnant and, by analogy, no such thing as slightly schizophrenic. 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 that are necessary for the detection of discontinuities in the data become available. The possible presence of discontinuities is a very important matter, and one often wonders whether the discontinuity inheres in the behavior of the patient or in the behavior of the society. It may depend on the attitude of the community or the examiner rather than on the patient’s behavior. Furthermore, discontinuity may exist not in a given variable but in the interrelationships between several variables. It is clear that the source of these discontinuities, whether in the observer or in the patient, is crucial for comparing different cultural groups. However, only the development of better methods or criteria for classification can determine what path to take in the future.

That the conflict between the dimensional and the typological approach may rest on a pseudoproblem is suggested by an example from pomology. Distinguishing between "good" and "bad" apples by "apple knocking" or by some other type of global inspection, a basically typological problem, is now replaced by a dimensional approach in which the light absorption properties of the apple are measured for two monochromatic light waves, and the difference in the degree of absorption of one pair of wavelengths is calibrated for maturity and the differential in another pair of wavelengths for the presence of water core (Birth and Morris, 1963; Birth, 1965). With the help of ratings for edibility and appearance, the dial readings can be calibrated and the classification of apples into several categories can be automated. Here at least, the discontinuity lies in the taster rather than in the apple! I am not so disheartened by the suggestion that even if we find typologies, some simple transformation may eventually make them vanish. I have great faith, perhaps unfounded, that the master mathematician of the future can find transformations that will convert discontinuities, typologies and what not into continuous dimensions and vice versa. But the fact that at the present state of our knowledge some typological analyses seem to do better than dimensional analyses, as several empirical papers have
shown (American Psychiatric Association in collaboration with Psychopharmacology Service Center of the National Institute of Mental Health, 1965) is sufficient warrant for their usefulness. Furthermore, disease entities, if they exist, are more acceptable to the typological than to the dimensional approach. The clinician, unlike the psychometrician, is forever searching for individual similarities rather than individual differences. The usual question uppermost in his mind when facing a new case is first, whom does he resemble? and not, how does he differ? The search for differences usually comes later. Individual differences, however, attracted psychometrically minded men like James McKean Cattell, despite the fact that up to that point in history, his mentor, Wundt, was searching for the underlying similarities in the responses of his subjects and regarded individual differences as due to error.

Program of Biometric Evaluation

It has been pointed out that in order to understand something, we must first measure it. While there is some controversy regarding the universal acceptance of this dictum, it is one of the fundamental premises of the biometric approach, as Galton, the grandfather of the biometric method has pointed out:

"General impressions are never to be trusted. Unfortunately, when they are of long standing, they become fixed rules of life and assume a prescriptive right not to be questioned. Consequently, those who are not accustomed to original inquiry entertain a hatred and horror of statistics. They cannot endure the idea of submitting their sacred impressions to coldblooded verification. But it is the triumph of scientific men to rise superior to such superstitions, to desire tests by which the value of beliefs may be ascertained, and to feel sufficiently masters of themselves to discard contemptuously whatever may be found to be untrue" (Galton, 1879). But measurement that is worth while has to be based upon some underlying scientific model for encompassing the data which measurement would yield. While measurements obtained under the guidance of one model may sometimes be useful in the evaluation of hypotheses of other models, as was the case with the data collected by Copernicus, this is rather unusual, and for the most part the value of measurement is pegged to the model from which it emanates. For this reason, it is important at least to catalog the variety of models that are now available in the field of psychopathology and note what type of measurement might rise from each. A quick review of the field indicates that the following models are extant: (1) The social-cultural model, (2) the developmental model, (3) the learning theory model, (4) the hereditary model, (5) the internal environment model, and (6) the neurophysiological or brain model.

The social-cultural model is built on the assumption that all mankind is vulnerable to mental disorders, that given sufficient deprivation and stress-producing loads or other alterations in our environment, our behavior would be altered to the point where the ability to continue to live normally as independent individuals in society is endangered. Under this model certain
expectancies exist with regard to social-cultural norms, and an individual who deviates from these social-cultural norms is regarded as potentially mentally ill. The techniques that are necessary to determine the presence of such deviation are primarily culture-dependent techniques, and we shall turn to them when we discuss the tools that have been developed for this purpose.

The second model, the developmental model, and the third, the learning theory model, are built on the assumption that mental illness develops as a result of some specific deprivation or interference during the critical period of development or that the source of deviant behavior of the mental patient is to be sought in his reinforcement history. In contrast with the social-cultural model, which looks for deviations from expected social-cultural norms that may be indigenous to each cultural, the expectancies derived from the developmental model and from the learning theory model are not culture-bound in the same sense, and may be regarded as culture-fair. By this we mean that various expected behaviors from which the person deviates have transcultural translations or equivalents, and we can thus speak of greeting behavior and its deviations or of bereavement behavior and its deviations or of the response to reinforcement and its deviations. While the actual response to these forms of stimulation is not identical from culture to culture, it is possible to establish the correspondences across cultures. Hence we can call them culture-fair.

Finally, we have the genetic model, the internal environment model and the brain function model, which indicate that something is awry either genetically or internally or in the brain functioning of the mentally ill person. These deviations may be in a sense culture-independent insofar as an individual who is vulnerable to an illness may yield indications that he is deviant in regard to his functioning in one or more of these three areas regardless of his social-cultural milieu. These give rise to culture-free indicators which will be discussed later.

In attempting to develop a systematic approach to the assessment of patient behavior, a cataloging of the types of measures available and the methods for eliciting them experimentally was developed (Zubin, 1958). It is clear from this analysis that as one moves from the physiological to the conceptual level, more and more prior experience and learning becomes involved, and the memory storage of the brain is invoked. To the extent that the physiological level of response is not as dependent on memory storage derived from the cultural milieu as the conceptual, it is perhaps easier to find less culture-dependence in the physiological responses and in the next level, the sensory responses. In our search for culture-free indicators we have limited ourselves largely to the physiological and sensory levels. The culture-dependent techniques are to be sought largely on the conceptual level, or on the perceptual and psychomotor levels, if the encoded memories of past experience are involved. The culture-fair techniques cannot be brought into relationship with the other two types of techniques because the culture-fair technique involves a translation from one culture to another of the stimulus as well as the elicited response without altering the importance of deviations in the behavior in question with regard to psychopathology. The behavior
itself could range throughout all the levels from the physiological to the conceptual.

Another approach to describing the biometric program is to prepare instruments for the various members of the mental health team to obtain from them fullest information possible regarding their knowledge of the patient. Thus, if we prepare specific tools for the psychiatrist to record his observations during the interview, for the psychologist to record his measures and observations, for the nurse to record her observations of the patient's behavior on the ward, for the social worker to record the interaction of the patient with his family and with his environment, for the occupational therapist and the physical therapist to record such behavior as comes under their purview, we would, thus, capture important clinical information that now is not always recorded in objective fashion. The biometric program has provided each of these disciplines with specific tools for recording the behavior that they are most capable of observing and recording.

Results of the Biometric Approach

In presenting the results of our biometric approach to diagnosis, we shall adhere to the outline of a three-pronged attack, using culture-dependent, culture-fair and culture-free techniques. A more detailed examination of the tools we have prepared for measuring these indicators reveals that none of them is Simon pure and, therefore, a more refined definition is required.

In determining whether a given stimulus elicits a culturally deviant response from patients in different cultural settings, we must first make certain that the stimulus in question is actually identical in its significance and in its power to elicit responses in the different cultures. If a given stimulus invariably elicits from normals an identical response, regardless of culture, we can regard the response as a culture-free indicator. Deviations from the expected behavior could serve as an indicator of psychopathology. Thus, the response of neonates to the ferric chloride test for phenylalanine is a culture-free indicator of PKU. Similarly, deviation from the expected in the response of the pupil to an unpatterned light stimulus may prove to be a culture-free indicator of schizophrenia. This is most likely culture-free since conditioning the contraction of the pupil to light (as the unconditioned stimulus) has never proved reliable.

At the other end of the continuum are stimuli that elicit varying responses, depending upon the culture of the subject. Such stimuli and the responses they elicit must be examined against the norms of the culture in question before they can serve as an indication of the presence or absence of psychopathology. Even when the identical response is obtained to the identical stimulus, the inference regarding psychopathology will not always be the same. Thus, if a rural uneducated patient answers "yes" to the question: "Have you committed an unforgivable sin?" it does not bear the same significance as a "yes" answer from a college professor. Furthermore, deviations from the expected response to a given stimulus may also have different significance for psychopathology. Thus, in some cultures deviations in eating ritual may represent a greater breakdown of interpersonal relationships than devi-
ations in sex-behavior, while in other cultures, the reverse may be true. In such cases the behavior belongs in the culture-dependent category. The weight with regard to psychopathology attached to a given item in a given culture can vary from zero to a maximum value. In would be important to separate zero weights that indicate absence of a relationship to psychopathology from those which indicate that the given item is impossible in the culture under examination.

It might be well to point out at this juncture that the relevance of a given item to the presence of psychopathology may be independent of the value placed on the item in question by the local culture. Thus, seizures may be regarded as culturally acceptable by some groups, but this does not eliminate the relationship of this sign to epilepsy.

Finally, there are types of behaviors such as greeting behavior, bereave-ment behavior, and behavior in response to reinforcement or reward that are found in all cultures but that vary with regard to both the stimuli that evoke them and the responses that they elicit. Thus, extending the hand usually leads to a handshake in our culture, bowing may lead to a complementary bow in another culture, and among the Andaman Islanders the stimulus provided by the appearance or approach of a friend leads to profuse crying. In each of these cultures, the elicited behavior may have an equivalent value, and observed deviations from the expected may represent equal severity of illness, though the form of both the stimulus and the response is culture-dependent. These are the culture-fair indicators, whose stimuli and responses have to be translated into equivalent forms from culture to culture, but whose significance for detection of deviance remains constant across cultures. Thus, poor communicability of speech, though differing in form and content from culture to culture, nevertheless may prove to be a common indicator of mental disorder across cultures. In contrast with the culture-dependent and culture-free stimuli and responses that may represent the first level of raw observation of behavior, the culture-fair items represent a second level of analysis involving classification of stimuli as well as behavior into equivalent classes. All of these three types of indicators, the culture-free, the culture-dependent and the culture-fair, are at present necessary for the detection of psychopathology. Our goal may be to free the detection of psychopathology from the need to consider culture-bound indicators and in this way to make it possible to use identical stimuli for the detection of psychopathology transculturally. However, our ability to find such culture-free indicators is still only a hope. Meantime, we must depend largely on culture-dependent and culture-fair indicators to blaze the trail. It is also important to realize that in our present search we are not looking for causes. We are merely looking for indicants of psychopathology, regardless of whether they are causes or consequences or merely associated with psychopathology.

Culture-Dependent Techniques

The techniques to be discussed under this category are primarily culture-dependent, though they contain some items that overlap with the culture-fair and perhaps even with the culture-free category. As you can see, our tools were hammered out before we conceived of the culture related frame-
work, and it is now difficult to classify these techniques into nonoverlapping categories, though in our final analysis we shall attempt to make such a separation. We will tentatively place all of our interviewing and observational techniques in the category of culture-dependent techniques.

In considering the various culture-dependent tools that have been used historically for detecting and diagnosing psychopathology, the following developments may be considered. First, the identification of mental disorder, going back as far as primitive days, probably began with the face-to-face interview including observation as the basic tool. This is no place to trace the history of observational techniques and the interview and their development, but it is quite likely that throughout the entire period of man’s existence verbal communication or its absence served as one of the indications of various kinds of mental disorders. Up to World War I, the interview was the primary tool. Soon after World War I, it became quite clear that the interviewing techniques utilized for selection of military personnel as well as for industry were quite unreliable and invalid, especially when compared to the rather reliable and valid tests for measuring intelligence. As a result, the interview fell into disrepute in the eyes of many psychometricians, although it maintained its status in the eyes of clinicians. The psychometrician’s response to the failure of the interview was to invent the personality inventory and personal data forms, in which an individual gave self-reports of his behavior in a format which lent itself readily to conventional psychometric evaluations and interpretations. In this way, it was possible to avoid the kind of data which had to be interpreted through a clinical-theoretical scheme before they could be utilized successfully. Such instruments as the Bernreuter and eventually the MMPI rose out of this trend, and have now become quite useful in personnel work as well as in counseling. It is quite clear that these instruments, good as they are, can only serve the mild or borderline cases, since one of the problems of some of the severely disturbed mental patients is their inability to communicate their troubles directly. In fact, even normals sometimes give responses on such inventories that, unless followed up by rather careful interviewing, can be quite misleading. For example, Landis and Katz (1934) pointed out that the response to the question, “Do you feel uneasy in entering a room full of people?” separated the sample queried into three groups, one-third saying yes, one-third saying no, and one-third choosing the question mark as their answer. When questioned further it became quite clear that they all had felt essentially the same way about this item. All of these individuals felt a little uneasy at first but soon overcame this uneasiness. The first group referred to the initial period and checked “Yes.” The second group referred to the later period and therefore said “No,” while the third group, not knowing which period to judge, gave the doubtful answer. In mental patients, of course, this type of difficulty is magnified.

The trend now is to return to the interview but with fuller knowledge of the nature of verbal behavior and the nature of intercommunication between people. We can now develop techniques which will partake of the objectivity of the psychometric approach while preserving the richness of the clinical approach.

Our interviewing techniques differ basically from the usual clinical interviews insofar as they are systematic, structured, and cover the entire continuum of situations.
gamut of possible psychopathology. In contrast with freewheeling clinical interviews, which aim at a shortcut to diagnosis, the biometric interview tries to cover all of psychopathology systematically with each patient and in this way makes comparison of total scores and profiles feasible.

This discussion will be limited to only two of our interview techniques: The Structured Clinical Interview and The Mental Status Schedule.

Structured Clinical Interview (SCI)

The SCI was designed to serve as an individual test of social and psychological adjustment (Burdock and Hardesty, 1964a, 1966). It contains a fixed input of somewhat ambiguous open-ended questions that set a relatively mild tone of inquiry but that, nevertheless, provide the subject with an opportunity to express ideation and behavior from which the psychologist can judge the presence or absence of psychopathology. The SCI is intended for use both as a screening instrument (in the community as well as in admission services of hospitals and clinics) and as an assessment tool for determination of changes in psychopathology with the passage of time.

While the SCI provides measures of both level and pattern of psychopathology and, because of its brief and comparatively neutral impact on the subject, should be the first in a battery of tests, it may be desired to follow it with more probing instruments, like the Mental Status Schedule, to explore the source and depth of the psychopathology uncovered. However, psychopathology revealed in response to the neutral stimuli used in the SCI may be less influenced by the respondent's tendencies toward acquiescence or disavowal.

The SCI consists of an interview schedule together with an inventory of 179 items, which are incorporated as an integral part of the protocol to be marked true or not true by the examiner during the interview on the basis of the subject's answers. The interview usually takes about 20 minutes. Because the examiner records his judgments during the interview, no additional time is required for completion of the inventory after the subject has departed. The following is a typical section of the interview schedule:

(65) "Tell me something about your imagination."
"What kinds of things have you been thinking about?"

65. Reports that he engages in wishful thinking instead of working.
66. Reports that he broods over a certain unpleasant thought or feeling.
67. Reports or expresses weird or bizarre thought.
68. Reports that a certain irrelevant thought intrudes on his consciousness.
69. Reports that things seem unreal or dream-like.
70. Says that he feels as if he is outside of his body, or as his body does not belong to him.

In order to facilitate the examiner's task, the items of the inventory have been printed opposite those interview questions which experience with the technique has shown to be most likely to elicit the relevant behavior.
The SCI yields a total score which provides an overall measure of severity of symptoms. The items have also been clustered into 10 nonoverlapping subtests which describe areas of potential psychopathology in psychological terms: (1) Anger-Hostility, (2) Conceptual Dysfunction, (3) Fear and Worry, (4) Incongruous Behavior, (5) Incongruous Ideation, (6) Lethargy-Depression, (7) Perceptual Dysfunction, (8) Physical Complaints, (9) Self-Depreciation, and (10) Sexual Problems. The items were assigned to subtests on the basis of clinical and theoretical criteria; item assignments were verified by point biserial correlations which were cross-validated on a fresh sample. The reliability of the scores has been studied by comparing the results of two or more simultaneous but independent observers. Intraclass correlations for total scores extended from 0.83 to 0.92 for groups of patients ranging in size from 26 to 83. Interobserver reliabilities for the subtests ranged in the 0.70’s or 0.80’s with one exception, Perceptual Dysfunction, which has since been revised. With score on the Ward Behavior Inventory as a criterion for validity, correlations of 0.30 to 0.68 were found for the SCI, while for an unstructured interview the correlation with the same criterion was only 0.22.

Figure 1 illustrates certain systematic differences among normals, outpatients and inpatients in the proportions of items indicative of psychopathology. In general, the three groups are arranged in order of increasing pathology, from normals to outpatients to inpatients. However, while the
inpatients generally exceed the outpatients, the latter group scores higher than the inpatients on Fear-Worry, Physical Complaints, and Self-Depreciation. Since the outpatients are preponderantly neurotics, while the inpatients are mostly psychotics, it is not surprising that the psychotics show more Conceptual Dysfunction, Incongruous Behavior, Incongruous Ideation, Lethargy-Dejection, and Perceptual Dysfunction, while the neurotics are highest on Fear-Worry, Physical Complaints, and Self-Depreciation.

In order to facilitate study of individual and group differences, the raw scores, both total and subtests, were transformed to reduce skewness and then converted to standard form. Although total score represents the sum of the subtest scores, it was standardized independently as a measure of overall level of psychopathology. The 10 subtests in turn provide standard scores for pattern analysis.

Figure 2 illustrates how the technique can be used to compare individual subjects. The baseline represents the mean scores of a norm group of non-

![Clinical Profile Diagram]

**Fig. 2. Use of the structured clinical interview form to compare individual subjects.**
psychiatric subjects. Total scores are set off at the extreme left followed by the profile of subtests. The two patients represented in figure 2 were interviewed at time of admission to hospital. The overall level of pathology is about the same for both patients, or about two sigmas above the mean of the normals. However, their profiles show that the respective totals are compounded of different ingredients. The manic patient greatly exceeds the depressive in Incongruous Ideation and Sexual Problems, while the depressive shows extreme elevation on Physical Complaints, which are absent in the manic. The depressive shows considerably more Fear-Worry than the manic, although both are high. The two patients show about equal elevation on Anger-Hostility, Conceptual Dysfunction, Lethargy-Dejection, and Self-Depreciation. Neither patient shows any Incongruous Behavior or Perceptual Dysfunction.

Figure 3 presents successive profiles of a patient who was interviewed three times at intervals of 2 months. The first interview took place on
admission to hospital, the second interview at the request of the psychiatrist because of a marked change in her behavior. The third interview was held on a followup visit to the hospital after the patient's release. The three total scores show a progressive decline in overall level of psychopathology from a high of two sigmas above normal to less than one sigma above. The profile of subtest scores shows two features which remain relatively high on all three occasions: Anger-Hostility and Self-Depreciation. In her manic phase at time of admission the patient showed extreme Incongruous Ideation and had numerous Physical Complaints. She also manifested considerable Conceptual Dysfunction, Incongruous Behavior, and Fear-Worry. Two months later, in a depressed phase, Incongruous Ideation, Conceptual Dysfunction, and Incongruous Behavior together with Physical Complaints had all declined to normal, but Lethargy-Dejection had risen. By the time of the followup interview, Fear-Worry had returned to normal, but Conceptual Dysfunction had increased again to a significant extent and, together with Anger-Hostility and Self-Depreciation, reflected the persistence of deviation from the normal reference group.

Figure 4 illustrates the potentiality of the SCI for screening. Profiles of two normals are shown side by side. One of these, a relatively low normal, has some elevation on Anger-Hostility and Lethargy-Dejection. The other subject has a total score more than one-and-a-half standard deviations above the mean. His profile shows extremely elevated scores on Anger-Hostility,
Fear-Worry, Self-Depreciation, and Lethargy-Dejection together with suspiciously high amounts of Conceptual Dysfunction and Incongruous Behavior. In any screening program, the latter subject might well be referred for a detailed psychiatric evaluation.

In summary, the SCI has been designed to serve as an individual test for manifest psychopathology. It is a psychometric instrument that can provide reliable scores when it is administered and evaluated by a specially trained clinical psychologist. It may be used to compare patients with one another, to follow up changes in individual patients or to detect subjects with potential pathology in the community.

The Mental Status Schedule (MSS)

The Mental Status Schedule (Spitzer, Burdock, and Hardesty, 1964) 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 follow up of incomplete responses. Most of the questions 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 make 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.

The entire schedule is too large to include in this chapter. Instead, I have selected two typical portions. There are two major sections: the Interview Section and the Observation Section.

The following is a representative portion of the Mental Status Schedule:

| How are you getting along with people? | 27 complains about the way peers or strangers treat him. |
| What kinds of trouble do you have with people? | 28 complains extensively about the way people in positions of authority or power treat him (e.g., staff members, police, employer). |
| Whom do you feel you can trust most? | 29 complains extensively about members of family, friend, or associate. |
| | 30 indicates he cannot trust other people or that he is unduly suspicious of their intent. |

The left side shows the questions which the examiner asks and on the right are the pre-coded answers which the examiner marks True or False according to the response given by the patient.
The following is another representative portion of the Schedule, showing examples of the behavioral items which the examiner observed during the interview:

RATE OF SPEECH
211 Speaks extremely rapidly and with infrequent pauses.
212 Speaks extremely slowly.

EMOTION
213 Has a sad expression or holds his body in a dejected or despondent posture.
214 Talks of his problem with no outward sign of emotion.
215 Laughs inappropriately during discussion of a serious topic (do not include simple embarrassment).

QUANTITY OF SPEECH
207 Makes no answer to most or all questions.
208 Answers questions with single words or brief phrases only.
209 Talks on and on and keeps resisting interruption.

These are also marked True or False depending upon whether the patient exhibited the behavior in question.

An MSS scoring system has been developed based on a series of factor analyses of 2,000 MSS protocols. The 248 MSS items were first grouped into 145 clusters of from one to four items, based on the following criteria:

(1) All items which reflected the same type of behavior, even though they dealt with different aspects of the behavior, were grouped together in the same category.

(2) Items of reported behavior were not grouped with items of behavior observed directly by the examiner.

(3) Items which were associated with different diagnostic syndromes were kept apart.

Correlation matrices for the 145 clusters of MSS items were obtained on two randomly matched samples of 1,000 cases each, and factors were extracted by both varimax and biquartimin rotation of the first three, 10, and 20 principal components.

Two sets of factor scales were constructed. The first set consisted of three macro scales based on the first (largest) three factors rotated obliquely. Items which had their highest loading on the same factor in both samples of 1,000 were considered for inclusion in the scales if the loadings were 0.24 or above for both analyses.

The first factor is characterized by reports of worrying, depression, guilt, anxiety, feelings of inadequacy, social discomfort, impaired concentration, suicidal preoccupation, indecisiveness and other similarly dysphoric complaints. It was therefore named Feelings-Concerns.

The second factor is characterized by evidence of intense preoccupation, disorientation, impaired recent and remote memory, incoherence, and impoverished communication. This factor was named Confusion-Retardation, although many of its elements are subsumed under the concept of withdrawal.

The third factor is characterized by overt psychotic manifestations involving delusions of varying content, hallucinations of all types, and impaired insight. This factor was called Delusions-Hallucinations.

The correlations among these three scales approached zero (0.01–0.13). It is therefore consonant with the finding that an orthogonal (varimax)
rotation of the first three principal components yielded essentially the same factors. These three scales reflect the three major dimensions measured by the MSS and can be used when an evaluation of total severity is desired.

The second set of MMS scales consists of the following 15 scales:

1. Retardation—Emotional Withdrawal.
2. Agitation—Excitement.
3. Hostile Belligerence.
4. Inappropriate or Bizarre Behavior or Appearance.
5. Disordered Form of Speech.
6. Disorientation—Memory.
7. Depression—Anxiety.
8. Somatic Concerns.
10. Social Isolation.
11. Suspicion—Persecution or Hallucinations.
13. Sociopathic Impulses or Acts.
15. Insight Deficiency.

These scales were derived according to the following principles: (1) Factors appearing with great stability in all four rotations of the first principal components in the two matched groups of 1,000 patients each and in the two methods of rotation were made into the following four scales (Retardation—Emotional Withdrawal, Disorientation—Memory, Insight Deficiency, and Hostile Belligerence), (2) factors appearing in three of the four rotations with partial support in the fourth were made into the following three scales (Disordered Form of Speech, Somatic Concern, and Suicide—Self-Mutilation), (3) several scales were constructed which are more loosely based on evidence from the above factor analyses but emerged from clinical considerations and the results of the factor-analytic work of others (Agitation—Excitement, Social Isolation, Grandiosity, Suspicion—Persecution or Hallucinations, Sociopathic Impulses or Acts), (4) one scale (Impairment or Restriction of Daily Routine) was included, although it was not derived from the factor analysis, so as to permit comparability with the corresponding scale of another instrument, the Psychiatric Status Schedule, in which the dimension is greatly expanded.

The usefulness of this type of interview for cross-cultural studies is evidenced by a comparison of a sample of New York and of Kentucky patients (fig. 5). On three of the eight scales shown, the New York patients exhibited more pathology than the Kentucky patients. The probabilities were calculated by the Ridit method.

Evidence of the degree of agreement attained by different raters is shown in table 4. This table is based on the simultaneous observation of 15 psychiatric patients by three psychiatrists. It demonstrates the reliability of six of the MSS factor analytically derived scales in a situation where one psychiatrist interviews the patient and he and two observing psychiatrists make their judgment of the MSS items.
The ability of this instrument to detect changes in behavior under specific treatment (drug and milieu therapy) is shown in Table 5. This table demonstrates the usefulness of the MSS for studying specific items of behavior or target symptoms before and after different forms of therapy. Such item for item comparison may reveal types of behavior which respond to a specific form of therapy. It is doubtful whether global judgments of patient behavior would reveal these specific changes.

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 rated 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.
Table 4

Intraclass Correlation Coefficient Based on Three Psychiatrists' Simultaneous Observation of Fifteen Psychiatric Patients

<table>
<thead>
<tr>
<th>MSS scale</th>
<th>Reliability R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression-anxiety</td>
<td>0.99</td>
</tr>
<tr>
<td>Insight deficiency</td>
<td>.95</td>
</tr>
<tr>
<td>Disorientation-memory</td>
<td>.95</td>
</tr>
<tr>
<td>Hostile belligerence</td>
<td>.68</td>
</tr>
<tr>
<td>Retardation-emotional withdrawal</td>
<td>.73</td>
</tr>
<tr>
<td>Social isolation</td>
<td>.87</td>
</tr>
</tbody>
</table>

Table 5

Frequency of Selected MSS Items for "Successes" in Two Groups Before and After Treatment

<table>
<thead>
<tr>
<th>MSS item No.</th>
<th>MSS item abbreviation</th>
<th>Drug Pre-</th>
<th>Drug Post-</th>
<th>Milieu Pre-</th>
<th>Milieu Post-</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Feels restless</td>
<td>91.3</td>
<td>47.8</td>
<td>87.5</td>
<td>68.7</td>
</tr>
<tr>
<td>99</td>
<td>No interest</td>
<td>39.1</td>
<td>4.3</td>
<td>43.7</td>
<td>18.7</td>
</tr>
<tr>
<td>107</td>
<td>Ideas of self-injury</td>
<td>17.4</td>
<td>0</td>
<td>25.0</td>
<td>6.2</td>
</tr>
<tr>
<td>117</td>
<td>Hallucinates voices</td>
<td>34.8</td>
<td>12.0</td>
<td>37.5</td>
<td>18.7</td>
</tr>
</tbody>
</table>

The usefulness of these systematic structured interviews in determining the basis on which diagnoses are arrived at is demonstrated in figure 6, reporting an experiment conducted by Katz (Katz, et al., 1966). 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 introspunctiveness, perceptual distortion, motor disturbances, hostile belligerence, apathy and retardation, grandiose expansiveness and thinking disorganization (Lorr, 1953). Although the psychiatrists were all seasoned veterans of psychiatry, 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; namely, 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. We are planning an objective approach to the estimation of apathy independently of the interview.

Culture-Fair Techniques

In addition to direct interviewing with the structured instruments, certain focused 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, bereavement, 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; for example, mmm, 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) (fig. 7).
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, the rate of conditioning of schizophrenic patients; that is, 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 Portnoy, 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, which consists of systematically deleting words from continuous speech passages and having other subjects guess the words which have been so deleted. Schizophrenics and normal samples were compared, and we found that normals communicate better than schizophrenics because a larger number of correct words were guessed. Furthermore, differentiation between the patient and the normal was better for the second 100 words of the sample than for the first 100 words (Salzinger, Portnoy, 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 better differentiated 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, Portnoy, 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 to cross out words they feel do not belong. The results indicated that schizophrenic speech is characterized by having a larger number of intrusions; that is, words crossed out, than normal speech samples. The 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 had 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 in language he shows "loose association" and poor ability to communicate, because language communication requires that the speaker react to long-range stimuli and not simply to short-range associations.
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 culture-fair tests is likely to lie not in the function under measurement but in the subject’s approach to the testing situation; for example, 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, in those variables that 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-free techniques have been constructed to minimize.²

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 permitted to vary at random with respect to the experimental conditions. 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. The subject, therefore, made no discrimination; the response was identical regardless of the stimulus. 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 of stimulus sequence on reaction time. There were three types of sequences considered: (1) No shift, the stimulus being identical with the immediately preceding trend, (2) shift within the modality (red to green or vice versa, high to low tones or vice versa)—ipsimodal, and (3) shift between modalities—cross modal (Sutton and Zubin, 1965).

Results of one of these experiments are presented in figure 8. The Roman numerals on the abscissa refer to our three sequence categories, ipsimodal identical, ipsimodal nonidentical, and cross modal. Note that normals are faster on all accounts; this, of course, is the general finding. More important is the difference in slope. The influence of overall speed was eliminated by performing covariance analysis comparing groups on cross-modal stimuli with intramodality performance held constant. The patients are found to be disproportionately impaired in response to the condition of modality shift.

²Such 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.
In a more intensive replication, we considered separately the shift from light to sound and the shift from sound to light. We also divided the patients into process and reactive schizophrenics. In figure 9 it may be seen that essentially the same picture emerges in the reactions to sound stimuli as shown in figure 8. However, in the reaction to light the slope is the same for all groups. The covariance analysis shows a significantly greater cross-modal retardation for the patients on reactions to sound but not on reactions to light. The covariance analysis comparing process and reactive patients are 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 stimuli in the previous trials is a light. The females generally show the same trends, but usually their deviations from the normals do not attain statistical significance. These findings hold in both acute and chronic patients. This impairment is over and above the difference in overall level of response between groups. Comparable
analyses for reactions to light do not yield differences between patients and normals. These findings have also recently received independent confirmation in another laboratory by Kristofferson (unpublished manuscript) using a slightly different procedure.

In a comparable study done by Benton et al. (1962) on brain damaged subjects, 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 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 not 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 is one for which readiness has been achieved. When the stimuli do not fall in with these “expectancies,” there is a 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. Since all sequences in our programs were equally probable, best performance would occur if the subject assumed (correctly) that the next stimulus may be any one of the four used. Therefore, preparation for any one type of stimulus was maladaptive.

Our further studies with normals have made it apparent that in these experiments there are two effects at work rather than one. One of these effects may be called the probability effect, in which reaction time to the uncertain or less probable stimulus is lengthened, and the other may be called sensory effect. The sensory effect is revealed by the fact that manipulation of the probability of cross-modal and ipsi-modal sequences does not produce identical results. There is something special about cross-modal stimuli which makes them peculiarly vulnerable to the state of uncertainty 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 (Sutton et al., 1965).

A complete summary of our results in pupillography, and in certain other areas where cultural explanations of the observed differences are hard to find, would be too lengthy to include here. In a study we did several years ago, we studied the reactions of the dark adapted pupil to a cycle consisting of 1 second light and 3 seconds dark in a group of acute and chronic patients contrasted with normal controls. The results of this study showed that the pupil diameter at onset of light in each cycle was significantly larger in normal controls than in the acute patients, but there was no difference between the chronic patients and the normal controls (fig. 10). There was, however, considerable overlap between the groups. The height of the bar is the distance from the 25 to 75 centiles. The line shows the median, and the arrow shows the mean.

We have since then greatly improved the devices for measuring pupillary
changes and have developed a rather complex electronic measuring and analyzing system. We are at present in the process of collecting data from schizophrenic patients. In this study we are using a single 30 millisecond light flash as a stimulus which is presented after dark adaptation every 30 seconds. Ten to twenty of such reactions are then averaged in an average response computer.

Some preliminary data are shown in figure 11. Extent of contraction is shown on the abscissa, and the percentage of individuals obtaining a given degree of contraction or less; that is, cumulative, is shown on the ordinate. Note that for both male and female patients the pupil contracts less than for normals. Furthermore, with respect to the males, except for the one normal showing the least contraction, there is almost no overlap between patients and normals. The total overlap for both males and females is about 20 percent. These results are, therefore, highly significant statistically.

We are still somewhat suspicious that somewhere an artifact will be found to account for such astonishing results. One candidate for an artifact is the report that phenothiazines produce corneal opacity, and it is possible that this reduces light input to the retina. We are, therefore, undertaking measurement of corneal opacity to see whether there is any correlation with
our pupillary findings. We found also, as we did in our first study, a difference between patients and normals in the dark-adapted diameter of the pupil, but in both studies the overlap was extensive. However, there remains a larger difference between patients and normals in extent of contraction even when they are matched on initial diameter.

In a recent series of experiments, we explored the pupillary response to a 30 millisecond light flash flickering at different rates. When the dark interval between light flashes was between two-tenths and four-tenths of a second, we observed a phenomenon of delay between the light input and pupillary contraction of about 450 milliseconds, without a decrement in the transmitted information. Specifically, under these conditions the eye receives four flashes of light, and four contractions are obtained, but the last two contractions occur after all four light flashes have been presented. This suggests that there is a short term neutral storage mechanism, with highly efficient processing characteristics. Assuming that we might deal in schizophrenia with a nervous system of reduced information processing qualities, we tested a group of acute and chronic patients with this experimental procedure. The data indicated that most patients showed a deficiency in this “following” response. In a sample of 25 patients and 18 normals, only two normals had this deficiency while the recordings of only two patients were indistinguishable from those of the normals. Thus, out of the entire sample of 43 individuals, only four or 10 percent of the individuals were misclassified. Again we must enter a caveat that reduced light input brought about by retinal opacity could account for these differences.

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 could not attribute the result to poorer motivation on the part of the patients. We thought that the response to delayed auditory feedback (DAF) of speech might be such a function, since schizophrenics might be bothered less by the delay and, therefore, could perform better. Unfortunately, this is not the case; male schizophrenics have shown greater
interference than normals, while the female patients seem to purchase an unaffected reading speed, under DAF conditions, at the cost of more errors. We might also note that this test hardly qualifies for the culture-free label, since it seems to be sensitive to what are probably culturally determined sex differences between males and females.

Our more recent attempts to develop measures on which schizophrenic patients might excel seem to be meeting with greater successes, but these studies are still in the pilot stage. These data have been used in testing only half a dozen patients and an equal number of normals, but they show great promise. The Bunsen-Roscoe or Bloch's law seems to hold for normals when reaction time is used as the response to the varying supraliminal parameters of intensity and duration during a critical period of 10 to 20 milliseconds. In other words, reaction time is identical for these equal energy packages in normals, which means that normals cannot discriminate these stimuli. Schizophrenics apparently have not heard of this law and fail to obey it, or have a much shorter critical duration than normals, since for a few patients we have obtained orderly differences in reaction time, even for these equal energy packages.

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 displacement from center brought about by the two techniques. Preliminary results indicate that values cannot be found at which the schizophrenic will not discriminate the Δ I stimuli (intensity differences) from the Δ t stimuli (time differences). In other words, the schizophrenic is detecting differences that the normal cannot detect. 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 (1963) 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; for example, 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 baseline, 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 "Groups of Schizophrenias." The heterogeneity in the category of neurosis is too well known to require documentation.

Heretofore, 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 look for new methods applicable to heterogeneous populations. One approach is to 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.

Summary

The behavioral changes which accompany the development of mental disorder have been classified into three major divisions: culture-dependent, culture-fair, or culture-free. Present-day methods of detecting and diagnosing mental disorder are largely based on culturally deviant types of behavior so that there is a need for developing culture-fair and culture-free techniques. Even the culture-bound techniques, however, suffer unduly from variations in their application which produce diagnostic variability that could be reduced by proper controls. Methods for introducing such controls are provided through such techniques as standard interviews and inventories of behavior that can be objectively scored and evaluated. Experimental methods have been developed for studying verbal behavior itself, the carrier-wave of culture. These can be applied to differentiating diagnostic groups. Advances in the development of culture-free techniques are not as great. Yet, certain derivatives of reaction time studies, pupillography, evoked potentials to ulnar stimulation, etc., promise well to provide us with serviceable tools. When these objective methods are applied to populations of mentally ill, the new techniques for fractionating populations into homogeneous subgroups may yield new classifications of the mentally ill which will be more homogeneous and more suitable for future investigations.