FOREWORD

Joseph Zubin

"Can Psychopathology be measured," the topic of the symposium reported in this volume, is a rhetorical question. Those who regard psychopathology as a vicissitude of life, occurring as a result of a multitude of factors, none of which is crucial in itself, will deny the possibility of measurement and control. In fact, they regard measurement as an idea borrowed from physical science, inapplicable to human behavior, normal or abnormal. Those who believe in the essential lawfulness of human behavior, both normal and abnormal, will answer the question in the affirmative. Then there are the middle-of-the-road investigators who will admit measurement to some areas of psychopathology—the peripheral, external areas—but will deny it to other areas. The only reason for raising the question in the first place, is to focus attention on the many areas where measurement for better or for worse has already entered, and to call attention to the areas which still escape our measuring devices. Perhaps some areas of psychopathology will remain forever "private events," available only to the patient himself or to gifted therapists. This need not deter our attempt to measure, for as the Abbe Gialiari once said: Science is "plutôt destiné à étudier, qu'à connaître, à chercher qu'à trouver la vérité."*
A BIOMETRIC APPROACH TO PSYCHOPATHOLOGY*

Joseph Zubin

There are a variety of approaches that have been used in the past for detecting and evaluating psychopathology. Among these may be mentioned the psychiatric approach, the psychological approach, physiological approach, phenomenological approach, clinical approach, and many others. The biometric approach, which we represent, was hardly heard of in the halls of psychopathology before the end of World War II. Biometrics is a science which applies measurement to living organisms. The biometric approach to psychopathology is essentially an attempt at a scientific evaluation of the behavior of mental patients, including their characteristic thoughts and emotions, as well as overt behavior. The scientific evaluation of the behavior of mental patients now falls into the interstitial area between the social, medical, and physical sciences, and there is no one discipline in which the objective measurement of this behavior can be pursued regardless of territorial rights. Biometrics research is trying to provide a platform for the study of the behavior of mental patients with whatever measures become available. While no justification for introducing objective measures is necessary, one need only point to the three or four billion dollars spent annually for mental disease in this country and the two-thirds of the budget of New York State spent on mental hygiene to indicate the importance of measurement in this field.

Furthermore, a veritable revolution has taken place in psychopathology during the last decade. In order to keep up with these changes, to evaluate their direction, and guide their development, new measuring devices are required. The static measures of yesterday can hardly meet the requirements of today's dynamic upheavals in the care, treatment, and research of the mentally ill. Perhaps a comparison between the psychopathological revolutions and the more common political revolutions may prove enlightening.

In most political revolutions, the energy released accomplishes much good and some, perhaps, unnecessary evil. The status quo is challenged, law and order are at least temporarily overthrown, prison doors are unlocked and thrown open, newcomers previously belonging to the outcast group come to the fore, the national treasury is confiscated, and finally the revolutionists settle down to become conservatives and be eventually overthrown in turn.

The revolution in psychopathology has not been as drastic as

*This study was supported in part by Grant M-1541 from the National Institute of Mental Health.
political revolutions in certain countries are, but certain parallels between these two types of revolutions can be readily found. Unlike political revolutions, the psychopathological revolution occurred quietly, subtly, without public furor, but its effects are no less telling.

Law and order, as represented in our diagnostic classifications, and type of patient admitted and released, have undergone a considerable change. The diagnostic nomenclature has undergone a serious face lifting in the mental diseases and radical extirpation in the mental deficiencies. Whether these operations were a success is still debatable. The hospital doors have been thrown open, the prison-aspect of our institutions is gone, release rates have been doubled in the mental diseases, and admission rates have risen in both types of disorders. Clinical psychologists, sociologists, anthropologists, statisticians, biometricians, biochemists, geneticists, pharmacologists, behavioral scientists, and others, seldom involved in research in psychopathology during the thirties, are now recognized as important contributors. The separation of mind and brain—that a firm tenet of the thirties—is no longer regarded as a stumbling block in integrated attacks on mental disorders. Funds for research, rather scarce in the thirties, are now, if not ample, at least available from the national treasury.

It can readily be seen that no single discipline can hope to cope with the wide area of measurement called for. For this reason, biometrics research has corralled the interests of anthropology, sociology, psychiatry, social psychology, experimental psychology, statistics, psychopharmacology, neurology, internal medicine, and a variety of other disciplines that are needed in order to encompass the wide spectrum of measurement required.

The overall program of biometrics research involves two aspects, one of practical and the other of theoretical import: objectification of current clinical practice; provision of scientific models for etiology, and testing of their implications.

In connection with the first of these, it is generally recognized that there are many important clinical concepts, very useful in everyday clinical practice, which are usually evaluated subjectively and intuitively. One function of biometrics research is to provide objective measures of the factors underlying these clinical concepts and impressions. Examples of this would include our work on the development of an index of type of onset of illness, based on the study of premorbid adolescent friendship patterns, and, our investigation of the cloze procedure as a means of comparing the "intelligibility" of normal and schizophrenic speech.
Also, within this first category lies the problem of standardization of the observations upon which diagnosis, prognosis, and degree of improvement are based. Here we are developing inventories, ranging in content from mental status to ward behavior.

One of the problems to which these methods are now being applied is a sampling survey of hospital activities. The purpose of this study is to provide more adequate information about daily hospital routine, kinds of patients admitted, released, placed on treatment, moved within the hospital, etc., than has been available. To determine whether the tranquilizers distributed to the patients are actually ingested, the observational techniques will be supplemented by urine-analysis methods. Attitudes of the personnel toward the patients will be gauged by specially devised scales. It is hoped that these methods will provide the kind of hospital statistics which will be of most use. Burdock and Hardesty, whose paper is first in this monograph, will discuss observational techniques, especially with regard to children.

The second aspect of biometrics research has to do with the development and investigation of various scientific models for psychopathology. The models being considered by our group are: (1) neurophysiological, (2) learning theory, (3) developmental, (4) sociocultural, and (5) epidemiological. We have not yet dealt with the heredity model nor with internal environment as a model, except incidently.

Our research program has included several studies investigating the neurophysiological model. This model is represented by two of the papers to follow. First, Hakerem and Sutton will present their work in pupillography, and then Sutton and Roehrig will discuss their studies of delayed auditory feedback. These two studies are part of a general investigation of the physiological, sensory, perceptual, and psychomotor characteristics of schizophrenic patients, as compared with normals under specified conditions of stimulation.

The learning theory model has led Salzinger and his coworkers, Portnoy and Feldman, to investigate verbal behavior and its modification, mainly through operant conditioning procedures. Their contribution to this book will deal with the use of the "cloze procedure" to study the alleged impairment of verbal response chains in schizophrenic speech.

The developmental model is not represented by a paper in this volume, but, an example of the work in this area is our study of adolescent friendship patterns. We are investigating the types of friendship patterns which characterize adolescents who later develop a schizophrenic illness, and trying to determine whether process
schizophrenia can be separated from the reactive type by knowledge of these patterns.

Within the sociocultural model we have sought measures of social deviance. Hammer and Salzinger will present a paper discussing some statistical properties of schizophrenic language in an attempt to arrive at a measure of social deviance.

Also, within the framework of the sociocultural model, we have carried out two studies of the relation between social factors and institutionalization. Hammer conducted an investigation of the influence of a mental patient's interpersonal network of family and friends upon time of hospitalization. The study to be presented here by Walton and Bennett, assisted by Nahemow, will concern the role of prior isolation on degree of adjustment to a home for the aged.

The epidemiological model is a supermodel, in which each of the other models is weighted differentially in the importance of its contribution to a given disease. One of our main goals is the ultimate integration of the submodels into a comprehensive picture of the various mental disorders. This is, of course, contingent upon the future refinement of techniques and accumulation of findings under each of the submodels, covering the range described — from neurophysiology to culture.
PUPILLARY REACTIONS TO LIGHT IN SCHIZOPHRENIC PATIENTS AND NORMALS*

Gad Hakerem, Samuel Sutton, Joseph Zubin

In the literature on psychopathology one finds numerous references to abnormalities in the pupillary reactions of patients of all diagnostic categories. Bumke in 1904 and Bach in 1908 summarized the pupillary studies of mentally ill patients which had appeared up to that time. It appears from these two reviews that there was no consensus as to the type of pupillary abnormalities which were present in psychiatric patients. Often contradictory observations were reported for a particular diagnostic category. Both unusually small and large pupils were reported for schizophrenic patients. Bumke tried to relate specific pupillary abnormalities which were present at the beginning of the illness to the eventual outcome of the psychosis. He proposed that the absence of pupillary dilation to painful stimuli in the early stages of the illness would be indicative of poor prognosis. Unfortunately, there are no longitudinal studies to verify this statement. In 1907, Westphal described a pupillary abnormality he had observed in catatonic schizophrenics. In these patients the reactivity of the pupil to light stimuli varied from moment to moment - at one moment the reaction might be prompt and extensive, at another there might be absolute rigidity. He called this symptom the "catatonic pupil" or "spasmus mobilis." Levine and Schilder in 1942 reported similar findings. An extensive study of the pupillary reactions of schizophrenic patients to a number of sensory and conceptual stimuli was performed by Lowenstein and Westphal in 1933. The authors compared the pupillary reactions of these patients with the reactions of normals and found greater variability within the patient group. They were able to distinguish four types of pupillary contractions and dilations in response to the onset and cessation of a light stimulus. One of these is more frequently found in patients and consists of a sluggish contraction in response to light and a slow redilation after cessation of the stimulus.

The pupil is a muscular organ which is innervated by both divisions of the autonomic nervous system (sympathetic and parasympathetic), and its diameter at any moment in time is the resultant of the relative excitation of these two components. The contraction to light may be considered the parasympathetic component which acts in opposition to two sources of sympathetic control of pupillary di-

*Supported in part by USPHS Grants M-1541 and MY-4759 from the National Institute of Mental Health. The authors are indebted to Drs. Lowenstein and Loewenfeld for their invaluable suggestions for the conduct of this experiment.
lotion. One of these is the excitation of the dilator muscle via the cervical sympathetic chain and the second is the tonic inhibition of the contractor muscle via the Westphal Edinger nucleus. The tonic inhibition is under the control of the hypothalamus and descending reticular formation and is referred to as supranuclear or central sympathetic inhibition.

The observations of deviant reactions in psychiatric patients as well as the consideration of the nervous pathways controlling pupillary diameter indicate the usefulness of further study of the pupillary reaction in schizophrenic patients. Another reason for the use of pupillary reactions as an indicator of autonomic reactivity in psychotic patients is the fact that instruments have been available to obtain objective recordings of pupillary diameter. Lowenstein (1956) is mainly responsible for the development of methods which allow the recording of pupillary diameter in complete darkness and with a minimum of error.

The experimental group consisted of 37 acute and 41 chronic schizophrenic patients. The 22 normal controls were psychologists, graduate students, nurses and hospital employees. Table 1 gives the distribution of subjects by sex and age.

**Table 1**

<table>
<thead>
<tr>
<th></th>
<th>Normals</th>
<th>Acutes</th>
<th>Chronics</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Mean Age</td>
<td>N Mean Age</td>
<td>N Mean Age</td>
<td>N Mean Age</td>
</tr>
<tr>
<td>Male</td>
<td>13 25</td>
<td>16 27</td>
<td>27 29</td>
<td>56 28</td>
</tr>
<tr>
<td>Female</td>
<td>9 28</td>
<td>21 28</td>
<td>14 32</td>
<td>44 29</td>
</tr>
<tr>
<td>Total</td>
<td>22 26</td>
<td>37 28</td>
<td>41 30</td>
<td>100 28</td>
</tr>
</tbody>
</table>

Since the pupillary reaction changes with age (Kumnik, 1956), we limited our groups to the age range between 18 and 37. The acute patients were first admissions to Brooklyn State Hospital and had a history of recent onset of the psychosis. The chronic patients had been in the hospital for at least five years without interruption. In order to decrease the heterogeneity of the patient groups, only patients who manifested auditory hallucinations or who produced delusional material during an interview immediately before testing were selected. None of the patients had been given tranquilizing
drugs or somatic treatment in the two weeks preceding the experimental session, and none had any known neurological or ophtalmological disorder.

The recording apparatus consisted of a 35 mm. motion picture camera which was driven by a synchronous motor at a speed of 10 frames per second. A double-lens system (Wollensak f. 4.5 F1 101 mm.) was used to photograph each eye separately, thus permitting more effective use of the exposed film area. Kodak infrared sensitive film (421 H-IF) was used and condenser lamps with General Electric bulbs (6.5 V.) served as photographic light sources. A Kodak Wratten 88A filter which passes only wavelengths above 720 mµ was mounted in front of each lamp. The emitted light was visible to the subject as two dark red spots if looked at directly. The infrared light was directed on the eye from below at an angle of 45 degrees.

A stimulus light source was mounted on the camera and a collimated beam directed on the subject’s left or right eye. At the level of the cornea, the light patch was 10 mm. in diameter and 15 ft. candle in intensity. A coupled episcotister controlled the duration of the light and dark periods.

The subject was seated in a comfortable chair which could be adjusted for height. Fixed position of the head was maintained by a dentist-type head rest and a chin rest. A pinpoint beam of light was directed on the subject’s forehead on which a small black dot had been painted. By this device the position of the subject relative to the camera and to the stimulus was monitored and controlled.

The subject was dark-adapted for 10 minutes and then presented with a series of eight cycles of one second of light followed by three seconds of darkness. This is the sequence used by Lowenstein in his numerous experiments, and this procedure was adapted in order to permit comparison with his work. First the right eye was stimulated and then, after a new dark adaptation period, the left eye was stimulated. Both eyes were always photographed. Since we did not find an inequality of pupillary size (anisecoria) in any of our subjects, we considered only the measurement obtained from the right pupil. The film was processed and the pupillograms recorded by a method described by Lowenstein (1956).

Results

An inspection of the obtained pupillograms revealed large variability within each subject’s record and across subjects in all three groups. In Figure 1, parts of such records are shown for illustration. The three subjects whose records are shown were those indi-
Hakerem et al.: Pupillary Reactions to Light

Individuals whose average pupillary diameter between light flashes was at the median of their group. The first reaction after dark adaptation, the fourth reaction, and the seventh reaction are presented. As may

**FIGURE 1.** Sample reactions to light of one normal, one acute, and one chronic schizophrenic patients. The first, fourth, and seventh reactions of a series of eight reactions are presented. Light is present only during the first second of each cycle. Light is presented to the right eye and the measurements shown are for the right pupil.
be seen, the normal subject has a larger pupil and a slower rate of contraction than either patient. The patients show a slower recovery after cessation of light. No clear differences emerge between the three subjects in their reaction to the sequential light stimulation. Because of the large variation in pupillary reactions within and between subjects, an attempt to compare groups on each reaction did not appear to be a fruitful procedure. Therefore, several aspects of the pupillary curves were selected, and averages were obtained across all 16 reactions for each subject on these measures only.

The measures used were selected because of their theoretical importance as indicators of the balance between sympathetic and parasympathetic activity and reactivity.

The following indices were used:

- **Initial diameter:** the size of the pupil at the moment the light stimulus is presented.
- **Extent of contraction:** the difference between diameter at stimulus onset and point of maximal contraction.
- **Time of contraction:** the time it takes the pupil to reach its point of maximal contraction.
- **Peak speed:** the maximal decrease in pupillary diameter achieved in any tenth of a second during the contraction phase.
- **Flatness of curve:** the time the pupil remains at maximal contraction. This is obtained by measuring the length of time during which the pupillary diameter remains within 0.1mm. of its size at maximal contraction.

**Figure 2** summarizes the findings for three groups on initial diameter. The median of each group is indicated by the crossbar, and the arrow shows the position of the mean so that comparison provides an estimate of skewness. The height of the vertical bar gives the distance between the 25th and 75th centiles. It may be seen that the variability is largest for the acute patients. Each group was compared with every other group by Mann-Whitney U tests. Significance levels of the comparisons are shown in the figure. There is no difference between the normal controls and the chronic patients, while the acute patients have a smaller initial diameter than either the normals ($p < 0.002$) or the chronic patients ($< 0.05$). **Figure 3** shows the same measures for extent of contraction and **Figure 4** shows the measures for peak speed. No differences between groups are present. On peak speed, but not on extent of contraction, the acute patients again show the largest variability.
NORMAL VS. ACUTE: $P<0.002$
NORMAL VS. CHRONIC: $P=N S$
ACUTE VS. CHRONIC: $P<0.05$

FIGURE 2. Average initial diameter for 16 reactions. Initial diameter is defined as the size of the pupil at the moment the light stimulus is presented.

NORMAL VS. ACUTE: $P=N S$
NORMAL VS. CHRONIC: $P=N S$
ACUTE VS. CHRONIC: $P=N S$

FIGURE 3. Average extent of contraction for 16 reactions. Extent of contraction is defined as the difference between diameter at stimulus onset and point of maximal contraction.
FIGURE 4. Average peak speed for 16 reactions. Peak speed is defined as the maximal decrease in pupillary diameter achieved in any tenth of a second.

FIGURE 5 shows the data for the flatness measure. This is the length of time that the pupil remained within 0.1 mm of its smallest diameter before redilating in response to the cessation of light. Inspection of the figure indicates that normals seem to redilate earlier after having reached the point of maximal contraction. However, none of the group differences reach significance.

The range of the measurements for time of contraction, i.e., the time elapsed from onset of stimulus until the pupil reaches its smallest diameter, was rather small (0.8-1.3 sec.), and the unit of measurement was a tenth of a second. The data are, therefore, presented in the form of a bar graph (FIGURE 6). This graph shows the percentage of individuals in each of the three groups at any one of the six discreet measurements between 0.8 and 1.3 seconds. It is apparent from this figure that the normals take longer to reach maximal contraction. The Mann-Whitney test shows significant differences between normals and acutes ($p < 0.0005$) and between normals andchronics ($p < 0.001$), while the patient groups do not differ from each other.

In summary then, the acute patients have a smaller initial diameter and reach the point of maximum contraction earlier than the normals (time of contraction). The chronic patients have the same initial diameter as the normals and therefore they also show a significantly larger initial diameter than the acute patients. On the time of contraction measure, the chronic patients are similar to the acute patients and therefore show a more rapid contraction than the
FIGURE 5. Average flatness of curve for 16 reactions. Flatness of curve is defined as the time in seconds the pupil remains within 0.1mm of maximal contraction.

FIGURE 6. Distribution of the average time of contraction for 16 reactions. Time of contraction is the time it takes the pupil to reach its point of maximal contraction.
normals. There are no differences between groups on any of the other measures.

Since the extent of contraction does not discriminate between groups, while the time at which minimum diameter is reached is shorter in the patient groups, one would expect that there should be a difference between patients and normals on the peak-speed measure. However, no such difference was found. This seems to be due to the fact that the speed of contraction is biphasic. The initial speed of contraction seems to be the same in both groups but the speed decreases more slowly and the contraction ends later in the normals.

The average initial diameter, as used above, is a complicated measure, since for the first of the series of eight reactions the initial diameter represents the dark-adapted diameter while for subsequent reactions it reflects the ability of the pupil to redilate during the three-second dark period between stimuli. Therefore, each of the 16 reactions were analyzed separately. The results showed that the relative position of the three groups was the same whether one considered the dark-adapted initial diameter or whether one considered the initial diameter at the beginning of each light-dark cycle.

The average initial diameter would also mask any differences between groups in the degree of fatigue, i.e., in the ability of the pupil to recover from repetitive light stimulation. The data were, therefore, analyzed from this point of view. No differences between groups were found.

Correlations (rank order) were computed between initial diameter and each of the other measures. The values were as follows: with time of contraction 0.37 ($p<0.01$), with extent of contraction 0.27 ($p<0.01$), with flatness − 0.25 ($p<0.05$), and with peak speed 0.24 ($p<0.05$). Of course, the correlation between initial diameter and point of maximum contraction is very high, rho = 0.85 ($p<0.01$). Except for the point of maximum contraction the correlations are low and therefore, initial diameter only determines part of the variance on the other measures.

Discussion

The results of this study can be interpreted by viewing the differences between the groups on the basis of prevailing knowledge of the pupillary reflex arc. Lowenstein and Loewenfeld (1950) have related the dynamic and static behavior of the pupil to specific centers and pathways of the nervous system. Based on their scheme, it would appear that acute patients show a deficiency in the inhib-
itory control of the Westphal Edinger nucleus by higher centers. This is expressed both in the significantly smaller initial diameter, as well as in the shorter time of contraction for the acute patients. Since there are no differences in extent of contraction or in peak speed between groups, it must be assumed that a high contraction speed persisted for a longer time in the patients and was not reduced by inhibitory elements acting on the Westphal-Edinger nucleus. Another indicator of deficient inhibition of the Westphal-Edinger nucleus would be a higher score on the flatness measure, i.e., the pupil would be expected to remain at its contracted diameter for a longer period of time in patients than in normals. By inspection of the records, it was observed that this persistence at the contracted diameter occurred in many single reactions in a number of patients, but was rarely present in normals (see Figure 1). After the averaging, the statistical analysis showed no difference between the three groups. (In a comparison between normals and all patients the differences did, in fact, approach significance.)

It is interesting to note that in the static measure of initial diameter chronic patients behave like the normals, while in their reaction to light (time of contraction), i.e., in a dynamic measure, they behave like the acute patients. Apparently, it is necessary to load the pupillary system and measure its dynamic action to observe a difference in the pupillary reactions of chronic patients and normals.

Lowenstein and Westphal (1933) have described a pupillary reaction type which is characterized by small initial diameter, fast contraction, and a prolonged stay of the pupil at its smallest diameter. They called this type of pupillary reaction, tonohaptic, and reported finding it in several catatonic patients. In clinical neurology, this type of reaction has been found in patients with lesions or disorders of the basal ganglia. It can be experimentally produced in animals by cutting the fibers from higher centers descending to the Westphal-Edinger nucleus. Except for the inconclusive findings on flatness, the present data are consistent with Lowenstein’s description of the tonohaptic curve. In the present study, diagnosis of the schizophrenic sample on the basis of sub-categories such as catatonia was not available and it was, therefore, not possible to make a direct check on Lowenstein and Westphal’s observation.

Lowenstein and Westphal also reported that they found a high frequency of their Type-IV response in the patient group. This type of pupillary reaction consists of a slow and inextensive contraction and a slow dilation. In the present study, such a trend was not discernible. We tend, however, to confirm Lowenstein and Westphal’s
statement that there is greater intrasubject as well as intersubject variability in the patient group.

A series of control experiments are necessary before a hypothesis of reduced inhibition of the Westphal-Edinger nucleus in schizophrenic patients can be accepted. For example, although reasoning from animal experiments would suggest that the sympathetic deficiency found in patients is a deficiency of central rather than peripheral origin, direct experimentation with stimuli mediated by the peripheral sympathetic system would be desirable.

Rubin (1962) has recently used pupillary measurements to explore the hypothesis that a disturbance in adrenergic-cholinergic reactivity to stress may be present in the functional psychoses. He used the reaction pattern of his normal controls as his criterion and found that his patients showed excess, as well as deficiencies, in his contraction and dilation indicators. However, it is not clear to the present writers in what way Rubin's findings are different from the general observation of greater variability among patients.

**Summary**

The pupillary reactions to light were studied in 37 acute and 41 chronic schizophrenic patients and compared with those of 22 normal controls. Several aspects of the pupillary reaction curves were selected as indicators of the balance between sympathetic and parasympathetic activity and reactivity. The three groups were compared on these measures, and significant differences between groups in initial diameter and in the time the pupil took to reach maximal contraction were found. The data were interpreted on the basis of Lowenstein's analysis of the neural centers and pathways controlling the pupillary mechanism.

It appears that the patient groups show a deficiency in the inhibitory functions acting on the Westphal-Edinger nucleus. This deficiency seems to be more manifest in the acute patients who show a smaller initial diameter and more rapid time of contraction than the controls. The chronic patients do not differ from the normal controls in the static measure (initial diameter) but react like the acute patients on the dynamic measure (time of contraction). Further experiments are necessary to validate the observations made here and to elucidate the exact nature of the differences in the pupillary reactions of schizophrenic patients and normal controls.

The present data is in agreement with studies by other authors in showing greater intra- and inter-subject variability in the patient group.
References


DELAYED AUDITORY FEEDBACK OF SPEECH IN
SCHIZOPHRENICS AND NORMALS*

Samuel Sutton, William C. Roehrig, Jeffrey Kramer

Introduction

Since Lee (1958) first reported the disruptive effects on speech of introducing a temporal delay in a speaker’s air-conducted sidetone (often referred to as delayed auditory feedback), there has been increasing interest in this phenomenon from the standpoint of pure research as well as practical use. For example, it has been used in the detection of feigned deafness, and as a stressor in studies of personality differences, drug effects, etc. The main disruptive effects of delayed auditory feedback (DAF) are repetition of certain speech sounds (i.e., “stuttering” of certain consonants and prolongation of certain vowels), general slowing of speech with occasional blocking, and increased vocal intensity (Chase, et al., 1959).

A number of studies have indicated that effects of the delayed feedback cannot be viewed simply as a distraction, but rather that the delay in feedback per se is critical in producing the disturbance. Evidence for the criticalness of this time-locked aspect comes from a variety of experiments such as those which demonstrate the relative ineffectiveness for producing speech disturbance of extraneous stimuli, such as white noise, the voice of another speaker, or even a recording of the speaker’s own voice repeating the words on a previous occasion (Butler & Galloway, 1957). It has also been shown that similar disturbances in performance of a motor task other than speech can be obtained by delaying the auditory feedback from that task (Chase, et al., 1961a); or, in fact, by delaying feedback in any of several sensory channels (Chase, et al., 1961b).

It is now known that for maximally disruptive effects the delayed air-conducted sidetone should be amplified considerably so as to mask or override synchronous feedback via kinesthetic, proprioceptive, and bone-conduction channels. The exact value of the temporal delay has been shown to be not too critical in the sense that a wide range of delays will cause some disturbance of performance. However, there are large individual differences in the temporal delay necessary for maximally disruptive effects.

Several studies have indicated differences between groups in the effect of delayed auditory feedback on speech performance. One

*Supported in part under grants B-2699 (1686) from the National Institute of Neurological Diseases and Blindness and M-1541 from the National Institute of Mental Health.
of the more dramatic findings is the fact that one third of a group of stutterers tested (data as yet unpublished) actually improved, in other words, decreased their stuttering under DAF. We have also found in normal children that the amount of speech disturbance under DAF increases with age (Chase, et al., 1961c).

The large differences between individuals and between groups of individuals in the amount of speech disturbance due to delayed auditory feedback, and the differences in the delay times producing maximal speech disturbance can be attributed either to personality differences (which include the degree of ability to ignore auditory sidetone and concentrate on feedback via bone-conduction and kinesthetic and proprioceptive channels) or to some specific neurophysiological differences. The authors' preference is the latter. It seems unlikely that the stress of DAF could cause the speech of a third of the stutterers tested to improve dramatically merely because their personalities were that different from other stutterers or from normals. In conjunction with this, the study with children of different ages strongly implies physiological maturation. There are other known physiological changes in the brain during this age-span (e.g., EEG changes). The differences between children aged 4-6 years and those aged 7-9 years can probably not be explained by learning processes alone, and it is doubtful that they can be explained by personality differences. Spilka (1954) attempted to relate the changes under DAF in the vocal responses of 150 college males to various personality variables: (1) self-percept stability, (2) schizoid tendencies, (3) paranoid tendencies, and (4) rigidity. He reported that the only investigated voice variable whose change under DAF was significantly related to personality functioning was that of vocal intensity variation.

The present study is addressed to the question whether there are quantitative or qualitative differences between the performance of schizophrenics and normals under delayed auditory feedback. If one were to hypothesize along lines of personality differences, it might be expected that schizophrenics will be less affected by delayed auditory feedback since they are often said to be out of touch with reality, relatively uninterested in the external environment, and seem to be poorly monitoring their speech (and thoughts) from the standpoint of communication with others, using neologisms, word salads, etc. Goldfarb and Braunstein (1958) reported in their study that schizophrenic children were less affected by delayed auditory feedback than were normal children (ages 8-10½ years). However, no data are as yet available on the performance of schizophrenic adults under delayed auditory feedback.
Differences between schizophrenics and normals in the utilization of sensory feedback would also not be inconsistent with neurophysiological hypotheses with regard to altered brain function in schizophrenia. Any hypothesized malfunction at a basic neurophysiological level such as altered synaptic transmission time, reduced effectiveness of inhibitory synaptic transmission, altered "switching" time, or higher level of neural "noise," might be expected to reflect itself in altered utilization or time dependence of sensory feedback.

In either case, whether physiological or personality differences turn out to explain differences in disturbance under delayed auditory feedback, the present study is addressed to the prior question of whether any differences do in fact exist between the performances of adult schizophrenics and normals under delayed auditory feedback.

**Apparatus and Procedure**

Each subject read orally two passages; first, under the control condition of synchronous feedback, and then under the DAF condition with a temporal delay of 0.15 sec. His instructions were to read the passages out loud in his normal reading manner. The subject was not told about the feedback delay and if he stopped reading when he first encountered the delay condition, the experimenter indicated that he was to continue as best he could.

The first passage was 98 words in length and the second passage was 259 words in length. The subjects spoke into an Altec M 11 microphone which was attached to the headset holding the earphones so that the fixed distance between the subject's mouth and the microphone could not be altered by movement of the head. The output of the microphone was fed into an Ampex 601 tape recorder where the combination of tape speed (7½ ips) and separation of record and playback heads introduced a 0.15 sec. delay. The output of the playback head was fed through an EICO model HF 60 amplifier to a pair of Permoflux PDR 8 earphones with doughnut cushions. The amplification introduced between the microphone and earphones was such that the ratio of voltages, measured for a 400 cycle tone at these two points, was 78 db. This calibration was repeated daily.

The experimental group consisted of 11 male schizophrenic patients with a mean age of 30 years, and 17 female schizophrenic patients with a mean age of 32 years. All of the patients were tested shortly after admission to the hospital and were not yet under any drug or therapy during the week in which they were tested. The normal controls consisted of 13 male hospital personnel with a mean
Sutton et al.: Auditory Feedback

age of 34 years, and 17 female hospital personnel with a mean age of 28 years. The average educational levels for the four subgroups were between two and four years of high school.

The tape recording of each subject’s speech was replayed until the following measurements were obtained: for each passage, the reading time in seconds under each condition; and for Passage I only, the number of words spoken with no error, and the number of errors made in each of the following eight categories:

1. $a$ - Insertion of an extraneous sound
2. $+$ - Repetitive addition - stuttering
3. $p$ - Prolongation of a sound
4. $s$ - Substitution of an incorrect sound for the correct one
5. $o$ - Slighting, as in the dropping of the final $d$ in the word legend.
6. $\sim$ - Shifted juncture - when the end of one word is dropped and used as the beginning of the next word - “roun darch” for “round arch”
7. $\Leftrightarrow$ - Contraction - “he’s” for “he is”
8. $o$ - Omission of a word or part thereof.

Unless otherwise specified, nonparametric tests were used to analyze the data, the Mann-Whitney U Test was used to test significance (two-tail) of differences between groups and the correlations computed were Spearman rho’s (Seigel, 1956).

**Results**

*Time measure.* The time taken by each subject to read Passages I and II aloud under the Control (normal feedback) condition, and the DAF (0.15 sec. feedback delay) condition was measured, omitting pauses of longer than three seconds (these usually represented adjustments of earphones, etc.). Also omitted was the time taken by the subject to laugh or make impromptu remarks (e.g., “Oh, I made a mistake.”). Testing group differences, it was found that the male and female schizophrenics did not differ significantly under either the Control or the DAF conditions for either passage. The male and female normals did not differ under the control condition, but did differ significantly ($p < 0.002$) under the DAF condition. Therefore, in the analysis, the male and female schizophrenics are treated as one group and compared separately with male normals and with female normals.

In **Figure 1,** the medians for each group and condition have been indicated by crossbars. As suggested by McNemar (1949), instead of the quartile deviation, the 25th and 75th percentiles have been indicated by the top and bottom of each bar to show the extent
FIGURE 1. Oral reading time in seconds under the control condition (synchronous feedback), the DAF condition (0.15 sec. delay), and the time difference between the two conditions (DAF minus control). The mean of each distribution is indicated by an arrow, the median by a cross bar, and the 25th and 75th percentiles by the bottom and top of each bar, respectively.

For Passage I (FIGURE 1) the combined schizophrenic group read slightly more slowly under the control condition than either the normal males ($p < 0.001$) or the normal females ($p < 0.001$). Under the DAF condition, the reading times of the schizophrenic group was not significantly different from those of the normal females but was significantly slower than the normal males ($p < 0.003$).

As may be seen in TABLE 1, the correlation between scores under control and DAF conditions was only 0.39 ($p < 0.05$) for the
TABLE 1

CORRELATIONS (RHO) BETWEEN EDUCATION AND RESPONSE MEASURES AND CORRELATIONS AMONG RESPONSE MEASURES

<table>
<thead>
<tr>
<th></th>
<th>Passage I</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Education vs. Time control</td>
<td>Education vs. Time DAF</td>
<td>Education vs. CW control</td>
<td>Education vs. CW DAF</td>
</tr>
<tr>
<td>Normal males</td>
<td>-.33</td>
<td>.32</td>
<td>.27</td>
<td>-.25</td>
</tr>
<tr>
<td>Normal females</td>
<td>-.11</td>
<td>.20</td>
<td>.13</td>
<td>-</td>
</tr>
<tr>
<td>Schizophrenics</td>
<td>-.14</td>
<td>.01</td>
<td>.63†</td>
<td>.56†</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Passage I</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CW control vs. Time difference vs. CW DAF</td>
<td>Time control vs. Time DAF</td>
<td>Passage I</td>
<td>Passage II</td>
</tr>
<tr>
<td>Normal males</td>
<td>.72†</td>
<td>.36</td>
<td>.53*</td>
<td>.51*</td>
</tr>
<tr>
<td>Normal females</td>
<td>.19</td>
<td>-.08</td>
<td>.20</td>
<td>0</td>
</tr>
<tr>
<td>Schizophrenics</td>
<td>.74†</td>
<td>.22</td>
<td>.39*</td>
<td>.56†</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Time difference vs. Time DAF</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Passage I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal males</td>
<td>.72†</td>
<td>.85†</td>
<td>.61*</td>
<td></td>
</tr>
<tr>
<td>Normal females</td>
<td>.71†</td>
<td>.87†</td>
<td>.93†</td>
<td></td>
</tr>
<tr>
<td>Schizophrenics</td>
<td>.76†</td>
<td>.88†</td>
<td>.82†</td>
<td></td>
</tr>
</tbody>
</table>

*Significant at .05 level of confidence.
†Significant at .01 level of confidence.

Sutton et al.: Auditory Feedback

schizophrenic group; somewhat higher for the normal males, rho = 0.53 (p < 0.05); and lower for the normal females, rho = 0.20 (NS). Difference scores were computed for each subject representing his increase in reading time as a result of the DAF condition. The distributions of these difference scores for the three groups have also been indicated in FIGURE 1 (DAF minus control). The normal males were significantly different from the normal females, but neither group of normals was significantly different from the schizophrenic group. The difference between the normal males and the schizo-
phrenic group approached significance however and, as will be seen below, this difference becomes significant ($p < 0.04$) for the longer Passage II.

If, in the absence of a nonparametric equivalent, analysis of covariance is used (despite the lack of both normality and homoscedasticity) to test whether the groups would differ under the DAF condition when statistically matched under the control condition, a similar pattern emerges — the schizophrenics differing significantly from the normal males, but not differing significantly from the normal females.

The results for the same subjects on the longer Passage (II) showed an almost identical pattern as on Passage I (see Figure 1), but the significance levels were somewhat higher. The normal males performed similarly to the normal females under the control condition, and both groups spoke faster than the schizophrenics (both $p$ values are less than 0.0005). Under the DAF condition, the scores for the schizophrenic group were not significantly different from those of the normal females, but the schizophrenic group spoke significantly more slowly than the normal males ($p < 0.0005$). The normal females spoke significantly more slowly under DAF than did the normal males ($p < 0.0002$).

As for the increase in speaking time due to the DAF condition (DAF minus control), the schizophrenic group was more slowed down than the normal males ($p < 0.04$) but did not differ significantly from the normal females. (As with Passage I, a covariance analysis yielded identical results.) The normal females were more slowed down by delay than were the normal males ($p < 0.02$).

For Passage II, the correlations between the control and the DAF conditions on the time scores were 0.56 ($p < 0.01$) for the schizophrenics, 0.51 ($p < 0.05$) for the normal males, and 0 for the normal females.

In summary then, for both passages the effect of reading under delayed auditory feedback is to slow down the schizophrenic group more than the normal males but not more than the normal females.

Correct words. Passage I was scored by two independent judges for number and type of errors, and number of correct words spoken with no error. The correlation between the correct word scores obtained by the two judges was $\rho = 0.95$ for the control condition and $\rho = 0.93$ for the DAF condition; hence the scores from the two judges were averaged. As with the time measure, it was found that the correct word scores for the male and female schizophrenics did not differ significantly under either the control or the DAF conditions, while for the normals, males and females did not differ.
under the control condition but differed significantly \((p<0.02)\) under the DAF condition. Therefore, as with the time measure, comparisons with the schizophrenic group were made separately for the normal males and the normal females.

The correct word score distributions have been indicated in Figure 2 in the same manner as were the time score distributions in Figures 1a and b. The schizophrenic group did not differ from the normal males under either the control or DAF conditions. The schizophrenic group was slightly but significantly different from the normal females under the control condition \((p<0.05)\) and considerably different under the DAF condition \((p<0.001)\).

![PASSAGE I](image_url)

**Figure 2.** Number of words enunciated correctly (out of a possible 98) under the control condition (synchronous feedback), the DAF condition \((0.15\text{ sec. delay})\), and the difference between the two (control minus DAF). The mean of each distribution is indicated by an arrow, the median by a cross bar, and the 25th and 75th percentiles by the bottom and top of each bar, respectively.
With respect to the decrease in correct words due to delay (control minus DAF), the schizophrenics were more affected than the normal females (p<0.01) but were not more affected than the normal males. The normal females were significantly more affected than the normal males (p<0.05). As with the time measure, if analysis of covariance is used to test whether the groups would differ under the DAF condition when statistically matched under the control condition, the same pattern emerges — the schizophrenics differing significantly from the normal females but not differing significantly from the normal males.

**Figure 3** shows the percentage of subjects in each group who made a particular type of error, and **Figure 4** shows the average number of errors made by those subjects who made that type of error. The largest increases from the control to the DAF condition were in the percentages of subjects making the errors of prolongation (p) and shifted juncture (w). There were no remarkable differences between the schizophrenics and the normals with respect to either the percentage of subjects making any particular type of error or the average number of errors made by those subjects. The normal females made higher correct word scores under the DAF condition than either the normal males or the schizophrenics because of a somewhat smaller percentage of subjects making the errors of shifted juncture (w) and omission (o), and a smaller average number of errors per subject for the errors of omission (o), substitution (s), and repetitive addition, or stuttering (+).

In a number of studies, a correct word rate was obtained for each subject by dividing the number of correct words in the passage by the total time to read the passage. One rationale for this procedure was that under DAF some individuals would speak very slowly, waiting for the "echo" of each syllable before phonating the next, and thus obtain a high correct word score whereas others would try to maintain a normal rate of speed at the expense of the correct word score. By using a combination of the two scores, the correct word rate, it was thought possible to compare all individuals on a single measure. In this study, no evidence was found to support such a concept. In fact there was no significant correlation between the increase in time from the control to the DAF condition and the decrease in correct word scores between the two conditions (Table 1). Use of correct word rate score in this study would have completely obscured the most significant findings, because on this score the normal females did not differ significantly from the normal males under either condition, control or DAF, or from the schizophrenics under the DAF condition.
FIGURE 3. Percentage of subjects in each group committing each type of error under the control condition (synchronous feedback), the DAF condition (0.15 sec. delay), and the difference between the two conditions (DAF minus control).

FIGURE 4. Average number of errors, computed only for those subjects who made a particular type of error.
Age was found not to be correlated with either time or correct word scores under either condition. As may be seen in TABLE 1, there was no significant correlation between education and time score for any group, but there was a moderate correlation between education and correct word scores for the schizophrenic group only. It may also be noted in TABLE 1, that the correlations between the control and DAF conditions for both the time and correct word measures were much lower for the normal females than for the other two groups. There is no apparent explanation for this curious fact. The correlations between the time measures for Passages I and II were moderate to high for both the control and DAF conditions and were similar for the three groups. This would seem to indicate that Passage I is long enough for experiments of this type. However, it will be remembered that the group differences on the time measure were somewhat more significant for the longer Passage II.

Conclusions

The schizophrenic males and the schizophrenic females did not differ under either the control or the DAF conditions for both the time scores and the correct word scores; whereas the normal males and the normal females performed similarly under the control condition for both the time and the correct word scores, but differed considerably under the DAF condition, and in opposite directions for the time and correct word scores. Using Italian speakers, Black (1955) found that there was a sex difference among normals with females being more retarded by delay than males, but the present writers are not aware of any comparable study of sex differences for English or General American speakers.

The schizophrenics and the normal females were not significantly different with respect to the time scores under DAF, nor with respect to the time difference scores between the control and DAF conditions, whereas the normal males were less affected by delayed auditory feedback than either of the other groups with respect to both time measures. With correct word scores, the pattern was reversed, the normal males performing similarly to the schizophrenics for both correct word and correct word difference scores, and the normal females doing significantly better (i.e., less affected by DAF) than either of the other groups. The writers can offer no satisfactory explanation for this curious reversal, but feel some confidence in its existence, in view of the high significance levels obtained for the group differences, as well as the repeatability of the pattern of time scores of the groups on Passage II.

With respect to the original rationale for this experiment, it may
be definitely concluded that schizophrenics were not less affected by delayed auditory feedback than were normals, since they were more affected than the normal males (but not more than the normal females) on the time measure, and more affected than the normal females (but not more than the normal males) on the correct word measure.

**Summary**

The speech performance of schizophrenic patients was compared with that of normal individuals under the conditions of undelayed and of delayed (0.15 sec. delay) auditory feedback (DAF). Performance was compared on the basis of oral reading time, number of words enunciated correctly, and number and type of errors made. It was concluded that:

1. The speech of normal females is more slowed by DAF than is the speech of normal males.
2. The speech of normal females is less affected by DAF than is that of normal males with respect to the number of words enunciated correctly.
3. DAF produced equal slowing of speech for both male and female schizophrenics.
4. DAF produced an equal effect with respect to the number of words enunciated correctly for both male and female schizophrenics.
5. The speech of schizophrenics is more slowed by DAF than is the speech of normal males, but not more slowed than the speech of normal females.
6. With respect to the number of words enunciated correctly, the speech of schizophrenics is more affected by DAF than is the speech of normal females, but not more affected than the speech of normal males.
7. There were no remarkable differences in the number and type of errors between schizophrenics and normals.

**Acknowledgments**

The authors are indebted to Dr. Richard Allen Chase for his assistance with experimental design and procedure; to Mrs. Andrea Rosen and Miss Marion Fishman for assistance with analysis of the tape records; to Miss Joyce Kerr for assistance with the statistical analysis; to Dr. Nathan Beckenstein, Director of the Brooklyn State Hospital, for making available the patient population and testing facilities; and to Dr. Joseph Zubin for his helpful suggestions and encouragement of this work.
References


In the beginning was the Word, and it must have been positively reinforced, because man has been talking incessantly ever since. It comes as no surprise, therefore, that students of psychopathology spend so much of their time and energy in examining the language of abnormal people.

Schizophrenic speech in particular has given rise to many studies and theories. It has been described as "word salad," as too personal and as too impersonal, as noncommunicative, idiosyncratic, repetitive, dissociative and as overly generalized and, of course, overly concrete. In the face of this overflow of theories and contradiction of facts, we decided to study some of the conditions under which schizophrenic speech samples are collected and to evaluate at least some of the variables which might influence characteristics of schizophrenic speech.

We chose as our first subject of study the situation in which data about schizophrenic verbal behavior is most often collected, namely, the interview. In a different context (Salzinger, 1959a) we conceptualized what happens to the data after they have been collected in an interview in terms of a communication theory model. The interviewer-interviewee interaction constitutes the original message – the input – which passes through a series of filters and noise generators consisting of the interviewer's powers of observation and memory, his selective recording techniques, his methods of analysis of the data, and his summarizing measure; only after the information has been partly lost and partly distorted is there an output which consists of the clinician's judgment of diagnosis, prognosis, type of thought disorder, or what have you. Although this model gives rise to a series of ideas for research, as, for example, how much information is lost at each filter, it is a model about the

*This study was supported in part by Research Grant M4842, in part by Research Grant M4443(A), from the National Institutes of Health. We wish to thank Joseph Zubin for his help and interest in this research. We are especially indebted to Elsa E. Robinson and Constance R. Sutton of New York University for making available to us the subjects used to gather the cloze procedure data. The speech samples were originally collected in previous studies (Salzinger and Pisoni, 1960, 1961) supported by Research Grants M566(C3) and M1541 from the National Institutes of Health, and we are grateful to N. Beckenstein, Director of the Brooklyn State Hospital, to A. Glenn, L. Olinger and M. Portnoy for making available the hospitalized schizophrenic sample and to L. C. Kolb, R. Senescu and H. Schoenberg for making available the hospitalized normal sample. Thanks are also due to Roy O. Freedle, Robert Keisman and Phyllis Zlotogura for much of the analysis of the data of the present study.
unstructured interview in which the interviewer has no specific plan for his behavior but rather acts primarily on the basis of the immediate stimuli impinging upon him.

We therefore turned our attention to the experimental analysis of the influence of the interviewer's behavior upon that of the interviewee. Using descriptive behavior theory as a model, with particular reference to experimental studies in verbal behavior (Krasner, 1958; Salzinger, 1959b), we looked upon the interviewer as a source of reinforcement and discriminative stimuli and viewed the speech of the patient as containing responses that could be conditioned. In the course of these studies we found that the rate of emission of self-referred affect statements can be modified in schizophrenics by discriminative stimuli such as questions (Salzinger, 1956; described in Zubin, 1958) and by reinforcers such as verbal expressions of agreement, e.g., "mmhm, yeah, yes," (Salzinger, 1956, described in Zubin, 1958; Salzinger & Pisoni, 1958, 1960, 1961). The rate of emission of affect statements was also successfully modified in normal subjects, and a comparison of normal and schizophrenic subjects showed that the latter extinguished faster than the former (Salzinger & Pisoni, 1960). It was further shown that the process of conditioning verbal affect is independent of the temporal placement of the conditioning period, 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 & Pisoni, 1961). Most recently we devised a technique which makes interviewer questions unnecessary (Salzinger, et al., In Press). Under these conditions we obtained monologues — extended, uninterrupted samples of continuous speech — from schizophrenic patients, who were separated from the interviewer by a screen, and we were able to employ a more uniform reinforcer than in the interview studies cited above. Here we found that, depending on 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.

These studies have led us to a number of conclusions and inferences:

1. The interviewer can increase the amount of speech emitted by a patient and can thereby obtain more information without necessarily biasing the kind of information he will obtain.

2. The interviewer can increase the amount of speech in a particular area, that is, he can increase the frequency of occurrence of such a response class
as self-referred affect, and this in turn can affect his judgment of such factors as flatness of affect. In addition, the effectiveness of such reinforcing remarks as "mmhm" is very great in spite of the fact that they are considered neutral by some interviewers and applied without any attempt to control their occurrence systematically. (A very interesting research project, and one we have not yet had the courage to undertake, would be to try to control the questions and reinforcing behavior of the interviewer by manipulating the statements made by the interviewee). In other words, even though the interviewer reinforces the interviewee's behavior unknowingly, he may still do so in a lawful way. To the interviewer who considers sex problems to be the core of abnormality, the interviewee's statements about these will constitute reinforcing events for the interviewer, since they will tend to confirm his theory; for the interviewer who believes social causes to be crucial for abnormality, statements concerning these factors will act similarly. Each kind of interviewee statement would in turn be reinforced by the respective interviewer. In this way, the theory or bias of a given interviewer will determine what reinforces him, what he reinforces, and therefore what theory he will confirm.

In light of the results obtained in these interview studies and of some of the problems that were raised, it seemed desirable to investigate other aspects of verbal behavior in situations specifically designed for their study. In the course of these, we found that the constitution of a response class does not always follow common sense rules and that the response class boundaries must be determined by empirical means (Salzinger, et al., 1959). Affect statements conditioned in our interview studies were shown to constitute an integrated response class since the subclasses of positive and negative effect did not differ in terms of their conditionability (Portnoy & Salzinger, 1964). A technique was devised for reinforcing the speech of children, demonstrating the reinforcing power of a papier mache' clown for controlling different speech rates (Salzinger, S. et al., 1962). It was found that verbal and nonverbal responses may differ in the degree to which they form response classes in an operant conditioning situation, and that verbal responses were more easily conditioned than nonverbal responses (Salzinger, et al., 1964). This last finding suggests that verbal behavior may be the
more flexible response system and that there may be a close relationship between the relatively high level of development and the relatively greater conditionability of verbal behavior in human beings. In an unpublished experiment* carried out in our laboratory, operant conditioning techniques were applied to the vocalizations of a four-year-old boy who had never learned to say any words at all and who had been hospitalized initially for autism and later rediagnosed as mentally defective. In daily conditioning sessions over nearly nine months, we were able to increase his rate of vocalization in general and, by reinforcing successive approximations, were able to shape the articulation of at least a dozen words. Finally, in an experiment with dogs (Salzinger & Waller, 1962), the barking response was successfully conditioned and brought under discriminative control by operant procedures, and in addition a response chain was established consisting of 10 bar presses followed by 10 barks after which the reinforcement was delivered. Thus, it was possible to demonstrate in animals the operation of response-produced stimuli in chains not exclusively composed of motor responses.

In the course of our research it became clear that, despite the great similarity between the conditioning of vocalization in animals and the conditioning of vocal and verbal behavior in human beings, a difference of degree of contextual effects still existed. The emission of verbal responses is, of course, not a random process, and while it is probably true that there is always some degree of sequential determination in the emission of successive vocal responses in animals, there is not much evidence now available for response-produced stimulus control in such sequences of vocalizations. In the verbal behavior of human beings there are undoubtedly a great many contextual constraints or well-learned chains of responses. Rubinstein and Aborn (1960) recently reviewed many studies which are relevant here. In our own work (Salzinger, et al., 1962) we have studied the effect of different degrees of contextual constraint on the emission of verbal responses by using the Miller-Selfridge (1950) passages of different statistical approximations to English as the independent variable and Taylor's (1953, 1954, 1956) "cloze" procedure to obtain data on the dependent variable. The cloze procedure originated as a measure of readability and consists essentially of having subjects guess the words which have been systematically deleted from a given passage. Thus, each passage can be characterized as evoking a given number of correct words. A large number of correct guesses of the deleted words would be evoked by a pas-

*This study was conducted by K. Salzinger, Suzanne Salzinger, Judith Eckman, R. S. Feldman and S. J. Coen.
sage of high redundancy or strong response chains, where the probability of having one response following the next is high; on the other hand, a small number of correct guesses would be evoked by a passage of low redundancy or weak response chains, where the probability of having one response following the next is low. The results of our study (Salzinger, et al., 1962) indicated that at least two factors determine whether a person will emit a correct response, one of these being the syntactical structure (chains of grammatical response classes) and the other being the content of surrounding words (chains of content response classes).

Three studies have dealt with the cloze procedure as a measure of the state of the speaker. In a study done in our laboratory (Salzinger, et al., 1961), chlorpromazine was used to modify the state of the speaker experimentally. Taking 200-word passages from the speech emitted by a normal subject under different dosages of chlorpromazine (varying from placebo and no drug to 125mg. in 25mg. steps), we deleted every fifth word and had groups of normal subjects guess what words were missing. The results showed a decrease in number of correct guesses with an increase in dosage, indicating a progressive weakening of response chains. In a study done by Osgood and Walker (1959) the cloze procedure was applied to a comparative study of suicide notes and control letters. These investigators found that male suicide notes evoked a larger number of correct completions than male control letters but found no significant differences for the respective female communications. Fillenbaum and Jones (1962), also using the cloze procedure, found that mutilated aphasic speech samples were generally less accurately completed than were mutilated normal control samples.

Thus, the cloze procedure discriminates between the writings of such authors as Gertrude Stein and Erskine Caldwell; it distinguishes between different approximations to English; and, finally, it reflects different states of the speaker. In other words, this technique is sensitive to differences in the structure of language. Furthermore, it tests directly how well one person communicates with another. The English language is of course very redundant, that is, successive words contain overlapping information. Just to give an example, one can generally understand another person on a telephone even if there is occasional interference. This means that the listener on the telephone effectively supplies the words which he cannot hear.

Now let us return to some of the characterizations of schizophrenic speech, such as “word salad” or dissociative, overly concrete or overly generalized speech, etc. All these terms have in common what Whitehorn and Zipf (1943) call the quality of being
"egocentric," that is, the speaker exerts relatively little effort to make himself understood. It is interesting to note that their conclusion was based on an inference which is open to some question. Using a log rank-log frequency plot, they showed that schizophrenic speech contains more repetition than normal speech. This was later confirmed in a study by Johnson and his students (1944) and by a similar technique in our own laboratory (Hammer & Salzinger, In press). Yet the assertion that greater repetition means that the speaker is using the same word to signify different things and that he is therefore not communicating well, is an interesting but untested proposition. Although one might expect the repetition of a very few words over and over again to convey little information, even relatively few words can be arranged and rearranged so as to convey quite a variety of messages. On the other hand, it seems reasonable that even though an individual emits a large number of different words, the arrangement of these words may be such as to convey very little—witness the phenomenon of word salad. Thus, because the measures of diversity of language—although of interest for other reasons—did not really examine directly the phenomenon of ability to communicate, and because the direct methods of examination of this phenomenon are for the most part clinical and intuitive, occurring in the uncontrolled interview situation and therefore almost impossible to substantiate, we decided to turn to the cloze procedure in order to examine the question of differences between schizophrenic and normal subjects in ability to communicate.

Procedure

The speech samples used in this study were part of either extended samples of continuous speech obtained by the monologue procedure, which we devised to make interviewer questions unnecessary (Salzinger, et al., In press), or uninterrupted portions of our standardized interviews (Salzinger & Pisoni, 1960). Briefly, the monologue procedure consists of instructing the subject to begin speaking about the reason for his being hospitalized and to continue talking about any topic he wishes, such as his friends or his hobbies, until told to stop. The interview procedure begins in much the same way, except that general questions are asked whenever the subject is silent for more than two seconds and unlike the monologue procedure, where the patient and interviewer are separated from each other by a screen, the interview is carried out under face to face conditions.

The cloze procedure samples were taken from the beginnings of the monologues or interviews of 13 schizophrenic patients (11 acute
Salzinger et al.: Verbal Behavior

and 2 chronic) at the Brooklyn State Hospital and 12 nonpsychiatric patients at Presbyterian Hospital, both in New York City. These subjects were selected so as to provide 13 schizophrenic-normal pairs (one normal subject was paired with two different schizophrenics) matched for sex, age, and formal education, and, where possible, for ethnic group. Two of the pairs were female, the remaining 11, male. The ages of the schizophrenics range from 19 to 44 years, with a median of 33; the ages of the normals range from 18 to 46 years, with a median of 32. The differences in age between the members of the pairs of subjects range from 0 to 16 years, with a median of 3. Years of formal education for the schizophrenics range from 8 to 16, with a median of 12; the range for the normals is from 6 to 16 years with a median of 12. The differences in years of education range from 0 to 4, with a median of 0.

The fact that both groups were hospitalized at the time of testing is, of course, another important variable on which all subjects were matched. For the 11 acute schizophrenics the number of days of continuous hospitalization prior to testing ranged from 0 (i.e., the subject was tested on his admission day) to 8, with a median of 0; for their matched normal subjects the range was 0 to 19, with a median of 4. The two chronic schizophrenics had been hospitalized for 463 and 1,365 days and their normal matches for 38 and 3 days, respectively.

The following measures were obtained for speech samples in order to assess the comparability of the normal and schizophrenic material on other variables that could influence the results of the present study.

**Type-token ratio (TTR).** This is the ratio of the number of different words (types) to total number of words (tokens) in a sample of a given length (Johnson, et al., 1944). In this case, separate TTR's were computed for the first and second 100 words in each passage. Wilcoxon matched-pairs signed-ranks tests (Siegel, 1956, for this and all subsequent statistical tests) showed that there were no significant differences in TTR's within the schizophrenic group and within the normal group between the first and second 100 words, or across groups in comparing separately the first and second 100 words* (all p-levels $>> 0.05$). For the schizophrenics, TTR's in the first 100 words ranged from 0.53 to 0.72 with a median of 0.60, and TTR's in the second 100 words ranged from 0.54 to 0.72 with a

*As mentioned previously, one normal subject appeared twice in the sample. Since he was paired once with one of the two chronic schizophrenics, both of whom, along with their normal matches, were eliminated from all group statistical comparisons in order to maximize the homogeneity of the sample, his data are included only once in each such comparison.
median of 0.62. For the normals, TTR’s in the first 100 words ranged from 0.56 to 0.73 with a median of 0.61, and TTR’s in the second 100 words ranged from 0.53 to 0.73 with a median of 0.67. It is interesting to note that one chronic schizophrenic and his normal match have TTR’s almost identical to the medians for the remainder of their respective groups in both the first and second 100 word samples. The other chronic schizophrenic and his normal match have TTR’s which are extremely low with respect to the rest of their groups. This is evidence for both the importance and the effectiveness of the matching procedure employed here.

Grammatical analysis. Based on the Fries (1952) method of analysis, separate counts were made of the parts of speech (which we will call lexical words) and function words in the first and second 100 words in each speech sample. Lexical words correspond roughly to nouns, pronouns, verbs, adjectives and adverbs; function words correspond roughly to articles, quantity words, auxiliary verbs, conjunctions and prepositions.

Wilcoxon matched-pairs tests showed that there were no significant differences in number of lexical words (or in number of function words) within the schizophrenic group and within the normal group between the first and second 100 words, or across groups in comparing separately the first and second 100 words* (all p-levels $>0.05$). For the schizophrenics, the number of lexical words in the first 100 words ranges from 55 to 72 with a median of 59, and the number of lexical words in the second 100 words ranges from 53 to 66 with a median of 62. For the normals, the number of lexical words in the first 100 words ranges from 50 to 66 with a median of 58, and the number of lexical words in the second 100 words ranges from 54 to 68 with a median of 58. The analogous values for the function words may be obtained by subtracting the lexical counts from 100. It is interesting that the two chronic schizophrenics, in the second 100 words, show the only instances of fewer lexical than function words (45 lexical words in one case, 48 in the other), but this will have to be followed up by examination of additional samples of chronic schizophrenic speech.

The speech samples described above were prepared for the cloze procedure experiment in the following way: The first 204 words, which preceded any reinforcement periods or interviewer utterances except, of course, the initial instructions, were mimeographed on a separate page, without any punctuation and with every fifth word deleted (except numbers and proper nouns) and a blank substituted.

*See page 851 for footnote.
making 40 blanks per passage.* The passages were then distributed to groups of college students who were told that they were being given samples of continuous speech and that they were to fill in each blank with whatever single word they thought belonged there.

Each subject received four passages consisting of two schizophrenic-normal pairs. One hour was allowed for their completion. The order of the passages within each set of four was systematically varied among the subjects. Including only those subjects for whom English was the native language, 17 protocols were obtained for two pairs of passages, 16 for six pairs, and 15 for the remaining five pairs.

For each passage, the following measures were obtained, based on the words placed in the blanks.

**Proportion of correct guesses out of total guesses (C score).** A "correct" guess was one that matched the deleted word exactly, since Taylor (1953) found that allowing synonyms or alternative forms of the deleted words did not increase the ability of the cloze procedure to differentiate the comprehensibility of different samples of verbal behavior.

In doing statistical comparisons between C scores, the number of correct guesses was obtained for each individual protocol. All pairs of these individual C scores were used in testing for differences between a given schizophrenic speech sample and its paired normal sample.

**Proportion of different incorrect guesses out of total incorrect guesses (D score).** This is the proportion of all incorrectly guessed words which were different from each other. This score is a measure of the variety of responses elicited by the cloze procedure.

**Results and Discussion**

**Figure 1.** shows the cumulative frequency of schizophrenic and normal subjects as a function of C scores for the first and for the second 100 words of the passages. Note that both curves for the normal subjects show a larger proportion of correct guesses than either one of the curves for the schizophrenic subjects, and that the two groups are further separated from each other for the second than for the first 100 words. The C scores of the normal subjects tended to increase, and the C scores of the schizophrenic patients decreased from the first to the second 100 words of the passages. The first 100 words of each passage followed directly after the initial.
The two chronic schizophrenics and their normal matches are not included. S-1 and S-2, and N-1 and N-2, refer respectively to the first and second 100 words of the schizophrenic and normal passages.

The results shown here for the group data can also be demonstrated in the analysis of the separate matched pairs. Figure 2 presents the data for the proportion of correct guesses in the total passage. Note that only one matched pair in the entire sample shows a greater C score for the schizophrenic passage and this is one of the two female matched pairs. Since there are only two female matched pairs in the sample, it is not possible to assess the significance of this fact. All the other matches demonstrate that the normal subjects communicate more effectively than the schizophrenic subjects. These differences are significant (using Wilcoxon matched-pairs tests) at the 0.08 level for one match, beyond the 0.05 level for two matches and beyond the 0.01 level for the other nine matches.
SCHIZOPHRENIC - NORMAL MATCHED PAIRS

FIGURE 2. Proportion of correct guesses to total guesses (C score) in total passage for each schizophrenic-normal matched pair. The two female pairs are indicated by FEM., the two chronic schizophrenics are indicated by CH., and the normal subject used in two pairs is indicated by (*).

Inspection of FIGURE 3, which presents the proportion of correct responses for the first 100 words only, demonstrates the same point as the cumulative graphs of FIGURE 1, namely, that the dif-

SCHIZOPHRENIC - NORMAL MATCHED PAIRS

FIGURE 3. Proportion of correct guesses to total guesses (C score) in the first 100 words of the passages for each schizophrenic-normal matched pair. The two female pairs are indicated by FEM., the two chronic schizophrenics are indicated by CH., and the normal subject used in two pairs is indicated by (*).
ferences between normals and schizophrenics are not as large for
the first 100 words as for the second. Three of the matches show
the schizophrenic speech sample to be more comprehensible than the
normal one; of the remaining 10 matches, in which there were more
correct guesses for the normal than for the schizophrenic speech,
one is significant at the 0.07 level, six at the 0.001 level or beyond,
and three were not significant, the p-levels all being 0.10 or greater.

Inspection of Figure 4, which presents the proportion of cor-
rect responses for the second 100 words, shows only one match,
again, one of the female pairs, for which the schizophrenic exceeds
the normal speech in comprehensibility. Of the 12 other pairs, nine
are significant at the 0.001 level or beyond, and three are not sig-
nificant, the p-levels being greater than 0.16.

It is also of interest to follow in detail the changes in compre-
hensibility taking place from the first to the second 100 words. On
the 12 normal passages, seven increased in comprehensibility – three
with a statistical significance level beyond 0.002, one equal to
0.02, one equal to 0.04 and two greater than 0.08; five normal pas-
sages decreased in comprehensibility – four with a statistical sig-
ificance level equal to 0.02, and one equal to 0.26. (The normal
passage which was used twice gave the same result on both occa-
sions, but is included only once in the above figures.)
Of the 13 schizophrenic speech samples, 11 decreased in comprehensibility from the first to the second 100 words - seven with a statistical significance level of 0.02 or beyond, two equal to 0.05, one equal to 0.09, and one equal to 0.64; the two schizophrenic samples which increased in comprehensibility did not do so significantly, having \( p \)-levels of 0.10 and 0.16. It is clear that this analysis, like the group data analysis, shows only a slight tendency for the normals to improve in comprehensibility, while the schizophrenics markedly decreased in comprehensibility as they had to depend more and more on the stimuli which they presented to themselves, rather than on the instructions or introductory questions which the experimenter presented to them.

The \( D \) scores show that there is relatively little differentiation between the schizophrenic and the normal group, or within each of the groups between the first and the second 100 words. Thus it would seem that the mutilated material exerts unequal contextual constraints upon the words actually emitted by the subjects but exerts very similar contextual constraints on the number of different incorrect responses.

Finally, an attempt was made to evaluate the relationship between the cloze procedure data and the Type-Token Ratio, the measure of repetitiveness which has been described by some to be an index of communicability. Earlier theoretical discussion of this point showed that this need not be true. We subjected this to empirical test by correlating \( TTR \) scores to cloze procedure \( C \) scores, separately for the first and second 100 words of both the schizophrenic and normal speech passages. None of these rank order correlations reached a magnitude which is statistically significant at the 0.05 level. Thus, it becomes clear that an index of repetitiveness does not in fact measure the same aspect of language as does the cloze procedure.

What, then, have we learned from this experiment about schizophrenic language? We have demonstrated the validity of the hypothesis of relatively lower communicability of schizophrenic speech - an hypothesis of which most previous data have been merely suggestive. Furthermore, it has been possible to specify differences between speech samples obtained at particular points in time under particular conditions. It will be interesting to extend the present investigation to speech samples selected over a greater time span in order to see whether, for example, the schizophrenic speech would continue to decrease in comprehensibility and at what point the normal speech might begin to show such a decrease, and also to study such differences under a greater variety of experimental
conditions. Normals and schizophrenics do not differ from each other under all circumstances, and it is important that those of us who are interested in psychopathology cease simply to canvass the schizophrenic's behavioral repertory for differences, but that we study the exact circumstances under which these differences occur, those under which they do not occur, and what in these different circumstances might explain the dissimilar results. We have suggested in our cloze procedure study that the reason the normal-schizophrenic differences were not as marked for the first as for the second hundred words, was that the effect of the external stimuli, namely, the introductory remarks of the experimenter, tended to diminish over time as the schizophrenic's own response-produced stimuli take over. Putting this into the more general language of external versus response-produced stimuli, one can review the literature on schizophrenia to see what other studies can be interpreted in this way. This is not the time to cover the area in detail, but we can mention the fact that the schizophrenic patient who "hears" voices is at that time primarily controlled by his response-produced stimuli. Also, two studies done in our own laboratory are relevant here. One (Salzinger & Pisoni, 1960) we have already mentioned, namely, that in which the schizophrenic patients were found to extinguish faster than the normals. The interpretation is the same as for the cloze procedure study: when the external controlling stimulus (in this case the verbal reinforcement) is left out, the schizophrenics very rapidly cease to be influenced by that external stimulus, while the normals continue to show its effect. In another study done at our laboratory (Salzinger, 1957), both schizophrenics and normals acted as subjects for a weight judgment experiment. They were asked to judge the heaviness of five weights under normal conditions and under conditions in which they were required to lift a much heavier anchor weight preceding each judged weight; they were also told to correct their judgments for the anchoring effect. In this study, too, the schizophrenics were similar to the normals only as long as the controlling stimuli were external. However, when the subjects' performance had to depend upon an instruction not immediately present, which is of course just another example of response-produced stimulus control, the schizophrenics again showed poorer performance than the normals.

What about the future of the cloze procedure for the study of psychopathology? One of the important arguments in its favor is, of course, the fact that the technique differentiates speech samples not only of chronic schizophrenic patients but also—and this is probably more important—of schizophrenics who do not have the
many obvious symptoms by which almost anybody can recognize a chronic patient. We will most likely investigate the usefulness of the $C$ score as a prognostic and diagnostic index, since it provides a great deal of information on the basis of very short speech samples. The only factor which might keep us from taking advantage of the practical possibilities of the technique is our interest in identifying at least some of the basic variables determining the cloze scores. A number of studies in this area are already in progress, and, with the help of linguistic techniques, we hope to be able to make more exact statements concerning the role of such multiword variables as structure versus the role of such single-word variables as frequency of occurrence in the English language.

References

of schizophrenics during the clinical interview. J. Abnorm. Soc. Psychol. 57: 84-90.


SOME FORMAL CHARACTERISTICS OF SCHIZOPHRENIC SPEECH AS A MEASURE OF SOCIAL DEVIANCE*

Muriel Hammer and Kurt Salzinger

In recent years, much discussion and research have been devoted to the question of the relationship between language and other aspects of culture, between language and "thought," between verbal and nonverbal behavior. The question to be considered here is whether certain statistical properties of speech are affected with any regularity in schizophrenia, an illness clinically characterized as involving disturbances of thought processes, and an illness which often manifests itself in social deviance of many kinds. Previous studies have contrasted certain formal characteristics of language usage between schizophrenic and normal subjects (e.g., Fairbanks, 1944; Feldstein & Jaffe, 1962; Gottschalk, et al., 1957; Johnson, 1944; Mann, 1944; Mowrer, 1953; White, 1949). In this study, we will examine some of the statistical aspects of speech in terms of their potential usefulness as an objective, quantitative, and broadly applicable index of social deviance.

Deviance from one's own group is not a measure of degree of "illness," but it is related to the ability to function in one's social environment and may, therefore, underlie criteria of "onset" and "outcome" of illness. An index of social deviance may not give any information about the patient's psychological or psychopathological condition; it may, nevertheless, be useful in terms of social prognosis, which accounts for some unknown portion of the factors influencing entry into a mental hospital, length of stay in the hospital, and "adjustment" during the post-hospital period. In addition, it may be useful to assess individuals in these terms in studies dealing comparatively with different social environments. For example, some of our own work has suggested that temporal patterns of hospitalization vary with sociocultural characteristics like the structure of the patient's ties within his family (Hammer, 1963-64) and the relative normative disparity between the patient's subculture and the dominant cultural system (Hammer, 1960). It would have been useful in that research to have had a measure of each patient's "distance" from his own group, in order to evaluate the degree to which systematic variations among the patients may have affected the sociocultural comparisons. An index of social deviance, applicable to individuals, might help to resolve this kind of problem.

*This study was supported in part by Grants M-4082-A and M-4842 from the National Institute of Mental Health.
The linguistic measures described are seen as providing possible initial indices for an area of deviance which does not suffer from the problems of potentially unreliable reporting that one encounters in trying to deal with many other areas of social behavior. For most social behavior, judgments of degree of deviance tend to be both impressionistic and implicitly based on the norms of a particular social group which may often be different from that of the persons being judged. Moreover, the related concepts of "norm" and "deviance" have a variety of meanings in current usage (e.g., Nadel, 1958), and "degree" of deviance differs according to the definition used. What is said to be proper behavior need not be the same as what is said to be expected behavior, and each of these may differ from what is observed to occur most frequently. The main advantages offered by speech as the behavior under investigation include the following: (1) it can be studied directly, as contrasted with the study of much of social behavior, which is, in varying degrees, dependent on reports of behavior not directly observed by the investigator; (2) many of the norms involved in speaking are not explicitly known to the speaker, and cannot be altered "at will;" (3) relatively large samples of speech can be obtained in shorter observation periods than are required for most social behavior; and (4) technical facilities exist for the recording of relevant aspects of the behavior. In addition, although deviance will be used here to refer only to a frequency-of-occurrence definition of norms, meaningful, precise operational definitions of the several other senses of "norm" or "deviance" can probably be worked out more readily for speech than for other kinds of social behavior.

That the speech of at least some schizophrenic patients is disturbed, is clearly observable, but systematic deviations from the speech of the patient's own normal language community have not been demonstrated. In studying schizophrenic patients' speech, some investigators have worked mainly with interpretation of manifest content (Kasanin, 1944), and have failed to define the language community explicitly. As will be indicated later, such definition seems important. Moreover, it may be more profitable to concentrate on those aspects of speech which do not readily enter the "awareness" of the speaker (e.g., the type-token ratio), and are probably less subject to his voluntary control than is the content of his speech. Considering the diverse functions language serves (e.g., "self-expression," "giving information," "being friendly," etc.) and the different requirements involved in each, one is led to expect differences not only in lexical content but also in form of organization of speech — between, for example, a scientific analysis.
and a love song, and perhaps also between the speech of schizophrenic patients and that of normal subjects.

If schizophrenia involves withdrawal, autism, and other similar characteristics often ascribed to the illness, an activity as socially dependent as speech should be affected. To the extent that the schizophrenic patient has withdrawn from social contact, his speech is no longer fully exposed to the ongoing subtle adjustments that may be presumed normally to continue throughout life, based on responses to others' speech and to a variety of nonverbal cues present in ordinary social interaction. Two kinds of results of such limitation of social interaction may be suggested: an increase in the relative use of those aspects of speech which have great habit strength in the premorbid stage, and, at the same time, an increase in idiosyncratic usage. The high habit-strength responses remain because of the large number of reinforcements they originally received, and the idiosyncratic responses come to the fore since the premorbid social restrictions (reinforcement contingencies controlled by a listener) are not currently effective.

Behavior of schizophrenic patients should show signs of being relatively more affected by their own immediately preceding behavior than by the general subject matter relevant to the external situation (Salzinger, Portnoy & Feldman, 1964). Their speech might, therefore, be expected to be more stereotyped and idiosyncratic than the speech of normal subjects of the same linguistic background. Comparison of schizophrenic and normal subjects on several measures concerned with somewhat different aspects of speech usage may yield a quantitative index of deviance which could be related to such variables as outcome of hospitalization.

The work reported here is preliminary research, involving the application of various word-count measures to the speech of a small number of subjects. No attempt has yet been made to relate these to any nonverbal index of deviance (except to the extent that this may be involved in the comparisons between schizophrenic and normal subjects), or to formulate composite speech measures. The primary intent of this paper is to describe a methodology in the analysis of normal and psychotic populations selected with attention to some of the background variables which may be relevant to defining a language subcommunity.

Procedure

In this initial work, a small sample of subjects was selected from groups of schizophrenic and nonpsychiatric patients from whom recorded monologues had been collected in our laboratory for another study (Salzinger, et al., In press).
It was considered desirable to use monologues rather than speech passages involving verbal interaction in order to eliminate, for the present time, the problems of evaluating the interviewer's (or other participant's) effects on the subjects' variations in utterance-length and content. Possible differences in the structure of speech of these monologues as compared with other kinds of interviews, and with a variety of more "natural" situations, will require further investigation.

The monologues from which the speech samples were taken had been elicited by asking the subjects to speak at length about any topics of interest to them, such as their illness, their families, their work, etc. (Salzinger, et al., In press). The monologues varied in length with a maximum time of 30 minutes. The experimental conditions were altered during the 30-minute period, and the maximum size sample that could be used for all subjects under the same condition (i.e., before reinforcement was delivered) was 900 words of the subject's speech, after the experimenter's explanations and preliminary questions, formed the speech sample for each subject.

The two groups of subjects consisted of schizophrenic patients at the Brooklyn State Hospital, and nonpsychiatric patients at Presbyterian Hospital, both in New York City. All subjects in the sample were born in the United States. Each of four schizophrenic subjects was matched with a normal subject in sex and ethnic background and as closely as possible in education and age. A supplementary sample of eight subjects was also analyzed although it was not composed of such matched pairs. The supplementary sample corresponds to the original four pairs in terms of sex and ethnic background, but not in terms of age and education. Furthermore, this supplementary sample does not consist of pairs of schizophrenic and normal subjects, and merely irregularly expands the size of the original four "groups" (Table 1).

The criteria of selection were intended to yield a gross approximation to speech subcommunities. In addition to speech variation associated with different regional and class dialects, there are occupational variants, family-group variants, age-group variants, and so on. Criteria adequate for defining actual speech subcommunities would necessitate field study not warranted for this initial investigation of the measures. On the other hand, the measures cannot be developed at all without some definition of the speech community. The requirement of birth in the United States simply limits the sample to native speakers of "American" English. (There could be some exceptions to this, but they did not appear to occur in the sample used.) Regional background was in part controlled, since
Table 1
Background Characteristics of the Schizophrenic and Normal Subjects

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Sex</th>
<th>Age</th>
<th>Education</th>
<th>Ethnic Group</th>
<th>Sex</th>
<th>Age</th>
<th>Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negro</td>
<td>male</td>
<td>27</td>
<td>10 yrs.</td>
<td>Negro</td>
<td>male</td>
<td>30</td>
<td>9 yrs.</td>
</tr>
<tr>
<td>Negro</td>
<td>female</td>
<td>40</td>
<td>10 yrs.</td>
<td>Negro</td>
<td>female</td>
<td>32</td>
<td>10 yrs.</td>
</tr>
<tr>
<td>Italian</td>
<td>female</td>
<td>44</td>
<td>8 yrs.</td>
<td>Italian</td>
<td>female</td>
<td>40</td>
<td>6 yrs.</td>
</tr>
<tr>
<td>Jewish</td>
<td>male</td>
<td>24</td>
<td>12 yrs.</td>
<td>Jewish</td>
<td>male</td>
<td>23</td>
<td>14 yrs.</td>
</tr>
</tbody>
</table>

Matched pairs

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Sex</th>
<th>Age</th>
<th>Education</th>
<th>Ethnic Group</th>
<th>Sex</th>
<th>Age</th>
<th>Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negro</td>
<td>male</td>
<td>36</td>
<td>12 yrs.</td>
<td>Negro</td>
<td>male</td>
<td>32</td>
<td>12 yrs.</td>
</tr>
<tr>
<td>Jewish</td>
<td>male</td>
<td>30</td>
<td>12 yrs.</td>
<td>Jewish</td>
<td>male</td>
<td>46</td>
<td>12 yrs.</td>
</tr>
<tr>
<td>Jewish</td>
<td>male</td>
<td>19</td>
<td>12 yrs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Unmatched subjects

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Sex</th>
<th>Age</th>
<th>Education</th>
<th>Ethnic Group</th>
<th>Sex</th>
<th>Age</th>
<th>Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jewish</td>
<td>male</td>
<td>25</td>
<td>12 yrs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

one group of subjects was Southern-born, and all subjects were New York City residents at the time they entered the hospitals. Ethnic background was the main available criterion of dialect, since it was not possible in this sample to obtain sufficient information for selection on the basis of region and class. The defining characteristics of ethnic group were dictated in part by the data available from hospital records. The three groups used were Negro, Jewish, and Italian, selected because they were considered to be relatively distinctive groupings and their populations in New York are sufficiently large to allow for the selection of matched subjects from a limited initial population. The Negro group included only Southern-born subjects (Florida, North Carolina, South Carolina, Texas, and Virginia). The Jewish group was selected to include subjects of East European background only — primarily Russia and Poland, which had a concentration of Yiddish-speaking Jews. For two of the six Jewish subjects, however, the judgment that their family background was East European had to be based on partial information only, such as the subject’s occasional use of Yiddish words; for the other four Jewish subjects, the records indicate Russian or Polish birthplace for at least one parent. Of the four subjects of
Italian background three had at least one parent born in Italy, while
the parents of the fourth subject were born in the United States.
No attempt will be made here to relate any differences found on
the speech measures to any other aspects of these ethnic groups.
It is ethnic difference per se rather than the specific group which
is of interest here.

Certain arbitrary decisions were tentatively made concerning
criteria for a "word," which was the basic unit of analysis. Mul-
tiple-word names whose parts are not used as names independently
of each other were considered to be single words (e.g., "New York"
was counted as one word, but "Mary Jones" was counted as two
since both "Mary" and "Jones" might be used separately to refer
to the same or another person). Related forms like "walk" and
"walks" were counted as different words. The effect of this de-
cision will require further exploration, since, a priori, the differ-
ence between "walk" and "walks" does not seem comparable to
the difference between "walk" and "perhaps." Furthermore, be-
cause the analysis was performed without regard to grammatical
category, certain words which should probably be distinguished are
treated as the same word (e.g., "walk" used as a noun and "walk"
used as a verb). In future analyses, this problem will be approached
by utilizing limited contextual identification. Questions about the
number and nature of relevant grammatical distinctions, size of
context, and operations for classification will need to be resolved.

Results and Discussion

Type-token ratio. The first measure considered, the type-token
ratio (TTR), is the proportion of different words (types) out of the
total number of words (tokens) in a passage. The type-token ratio
has been used previously in comparing mental patients with normal
subjects on both written and spoken passages (Fairbanks, 1944;
Feldstein & Jaffe, 1962; Mann, 1944; review by Mowrer, 1953), and
normal subjects have generally have been reported to have higher
average type-token ratios than schizophrenic patients. These studies
raise several questions, two of which have been investigated in
the present study: (1) Generally, the measure used as a score has
been the average TTR over a number of short speech segments —
typically, segments of 100 tokens. Such an average need not be
correlated with the overall TTR's, i.e., speech samples whose
average TTR over a number of 100-word segments is relatively high
need not have a high TTR for the total speech sample: at the hy-
pothetical extremes, an average TTR of 0.70 could be produced
by the use of the same 70 words repeated in each 100-word segment,
or by 70 different new words in each 100-word segment yielding overall \( TTR \)'s for 200 words of 0.35 or 0.70 (70 or 140 types), for 300 words, of 0.23 or 0.70 (70 or 210 types), and so on. On the assumption that sustained variation in speech may be the more sensitive measure, overall \( TTR \)'s as well as average \( TTR \)'s using 100-word segments have been computed and compared; (2) No special attention has been paid in previous studies to the possible influence of dialect differences on this measure. Subjects in this study have been matched for ethnic background—a partial control of dialect difference—as well as age, sex, and education.

The first hypothesis considered was that schizophrenic subjects tend to have lower type-token ratios than normal subjects of the same sex, age, ethnic, and educational background, and that this tendency is stronger for relatively large speech samples than for small ones. It was also hypothesized that failure to control for ethnic background, and possibly also for education, sex and age, would distort the results of a comparison between schizophrenic and normal subjects in a direction dependent upon the relative distribution of these factors within the schizophrenic and normal samples.

The four matched pairs of the first sample of subjects show a uniformly higher overall type-token ratio for 900 words for the normal than for the matched schizophrenic subject although the difference is small in the first and third pairs (Table 2).

The differences are not as consistent when considering the \( TTR \)'s for the first 100 words of the speech samples or the mean \( TTR \)'s for all the 100-word segments. As stated earlier, there is no necessary relationship between the mean \( TTR \)'s and the overall \( TTR \)'s. Using the combined sample of 16 subjects, the correlation* between the two measures is not very high either for the whole sample \( (r_s = 0.34) \) or for schizophrenics \( (r_s = 0.30) \), or normals \( (r_s = 0.21) \) separately. Mann (1944), however, used both 100-word segments and overall \( TTR \)'s on 2,800-word written passages, and found a high correlation between the mean \( TTR \)'s and the overall \( TTR \)'s. Chotlos (1944) also found high correlations among \( TTR \)'s of written passages for segments varying from 100 words to 3,000 words. He does, however, say, "...there are instances in which the value of the type-token ratio for 100-word segments places individuals near the top of the group for this measure while for these same individuals the value of the 3,000-word type-token ratio places them near the bottom of the group." Among the differences between

*The correlations shown here and elsewhere are intended only as descriptive statistics for this sample.
### Table 2
Comparison of Matched Schizophrenic and Normal Subjects on Type-Token Ratios for 900 Words and for 100 Words

<table>
<thead>
<tr>
<th></th>
<th>Overall TTR 900 words</th>
<th>TTR first 100 words</th>
<th>Mean TTR 100-word segments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Schizophrenic</td>
<td>Sch./nor.*</td>
</tr>
<tr>
<td>First pair</td>
<td>.332</td>
<td>.329</td>
<td>.99</td>
</tr>
<tr>
<td>Second pair</td>
<td>.349</td>
<td>.322</td>
<td>.92</td>
</tr>
<tr>
<td>Third pair</td>
<td>.294</td>
<td>.288</td>
<td>.98</td>
</tr>
<tr>
<td>Fourth pair</td>
<td>.360</td>
<td>.343</td>
<td>.95</td>
</tr>
</tbody>
</table>

*The ratio of the schizophrenic to the normal TTR indicates both the size and the direction of difference. Values below 1.00 indicate that the normal score is higher; values above 1.00 indicate that the normal score is lower.*
the studies which may account for the discrepancy between their results and those reported here are: (1) the number of subjects used: 16 in this study, 48 and 108 in the others; (2) overall length of speech samples: 900 words for this study and 2,800-3,000 for the others; (3) type of verbal behavior: speech for this study and written passages for the others; and (4) procedures for obtaining speech samples.

Having found uniformly higher overall TTR's for normal than for matched schizophrenic subjects it is of interest to consider the relationship between the overall TTR and the variables on which the subjects were matched. The distribution of cases in this small matched sample does not permit separate analyses by age, sex, educational level, and ethnic group, since these factors tend to vary together. Two of the four pairs belong to the same ethnic group and are at the same educational level. When these two pairs are combined for comparison with the other pairs (Table 3), there is a difference in overall TTR for 900 words among the resulting three groups as indicated by the rank order of the scores.*

*Although the statistical tests of significance which were applied to this and other comparisons tend to support the descriptive statements made here, the composition of the samples makes their use questionable, and they are consequently not reported.

### Table 3

**Overall TTR Scores and Their Rank Order for Three Different Groups: Matched Sample**

<table>
<thead>
<tr>
<th>Italian women, ages 40 and 44, educ. &lt; 8 years</th>
<th>Negro men and women ages 27-40, educ. 9-11 years</th>
<th>Jewish men ages 23 and 24, educ. 12-14 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>Rank</td>
<td>Score</td>
</tr>
<tr>
<td>.288</td>
<td>1</td>
<td>.322</td>
</tr>
<tr>
<td>.294</td>
<td>2</td>
<td>.329</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.332</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.349</td>
</tr>
<tr>
<td>Mean rank</td>
<td>1.5</td>
<td>4.8</td>
</tr>
</tbody>
</table>

When the supplementary sample is included, separate evaluation can be made by sex, age, education, and ethnic group, to get an indication of the relative importance of these factors in affecting the TTR for this sample. It should be pointed out that such evaluation does not form sufficient basis for inferences about the popula-
tions from which these subjects were selected because of the particular distribution of cases (Table 4) – for example, all the Italian subjects are women, all the Jewish subjects are men, and only the Negro group includes both men and women.

The following tables, therefore, are intended only to suggest which of the several possible groupings of subjects shows the strongest relationship to the overall type-token ratio for this sample.

**Table 4**

**Overall TTR by Group**

<table>
<thead>
<tr>
<th>Ethnic group</th>
<th>Age group</th>
<th>Schizophrenic subjects</th>
<th>Normal subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>Jewish</td>
<td>under 26</td>
<td>.343</td>
<td>.292</td>
</tr>
<tr>
<td>Jewish</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jewish</td>
<td>26-33</td>
<td>.328</td>
<td></td>
</tr>
<tr>
<td>Jewish</td>
<td>over 33</td>
<td>.389</td>
<td></td>
</tr>
<tr>
<td>Negro</td>
<td>under 26</td>
<td>.329</td>
<td></td>
</tr>
<tr>
<td>Negro</td>
<td>26-33</td>
<td></td>
<td>.322</td>
</tr>
<tr>
<td>Negro</td>
<td>over 33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italian</td>
<td>under 26</td>
<td>.331</td>
<td></td>
</tr>
<tr>
<td>Italian</td>
<td>26-33</td>
<td></td>
<td>.317</td>
</tr>
<tr>
<td>Italian</td>
<td>over 33</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*"High" education means at least 12 years of school, whereas "low" education means less than 12 years of school.

In Table 5, the most extreme values appear to be those related to ethnic background. For the 16 subjects combined, the ethnic groups differ most strongly, the men seem somewhat higher than the women, the younger subjects higher than the older subjects, and the better educated higher than the less educated. The schizophrenic subjects separately seem to follow a similar pattern, though less strongly on the ethnic variable. The normal subjects separately do not show the same pattern for age, but show a perfect relationship with ethnic group.

For the combined sample the normal subjects generally have higher TTR's than the schizophrenic subjects (Table 6), on both overall TTR's and mean TTR's. That the differences are not con-
consistent (both groups include both low and high ranks) would be expected from the fact that three differing ethnic groups are included.

In summary, neither the overall $TTR$ nor the mean $TTR$ consistently distinguishes schizophrenic from unmatched normal subjects. Mean $TTR$'s for 100-word speech segments also do not discriminate schizophrenic from matched normal subjects, while overall $TTR$'s are consistently higher for the normal than for the matched schizophrenic subject for four pairs matched for age, sex, education, and ethnic group. This result seems inconsistent with the results of Fairbanks (1944), who used mean $TTR$'s for 100-word segments of 3,000-word speech samples and found significant differences between schizophrenic and normal subjects. The difference in results may be due to the small sample in the present study. On the other hand, the inconsistency may be due to the fact that Fairbanks' sampling procedures were designed to obtain maximum differences between the groups. The normal subjects in that study were college freshmen selected for their high scores on tests of English reading, comprehension, and training, and the schizophrenic subjects included chronic patients. Feldstein and Jaffe (1962), using 25-word segments, found no significant difference in mean segmental $TTR$'s between male schizophrenic and normal groups of average intelligence matched for age and education. However, it is difficult to evaluate their results, since 25-word segments may not be comparable to the larger segments, and since the schizophrenic patients in their sample were receiving tranquilizers, whose effect on the $TTR$ is not really known. They suggest that this might have increased any differences between the groups of subjects; however, in a study done at our laboratory (Salzinger, et al., 1961) with a normal subject given different dosages of chlorpromazine from 0 to 125 mg., indications were that $TTR$ increases with dosage. Such an effect in the schizophrenic sample cited might well have eliminated real differences in $TTR$ between their samples. In considering the possible effects of differences in the length of the speech segment, it is interesting to note that Grummon (1950, in an unpublished dissertation, reported in Mowrer, 1953) found that 500-word and 1,000-word overall $TTR$'s were related to degree of improvement under therapy whereas mean 100-word $TTR$'s were not.

The samples used here do not permit more than tentative statements, but ethnic background does appear to affect the overall $TTR$. Age, sex, and education do not as clearly influence the $TTR$, but must be retested in other samples selected for that purpose. Since ethnic group and level of education are known to be related, samples must be chosen with special attention to these two variables if their effects are to be independently evaluated.
<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Sex</th>
<th>Education</th>
<th>Age Group</th>
<th>Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Under 26 years</td>
<td>&gt; 12 years</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>26-33 years</td>
<td>33 years</td>
<td></td>
</tr>
<tr>
<td>Italian Negro</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jewish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Sex</th>
<th>Education</th>
<th>Age Group</th>
<th>Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Under 26 years</td>
<td>&gt; 12 years</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>26-33 years</td>
<td>33 years</td>
<td></td>
</tr>
<tr>
<td>Italian Negro</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jewish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean rank: 9.8

Type: TABLE

Subject: RANKS BASED ON OVERALL TTR, SEPARATELY BY SEX, AGE, EDUCATION, AND ETHNIC GROUP: COMBINED SAMPLE
<table>
<thead>
<tr>
<th>Schizophrenic Subjects</th>
<th>Mean rank:</th>
<th>Normal Subjects</th>
<th>Mean rank:</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>6.0</td>
<td>9</td>
<td>6.0</td>
</tr>
<tr>
<td>8</td>
<td>2.7</td>
<td>7</td>
<td>3.7</td>
</tr>
<tr>
<td>7</td>
<td>5.7</td>
<td>6</td>
<td>5.0</td>
</tr>
<tr>
<td>6</td>
<td>3.3</td>
<td>5</td>
<td>3.5</td>
</tr>
<tr>
<td>5</td>
<td>4.8</td>
<td>4</td>
<td>1.0</td>
</tr>
<tr>
<td>4</td>
<td>4.0</td>
<td>3</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>5.0</td>
<td>2</td>
<td>2.7</td>
</tr>
<tr>
<td>2</td>
<td>1.0</td>
<td>1</td>
<td>2.7</td>
</tr>
<tr>
<td>1</td>
<td>4.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Minimum number of types accounting for a given proportion of the speech sample. This measure is similar to one used by Fairbanks (1944) and Mann (1944), and is also related to the analysis by Whitehorn and Zipf (1943) of the relationship between rank (based on frequency) and frequency of words in long written passages. The measure used in the present study, like that used by Fairbanks, is based on speech rather than written passages, and is cumulative, showing the minimum number of different words (types) comprising a given proportion of the total number of tokens in the speech sample.

Schizophrenic patients were expected to rely more heavily on their most familiar words, and thus use relatively fewer different words for any given proportion of the total speech sample. At the upper limit this measure is, of course, identical with the number of types indicated in the overall TTR; but for smaller proportions of the speech sample this and the TTR measure may differ. For the eight subjects of the matched sample, the correlation between the minimum number of types and the overall TTR is not high for the smaller proportions of the speech samples: for \( \frac{1}{4} \) of the tokens \( r_s = 0.11 \); for \( \frac{1}{2} \), \( r_s = 0.45 \). For large proportions of the speech sample, of course, the correlation must approach 1; at \( \frac{3}{4} \), \( r_s = 0.90 \).

Table 7 shows for the four pairs of matched subjects, the minimum number of types necessary to account for one-fourth, one-half, and three-fourths of the tokens in the whole speech sample, as
well as the total number for the 900-word samples. Except for the second pair, for which the values for schizophrenic and normal subjects are the same through half the sample, the values for normals consistently exceed the values for the matched schizophrenic subjects. (Figure 1).

These results are not directly comparable with those of Fairbanks (1944), whose analyses of a similar measure for spoken passages were based on group usage, while these are computed for each individual. They do, however, seem consistent with her findings. Mann’s (1944) related study of written passages did not yield similar results.

The minimum number of types for each of these proportions below 1.00* were grouped in Table 8 in the same age-ethnic-education categories as in Table 5. The rank order of subjects’ scores at three-fourths of the sample is perfectly associated with these groupings, and almost the same as the rank order for the TTR. As indicated earlier, however, at the higher proportions of the sample, these two measures are necessarily related. The rank orders at one-fourth and at one-half of the sample, unlike the overall TTR, do not indicate any relationship between this measure and the age-ethnic-education groupings.

The minimum number of types accounting for a given proportion of the speech sample reflects a different dimension from the type-token ratio, namely, the relative weight of the most frequently used words rather than the number of different words used.

*Weighted type-frequency measure. A number of statistical techniques have been worked out for the purpose of developing a characteristic constant related to the type-token measure, but independent of length of sample (Herdan, 1960). Some question has been raised about the appropriateness of some of these techniques (Chotlos, 1944); and it is also unclear how large a sample might be the minimum proper to such analysis. Apart from the problem of independence of sample length, however, such a measure as Yule's Characteristic K (Herdan, 1960) differs from the TTR in that it takes into account the frequency of occurrence of each type. The speech samples used here were of the same length (900 words) for each subject, and schizophrenic subjects were compared with matched normal subjects on this measure (Table 9).

Again, as with the overall TTR and the minimum types for a given proportion of tokens, there are no reversals of direction among the matched pairs. Each schizophrenic subject has a higher Char-

*At 1.00, of course, this measure is identical with the number of types used in the TTR.
<table>
<thead>
<tr>
<th></th>
<th>Schizo-phrenic</th>
<th>Normal</th>
<th>Whole sample (900 tokens)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First pair</td>
<td>5</td>
<td>8</td>
<td>29</td>
</tr>
<tr>
<td>Second pair</td>
<td>6</td>
<td>6</td>
<td>33</td>
</tr>
<tr>
<td>Third pair</td>
<td>6</td>
<td>8</td>
<td>26</td>
</tr>
<tr>
<td>Fourth pair</td>
<td>6</td>
<td>7</td>
<td>32</td>
</tr>
</tbody>
</table>

TABLE 7

Minimum number of types necessary to account for a given proportion of each 900-word speech sample: matched Schizophrenic and Normal subjects
FIGURE 1. Minimum number of types necessary to account for a given proportion (0.25, 0.50, 0.75, 1.00) of each 900-word speech sample, for four schizophrenic-normal matched pairs.
characteristic $K$ than his matched normal subject. Moreover, there is no overlap between the schizophrenic and the normal subjects, which suggests that this measure is more sensitive to manifestations of schizophrenia than is the $TTR$.

Tabulating these scores by age-ethnic-education grouping

**Table 8**

Minimum Number of Types (and Rank Orders) for $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$ of Speech Samples, for Three Different Groups

<table>
<thead>
<tr>
<th>Matched sample</th>
<th>One-quarter of sample</th>
<th>One-half of sample</th>
<th>Three-quarters of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italian women, ages 40 and 44, educ. &lt; 8 years</td>
<td>Negro men and women, ages 27-40, educ. 9-11 years</td>
<td>Jewish men, ages 23 and 24, educ. 12-14 years</td>
<td></td>
</tr>
<tr>
<td>Types</td>
<td>Rank</td>
<td>Types</td>
<td>Rank</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>6</td>
<td>3.5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>7.5</td>
<td>6</td>
<td>3.5</td>
</tr>
<tr>
<td>6</td>
<td>3.5</td>
<td>8</td>
<td>7.5</td>
</tr>
<tr>
<td>Mean rank:</td>
<td>5.5</td>
<td>3.9</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td>35</td>
<td>7</td>
<td>33</td>
<td>5</td>
</tr>
<tr>
<td>33</td>
<td>5</td>
<td>33</td>
<td>5</td>
</tr>
<tr>
<td>Mean rank:</td>
<td>4.0</td>
<td>4.3</td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>1</td>
<td>99</td>
<td>3</td>
</tr>
<tr>
<td>93</td>
<td>2</td>
<td>105</td>
<td>4</td>
</tr>
<tr>
<td>106</td>
<td>5</td>
<td>108</td>
<td>6</td>
</tr>
<tr>
<td>Mean rank:</td>
<td>1.5</td>
<td>4.5</td>
<td>7.5</td>
</tr>
</tbody>
</table>
$K = \frac{Er^2n_r - N}{N^2} \times 100$ where $r =$ frequency

$n_r =$ number of types with frequency $r$

$N =$ number of tokens in sample (900)

(Table 10), as was done with the TTR and the minimum types, shows no regular ordering.

Thus, although the TTR, the minimum number of types, and the $K$ are all measures of word repetition, the $TTR$, based on vocabulary used, seems highly dependent on ethnic grouping (and possibly other background factors), while the minimum number of types for proportions of one-fourth and one-half the tokens, and Yule's Char-

**Table 9**

**K-Values for Matched Schizophrenic and Normal Subjects**

<table>
<thead>
<tr>
<th></th>
<th>Schizophrenic subject</th>
<th>Normal subject</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>First pair</td>
<td>2.15</td>
<td>1.28</td>
<td>+.87</td>
</tr>
<tr>
<td>Second pair</td>
<td>1.66</td>
<td>1.42</td>
<td>+.24</td>
</tr>
<tr>
<td>Third pair</td>
<td>1.85</td>
<td>1.44</td>
<td>+.41</td>
</tr>
<tr>
<td>Fourth pair</td>
<td>1.51</td>
<td>1.47</td>
<td>+.04</td>
</tr>
</tbody>
</table>

**Table 10**

**K-Values and Their Rank Order for Three Different Groups**

<table>
<thead>
<tr>
<th>Matched Sample</th>
<th>Italian women, ages 40 and 44, educ. &lt; 8 years</th>
<th>Negro men and women, ages 27-40, educ. 9-11 years</th>
<th>Jewish men, ages 23 and 24 educ. 12-14 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K$</td>
<td>Rank</td>
<td>$K$</td>
<td>Rank</td>
</tr>
<tr>
<td>1.44</td>
<td>3</td>
<td>1.28</td>
<td>1</td>
</tr>
<tr>
<td>1.85</td>
<td>7</td>
<td>1.42</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.66</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.15</td>
<td>8</td>
</tr>
<tr>
<td>Mean rank:</td>
<td>5.0</td>
<td>4.3</td>
<td>4.5</td>
</tr>
</tbody>
</table>
acteristic $I$, both involving word frequencies, seem independent of these background factors.

**Word-sequence distributions.** In addition to the measures of repetitiveness already described, all of which deal with single words, an analysis was made of the amount of repetition of word sequences. It would be consistent with the hypothesis indicated earlier to expect that the schizophrenic subjects would tend to repeat longer word sequences than the normal subjects.

All possible two-word sequences (899 sequences in the 900-word passage) were used, and their frequency distribution was obtained. Similarly, 898 three-word sequences, 897 four-word sequences, etc., were used. The analysis was continued until all exactly repeated sequences were included.

Efficient comparison of the resulting distributions will depend upon the development of a single weighted measure based upon length of sequence and frequency of occurrence. Results of comparisons of several aspects of word sequence repetition are unclear. In two schizophrenic-normal pairs, the longest repeated sequence in the normal speech sample exceeded the length of that in the schizophrenic speech sample, in one pair the opposite held true, and in one pair the lengths were the same. **Table** 11 (a) shows the total number of "types" for sequences of one, two, three, and four words. (At about four or five words, the number of repetitions becomes quite small and probably quite unstable). The normal subject generally uses a higher number of "types" than his matched schizophrenic subject, but not without exception (**Figure** 2). In **Table** 11b the number of repeated single-word types has been used as a base for computing the ratios of two-, three-, and four-word repeated sequences. The schizophrenic subjects generally have higher ratios (**Figure** 3) and the exceptions follow a similar pattern to the exceptions in the number of "types," shown in **Figure** 2. These higher ratios indicate a relatively slower decrease in the amount of repetition as the sequence-length increases.

**Interval distributions.** The final measure on word distribution within the speech sample is the distance, in number of words, between recurrences of repeated words and word sequences. It was expected that if, as has been suggested, schizophrenic subjects are relatively more influenced than are normal subjects by their own immediately preceding behavior, their speech should not only be more repetitious, but the repetitions should show a more closely clustered pattern—i.e., a tendency to shorter intervals between repetitions.

A number of mathematical problems require resolution in order
**Table 11 a & b**

**Comparison of 1-, 2-, 3-, and 4-word sequences for four schizophrenic-normal matched pairs**

(a) Total types (number of different 1-, 2-, 3-, and 4-word sequences)

<table>
<thead>
<tr>
<th></th>
<th>1-word</th>
<th>2-word</th>
<th>3-word</th>
<th>4-word</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Schizophrenic</td>
<td>Normal</td>
<td>Schizophrenic</td>
<td>Normal</td>
</tr>
<tr>
<td>First pair</td>
<td>296</td>
<td>299</td>
<td>658</td>
<td>721</td>
</tr>
<tr>
<td>Second pair</td>
<td>290</td>
<td>314</td>
<td>676</td>
<td>729</td>
</tr>
<tr>
<td>Third pair</td>
<td>259</td>
<td>265</td>
<td>665</td>
<td>687</td>
</tr>
<tr>
<td>Fourth pair</td>
<td>309</td>
<td>324</td>
<td>739</td>
<td>722</td>
</tr>
</tbody>
</table>

(b) Ratio of repeated multi-word types to repeated single-word types

<table>
<thead>
<tr>
<th></th>
<th>2-word</th>
<th>3-word</th>
<th>4-word</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Schizophrenic</td>
<td>Normal</td>
<td>Schizophrenic</td>
</tr>
<tr>
<td>First pair</td>
<td>.94</td>
<td>.71</td>
<td>.39</td>
</tr>
<tr>
<td>Second pair</td>
<td>.90</td>
<td>.80</td>
<td>.36</td>
</tr>
<tr>
<td>Third pair</td>
<td>.87</td>
<td>.83</td>
<td>.36</td>
</tr>
<tr>
<td>Fourth pair</td>
<td>.71</td>
<td>.75</td>
<td>.27</td>
</tr>
</tbody>
</table>
to develop a single measure of clustering, primarily because the basic word-frequency distributions differ. The two gross measures which have so far been applied to these intervals have shown no consistent pattern. For the 16 subjects of the combined sample,
comparisons between schizophrenic and normal subjects show no systematic differences in mean interval lengths for words occurring at each of the frequencies. For the eight subjects of the matched sample, the intervals between recurrences of each of the 10 most frequent types for each subject seem to depart only randomly from what might be expected if the repetitions of a type were evenly distributed throughout the passage (Kolmogorov-Smirnov one-sample test, Siegel, 1956). The 10 most frequent types are, with few exceptions, syntactic (function) words (see Fries, 1952), such as "and," "the," "to," which may tend to a more even distribution than the content (lexical) words. The less frequent types, however, cannot be compared by the same procedure. The mean interval measure, on the other hand, can be used for all repeated types, but it is insensitive to any differences in distribution between the first and last occurrence of a type, i.e., the mean interval for any type for a given subject is simply the distance in number of words between the first and last occurrences divided by the number of intervals. Medians or modes may be preferable for this kind of analysis. However, it will probably be most suitable to develop appropriate adjustment factors which will allow analysis across all frequencies, resulting in a single measure of clustering of repetitions for a whole passage.

Common vocabulary. Another speech characteristic in which schizophrenic patients may manifest divergence from their social group is specific vocabulary usage.

In this analysis, each subject was paired with every other subject to obtain a score for the number of types common to both members of the pair.* The Italian group was omitted since it included only one normal subject, and could, therefore, not produce any score for a normal pair. The remaining 12 subjects were two Negro schizophrenics, four Negro normals, four Jewish schizophrenics and two Jewish normals. The scores for the 66 resulting pairs were used as the basis for several comparisons concerned with the relative similarity of schizophrenic with other schizophrenic subjects, normal with other normal subjects, and schizophrenic with normal subjects; with a corresponding analysis comparing the two ethnic groups; and with an attempt to assess the relative vocabulary "distances" among the subjects.

(1) Comparison of schizophrenic and normal subjects: The score for each subject consisted of the difference between the median

*Since frequencies of occurrence for each type were not taken into account in this score, it is possible that differences in total number of types may affect the scores.
value of his scores (number of types in common) with other subjects of the same group (schizophrenic with schizophrenics, normal with normals) and the median value of his scores with the subjects of the other group (schizophrenic with normals, normal with schizophrenics). Thus, positive differences would imply that the subject has greater similarity with the members of his own group, whereas negative differences would imply that he has greater similarity with the members of the other group. Table 12 shows that each schizophrenic subject is more similar on the average to the normal subjects than to the other schizophrenic subjects. The normal subjects, on the other hand, are more similar to each other than to the schizophrenic subjects.

**Table 12**

**Vocabulary Comparisons Between Schizophrenic and Normal Subjects**

<table>
<thead>
<tr>
<th>Schizophrenic subjects</th>
<th>Normal subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same-group median</td>
<td>Other-group median</td>
</tr>
<tr>
<td>100</td>
<td>107</td>
</tr>
<tr>
<td>101</td>
<td>112.5</td>
</tr>
<tr>
<td>101</td>
<td>103.5</td>
</tr>
<tr>
<td>100</td>
<td>113</td>
</tr>
<tr>
<td>114</td>
<td>120</td>
</tr>
<tr>
<td>98</td>
<td>99.5</td>
</tr>
</tbody>
</table>

(2) Comparison of Negro and Jewish subjects: When a similar comparison is made between the two ethnic groups (Table 13), there is no indication of greater similarity within or between the groups. There seems to be a tendency for the Negro subjects to be more similar to each other than to the Jewish subjects; the Jewish subjects show no consistent tendency in either direction. It should be noted, however, that this may in part be the effect of the larger proportion of schizophrenic subjects in the Jewish group.

(3) Distances between subjects: For each subject, a rank order of scores was obtained in which he occupied Rank 1, and the subject with whom he had the fewest words in common occupied Rank 12. A new "distance" score was obtained for each pair of subjects based on their relative similarity with respect to all other subjects. For example, the score for pair C-F was obtained by summing differences in ranks for each subject (Table 14).
### Table 13

**Vocabulary Comparisons, Negro and Jewish Subjects**

<table>
<thead>
<tr>
<th></th>
<th>Negro</th>
<th></th>
<th>Jewish</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Same-group median</td>
<td>Other-group median</td>
<td>Difference</td>
<td>Same-group median</td>
</tr>
<tr>
<td></td>
<td>107</td>
<td>99</td>
<td>+8</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>105</td>
<td>100</td>
<td>+5</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>115</td>
<td>113</td>
<td>+2</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>117</td>
<td>117</td>
<td>0</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>115</td>
<td>108.5</td>
<td>+6.5</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>123.5</td>
<td>-2.5</td>
<td>98</td>
</tr>
</tbody>
</table>

Thus, lower distance scores indicate greater similarity between subjects. The range of scores was 9-52. A score of 30 was arbitrarily chosen as the upper value for considering two subjects to be "connected." Of the resulting 19 "connections," the schizophrenic subjects each have from zero to three connections with any other subject, while the normal subjects each have four or five connections. The number of connections for each type of pair, and the

### Table 14

**Illustrative "Distance" Scores: Subject C with Subject F**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Number of words in common with subject C</th>
<th>Number of words in common with subject F</th>
<th>Rank order for subject C</th>
<th>Rank order for subject F</th>
<th>Differences in rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>120</td>
<td>108</td>
<td>4</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>114</td>
<td>101</td>
<td>7.5</td>
<td>9</td>
<td>1.5</td>
</tr>
<tr>
<td>C</td>
<td>323*</td>
<td>109</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td>114</td>
<td>102</td>
<td>7.5</td>
<td>7.5</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>121</td>
<td>105</td>
<td>3</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>F</td>
<td>109</td>
<td>290*</td>
<td>10</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>G</td>
<td>107</td>
<td>102</td>
<td>11</td>
<td>7.5</td>
<td>3.5</td>
</tr>
<tr>
<td>H</td>
<td>124</td>
<td>115</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>I</td>
<td>118</td>
<td>99</td>
<td>5</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>J</td>
<td>113</td>
<td>114</td>
<td>9</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>K</td>
<td>115</td>
<td>99</td>
<td>6</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>L</td>
<td>104</td>
<td>99</td>
<td>12</td>
<td>11</td>
<td>1</td>
</tr>
</tbody>
</table>

*The total number of types used by the subject.*
possible number of such pairs in this sample, is shown in Table 15. The normal subjects are closer to each other, in terms of this measure, than are any other combinations of subjects.

**Table 15**

<table>
<thead>
<tr>
<th></th>
<th>Possible pairs</th>
<th>&quot;Connected&quot; pairs</th>
<th>Proportion &quot;connected&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal-normal pairs</td>
<td>15</td>
<td>10</td>
<td>.67</td>
</tr>
<tr>
<td>Normal-schizophrenic pairs</td>
<td>36</td>
<td>7</td>
<td>.19</td>
</tr>
<tr>
<td>Schizophrenic-schizophrenic pairs</td>
<td>15</td>
<td>2</td>
<td>.13</td>
</tr>
</tbody>
</table>

Table 16 shows a similar comparison for Negro and Jewish subjects. There appears to be no sizeable difference in relative closeness of members of these two ethnic groups.

**Table 16**

<table>
<thead>
<tr>
<th></th>
<th>Possible pairs</th>
<th>&quot;Connected&quot; pairs</th>
<th>Proportion &quot;connected&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jewish-Jewish pairs</td>
<td>15</td>
<td>3</td>
<td>.20</td>
</tr>
<tr>
<td>Jewish-Negro pairs</td>
<td>36</td>
<td>11</td>
<td>.31</td>
</tr>
<tr>
<td>Negro-Negro pairs</td>
<td>15</td>
<td>5</td>
<td>.33</td>
</tr>
</tbody>
</table>

In summary, with respect to the shared vocabulary measures so far employed, differences between these ethnic groups seem to be unimportant. Normal subjects seem to be relatively more similar to each other, with schizophrenic subjects somewhat more diverse among themselves, as well as divergent from the normal subjects.

**Summary and Conclusions**

Several approaches to the measurement of some formal aspects of speech have been explored, preparatory to investigating their use as an index of social deviance. The value of using speech for this purpose lies first in its susceptibility to precise analysis, and second, in both its dependence and its influence upon social interaction.

The speech of schizophrenic patients has often been referred to
clinically as "stereotyped," and previous studies have given some quantitative confirmation to this judgment. The first three measures of word-use variation studied here (the type-token ration, minimum number of types, and Yule's Characteristic $K$) also seem to confirm this judgment, at least in comparing schizophrenic with normal subjects of the same sex who are similar in age, education, and ethnic background. The fact that ethnic background and possibly education seem to be stronger influences than schizophrenic illness on the type-token ratio suggests not only that great care must be taken in sample-selection for studying this phenomenon, but also, and more important, that great care must be taken in interpreting the results in terms of the processes that may have produced them. The minimum number of types and Yule's Characteristic $K$, on the other hand, seem less related to these background variables and more affected by schizophrenic illness.

The other distribution measures, word-sequence repetitiveness and interval length, require considerably more development statistically, as well as empirically, before they can be evaluated for their possible usefulness as stylistic indices.

The measures based on shared vocabulary seem not to be influenced by these ethnic group differences. Loosely interpreted, the schizophrenic subjects' greater distances on these measures may be seen as a reflection of greater idiosyncrasy of speech, which can, of course, occur along with the greater stereotypy manifested in the lower type-token ratios and minimum number of different words for various proportions of the passage, and in higher values of Yule's Characteristic $K$.

In addition to the obvious need for using larger samples and studying other languages and dialects, the results of this study indicate the following emphases for continuing research: (1) selection of samples for testing separately the influence of ethnic group, education, and other background variables on these aspects of speech; (2) testing these measures in speech samples derived from "ordinary" interviews and conversations, as well as monologues; (3) testing the relationship between standard vocabulary scores and these measures of speech variation; (4) testing the effect on these measures of differences in social context (as given or as experimentally manipulated), such as the relationship between the speaker and the listener, the number of listeners, the "purpose" of speaking (reinforcement contingency); (5) testing the relationship of these measures to a rough independent criterion of "deviance," such as length of hospitalization for schizophrenia; and (6) exploring the use of other kinds of measures, involving other aspects of speech usage.
It would be premature at this stage in our research to attempt to specify the kinds of measures which may be appropriate for the characterization of speech usage normatively for groups or for individuals in relation to groups. However, since only a very restricted set of measures has been dealt with here, some mention should perhaps be made of other kinds of measures and of the reasons why a variety of measures seems to be called for. It may be suggested — although, to our knowledge, without systematic empirical evidence — that different aspects of speech tend to be differentially subject to influence from the several kinds of groupings an individual may be a member of. For example, sound patterns (intonation, etc.) are necessarily acquired early and used regularly, and may therefore be most subject to familial influence, and thereby to the influence of the linguistic grouping (ethnic, regional, etc.) which is part of the familial background. Vocabularly, on the other hand, is probably subject to alteration — at least by addition — through most of life, and might better reflect “mass culture,” formal education, and peer group influences. Certain stylistic features, such as sentence length and complexity, sequences of word classes, and some of the word-use features dealt with in this paper, may reflect the “dominant mode” of language use for a given individual or group. (It is assumed here that there are different modes — implying more or less distinctive conventional features — examples of which might be objective description, mood expression, argument, conversation; and that there are differences in the relative importance of these modes for different people, such as journalists, poets, union leaders, farmers, and playboys. It is not assumed that the examples of different modes, or the implication that occupation is necessarily associated with such differences, are correct, but only that there are different dominant modes for different groups of people). If we want to measure an individual’s distance from a group, or a group’s distance from another group, or two individuals from each other, it seems most probable that a variety of measures, concerned with different aspects of speech, will be necessary.

Acknowledgment

The authors wish to acknowledge the able assistance of Verna Schmauder in the analysis of data, and J. Oppenheim in collecting the monologues of the normal subjects.

References

and comparative analysis of individual written language samples. Psychol. Monogr. 56: 77-112.


A CHILDREN’S BEHAVIOR DIAGNOSTIC INVENTORY*

Eugene I. Burdock and Anne S. Hardesty

The aim of the project under which the current research has been undertaken is the development of quantitative techniques of observation to be applied to mental patients as measures of severity of illness and indices of response to treatment. In pursuit of this goal several inventories have been developed for use in different settings. The Ward Behavior Rating Scale has capitalized on the ability of nurses to identify changes in behavior due to drug effects, when provided with a standardized inventory and appropriate training. A Mental Status Schedule and Inventory has been designed to provide a standard procedure for use by the psychiatrist in mental status assessments on admission and at various critical points in a patient’s hospital course.

Despite the many problems inherent in assessing severely disturbed adult patients, the judgment that a patient is performing adaptively is made somewhat less complex by the fact that there are generally recognized social norms for the behavior of adults. When the task is to assess mentally ill children, two components of behavior must be separated: the manifestations of illness (i.e., the behavioral correlates), and the developmental level (for the same behaviors have different adaptive significance at different ages.) To meet the needs for an instrument which would inventory maladaptive behaviors in children at different age levels, a Children’s Behavior Diagnostic Inventory has been constructed.

Initially, a list of items was drawn up which clinical experience suggested were significant of varying degrees of psychopathology. Psychiatric texts and articles were scanned for behaviors characteristic of emotional disorders of childhood (Eender, Caplan, Chess, English & Pearson, Kanner, Mahler, etc.). Each symptom was checked in more than one important source. If a symptom was described in only one source it was discarded, so that the resultant list consisted of generally recognized symptoms.

The symptoms were then screened to identify the age at which they first become significant of developmental anomaly. A number of sources (Macfarlane, Honzik, & Allen, 1954: Lapouse & Monk, 1958) make reference to ages at which certain asocial or problem behaviors may be expected to disappear in the majority of children. From other discussions extrapolations could be made, although specific norms were not available. The next step was to translate

*Supported in part by grant M-3546 from the National Institute of Mental Health.
the symptom into a behavioral item that was observable. The items were then grouped by age levels into six two-year intervals ranging from one to twelve.

The observer's task is to indicate the presence or absence of maladaptive behavior by encircling T (true) or NT (not true) for all those items up through the age level which includes the child's last birthday. Only behaviors actually observed are recorded. Each maladaptive behavior is given unit weight. Total score is the percentage of maladaptive behaviors reported within the corresponding age group. This scoring procedure allows comparisons across age groups, since the number of items increases with age.

The number of items at each age level increases most rapidly for the years up to and including age of entry into school. The biggest increment is in the category of Socialization. The numbers of items for each age level are as follows:

<table>
<thead>
<tr>
<th>Age level</th>
<th>No. of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>11</td>
</tr>
<tr>
<td>3-4</td>
<td>41</td>
</tr>
<tr>
<td>5-6</td>
<td>80</td>
</tr>
<tr>
<td>7-8</td>
<td>114</td>
</tr>
<tr>
<td>9-10</td>
<td>130</td>
</tr>
<tr>
<td>11-12</td>
<td>137</td>
</tr>
</tbody>
</table>

For descriptive purposes the items have been arbitrarily classified into six behavioral areas: (1) Vegetative Functioning (eating, sleeping, and toileting): Wets bed. Keeps refusing to eat. (2) Appearance and Mannerisms (facial expression, psychomotor activity): Engages in rhythmic motions (swaying, head rolling, etc.). Washes hands or does other things over and over again. Grimaces or gestures grotesquely. (3) Speech and Voice (quantity of speech, phonation): Utters no articulate sound. Screams. (4) Emotional Display (fear, anger): Cries or whimpers. Has a temper tantrum. (5) Socialization (interaction with others, conformity to social norm): Takes part in ongoing activities without being urged. Stays by himself. Runs away or plays truant. (6) Thought Processes (memory, attention, concentration, ideation): Acts perplexed or confused. Keeps forgetting detail, task, or event.

The number of items in each of the six behavioral areas for each age level is shown in Table 1.

In a preliminary study of reliability, 15 mentally ill children were observed by two nurses. The nurses served on different but overlapping shifts and had been instructed to confine their judgments to observations made on two consecutive days and to disre-
TABLE 1

DISTRIBUTION BY BEHAVIORAL AREA AND AGE LEVEL OF ITEMS ON THE CHILDREN'S BEHAVIOR DIAGNOSTIC INVENTORY

<table>
<thead>
<tr>
<th>Behavioral area</th>
<th>Age level</th>
<th>1-2</th>
<th>3-4</th>
<th>5-6</th>
<th>7-8</th>
<th>9-10</th>
<th>11-12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetative Functioning (VF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appearance &amp; Mannerisms (AM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech &amp; Voice (SV)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional Display (ED)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socialization (S)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thought Processes (TP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>11</td>
<td>41</td>
<td>80</td>
<td>114</td>
<td>130</td>
<td>137</td>
</tr>
</tbody>
</table>

In regard earlier knowledge of a child's behavior. When their observations were compared, a product-moment correlation of 0.72 was obtained. The results of an analysis of variance applied to the data and shown in TABLE 2.

TABLE 2

ANALYSIS OF VARIANCE OF SCORES ON CHILDREN'S BEHAVIOR DIAGNOSTIC INVENTORY

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Sum of squares</th>
<th>d. f.</th>
<th>Mean square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>2360.867</td>
<td>14</td>
<td>168.633</td>
</tr>
<tr>
<td>Observers</td>
<td>353.633</td>
<td>1</td>
<td>353.633 F = 9.468**</td>
</tr>
<tr>
<td>Residual</td>
<td>522.867</td>
<td>14</td>
<td>37.47</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>3237.367</td>
<td>29</td>
<td></td>
</tr>
</tbody>
</table>

\[
\frac{\sigma^2_{\text{obs}}}{N} = \frac{\text{MS}_{\text{obs}} - \text{MS}_{\text{res}}}{\text{N}} = 21.08
\]

\[
R_{\text{intraclass}} = \frac{\text{MS}_{pt} - \text{MS}_{\text{res}}}{\text{MS}_{pt} + \text{MS}_{\text{res}}} = .64
\]
Note that the difference between the two observers is statistically significant. The estimate of the component of variance due to observers is 21.08. The intraclass correlation is smaller than the product-moment correlation. It is a more conservative index because it is more sensitive to differences in the variability of the two observers. Its magnitude of 0.64 indicates that approximately two-thirds of the variance of a score is attributable to differences among patients.

The inventory was tried out on several samples of normal children. Table 3 presents their average percentage scores. This table gives a picture of the range of scores by age and sex in groups with complete and incomplete observations. For both groups, the boys' scores are consistently higher than the girls' scores.

There is a rise in score from age level three-four to age level five-six for both sexes, and in about the same proportions. There is some suggestion that scores decline from this high point and level off with increasing age, but the data are really too meager to support such an inference.

Table 4 compares the scores of samples of normal children with those of abnormal children. These data were collected through the cooperation of several agencies in order to obtain a tentative indication of the power of the Children's Behavior Diagnostic Inventory to discriminate between normals and abnormals. The normals are the same children shown in Table 2, but now grouped according to observational setting. The first three lines are based on observations by teachers in a nursery school, a public school, and a parochial school. The fourth line represents observations by student nurses of 21 children hospitalized for physical illnesses. Campers were rated by camp counselors. There were two campers whose scores were the lowest in any of the normal groups. There was a large gap between their scores and those of the rest of the campers.

The first group of 22 abnormal children were observed by psychiatric nurses on a psychiatric ward. They were of mixed diagnoses. The autistic children were in a special day-school program. They were observed by a psychiatrist. The last line on the table represents a small group of children in a day school for mentally handicapped children, who were observed by their remedial teachers. Among them were six children known to be brain-damaged. The range of IQ's was 25 to 58. Mean score for the six children with known brain damage was 22.8. For the others the mean score was 27.2.

The striking feature of Table 4 is the sharp separation be-
### Table 3

**Children's Behavior Diagnostic Inventory**

**Distribution of Scores of Normals by Age, Sex and Observation Setting**

#### Complete observation

(including eating and sleeping behavior)

<table>
<thead>
<tr>
<th>Age</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>X</td>
</tr>
<tr>
<td>3-4</td>
<td>27</td>
<td>8.3</td>
</tr>
<tr>
<td>5-6</td>
<td>8</td>
<td>12.1</td>
</tr>
</tbody>
</table>
| [7-8 | 1     | 4.0     | -    | 3     | 7.0     | 3.5  |}
| 9-10 | 16    | 10.0    | 7.7  | [2    | 10.0    | 2.8  |
| 11-12| 7     | 10.1    | 11.0 | [1    | 2.0     | -    |}
| Total| 59    | 9.4     | 8.0  | 39    | 6.1     | 5.1  |

#### Incomplete observation

(omitting eating and sleeping behavior)

<table>
<thead>
<tr>
<th>Age</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>X</td>
</tr>
<tr>
<td>7-8</td>
<td>14</td>
<td>6.4</td>
</tr>
<tr>
<td>9-10</td>
<td>7</td>
<td>7.1</td>
</tr>
</tbody>
</table>
| [11-12| 3    | 10.0    | 7.9  | 1     | 9.0     | -    |}
| Total| 24    | 7.0     | 4.9  | 17    | 4.5     | 3.0  |

Note: Brackets enclosed those estimates of mean and standard deviation based on less than four cases.

Between normals and abnormals with relatively little overlap despite the small sizes of the samples.

Finally, a small short-term replication of observations was carried out on five children on a psychiatric ward. The same observers, psychiatric nurses, observed the children twice at an interval of
Comparison of Normals and Abnormals on the Children's Behavior Diagnostic Inventory

<table>
<thead>
<tr>
<th>Sample</th>
<th>N</th>
<th>Mean age</th>
<th>Mean score</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursery school</td>
<td>59</td>
<td>4</td>
<td>7</td>
<td>6.7</td>
</tr>
<tr>
<td>Public school</td>
<td>15</td>
<td>8</td>
<td>7</td>
<td>4.5</td>
</tr>
<tr>
<td>Parochial school</td>
<td>29</td>
<td>9</td>
<td>6</td>
<td>4.5</td>
</tr>
<tr>
<td>Pediatric hospital</td>
<td>21</td>
<td>8</td>
<td>8</td>
<td>5.4</td>
</tr>
<tr>
<td>Summer camp for boys</td>
<td>21</td>
<td>10</td>
<td>11</td>
<td>8.5</td>
</tr>
<tr>
<td>Totals</td>
<td>145</td>
<td></td>
<td>7.4</td>
<td>6.2</td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatric institute</td>
<td>22</td>
<td>9</td>
<td>20</td>
<td>11.4</td>
</tr>
<tr>
<td>School for autistic</td>
<td>14</td>
<td>6</td>
<td>21</td>
<td>4.6</td>
</tr>
<tr>
<td>children</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School for mental</td>
<td>11</td>
<td>6</td>
<td>25</td>
<td>11.3</td>
</tr>
<tr>
<td>defectives</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>47</td>
<td></td>
<td>21.5</td>
<td>9.9</td>
</tr>
</tbody>
</table>

The Children's Behavior Diagnostic Inventory is one of a series of observational techniques developed by the Biometrics Research Unit. Its purpose is to provide a tool for the screening of psychiatric casualties from the general population and a criterion measure for assessing change in response to treatment. Several preliminary studies are reported: (1) estimate of reliability of observations; (2) distribution of scores of normal children; (3) comparison of normals with abnormals, and (4) comparison of initial scores with scores on four-month follow-up.

Summary

The Children's Behavior Diagnostic Inventory is one of a series of observational techniques developed by the Biometrics Research Unit. Its purpose is to provide a tool for the screening of psychiatric casualties from the general population and a criterion measure for assessing change in response to treatment. Several preliminary studies are reported: (1) estimate of reliability of observations; (2) distribution of scores of normal children; (3) comparison of normals with abnormals, and (4) comparison of initial scores with scores on four-month follow-up.
### TABLE 5

**COMPARISON OF INITIAL SCORE WITH SCORE ON FOUR-MONTH FOLLOW-UP FOR FIVE CHILDREN ON A PSYCHIATRIC WARD**

<table>
<thead>
<tr>
<th>Subject</th>
<th>No. of observers</th>
<th>Initial inventory Mean score</th>
<th>Rank</th>
<th>Follow-up inventory Mean score</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>26.7</td>
<td>1</td>
<td>29.0</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>24.0</td>
<td>2</td>
<td>21.5</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>21.7</td>
<td>3</td>
<td>26.7</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>13.0</td>
<td>4</td>
<td>17.0</td>
<td>4</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>10.5</td>
<td>5</td>
<td>15.5</td>
<td>5</td>
</tr>
</tbody>
</table>

$\bar{X} = 19.2$  
$s = 7.1$

These preliminary studies provide sufficient evidence of the inventory's potential to warrant an investment in controlled studies of its reliability and discriminating power. Reliability will be examined by comparing the judgments of specially trained observers who will base their judgments on simultaneous but independent observations. The power of the inventory to discriminate normals from abnormals will be investigated by contrasting the judgments of the same observers on samples drawn from the two populations.

### References


Aging is a complex phenomenon; information concerning this phenomenon is obtained by a number of investigative approaches. An increasingly important social problem, aging is being studied from the aspect of the change in status undergone by the elderly individual with diminishing social roles and privileges. Psychologically, attitude changes of the aging individual can be studied, reflecting his altered mentation and the changed social expectations directed at him in interpersonal relationships. Medically, the mental status is an important determinant of individual behavior, accessible for evaluation by means of the techniques of the psychiatric examination.

The present investigation is a collaborative one. A sample of aged persons was studied by each of the three approaches indicated above, a sociologist, a social psychologist, and a psychiatrist independently making observations. The investigation studied the factors determining and modifying the capacity of elderly persons to adjust socially. The present report focuses on the social effects produced by the various types of mental disorder.

**Psychiatric Impairment in Old Age**

Mental disorder in the aged is of two types: (a) Loss of intellectual ability resulting from organic brain impairment (senile and arteriosclerotic dementia). (b) Functional mental disorder with psychiatric symptoms of behavior disturbance resembling those present in psychiatric patients at any age.

The latter category may go undiagnosed, the assumption often being made that in cases of mental disorder in the aged the symptoms are due to organic brain deterioration (dementia), without specific investigations having been undertaken to detect whether functional mental illness is present. Follow-up and neuropathological studies have made untenable "...the view that cerebral disease has a ubiquitous etiological role in old age mental disorder."(Kay & Roth, p. 649)

Recently, there has been repeated emphasis that a distinction,
of great clinical importance, is possible between organic brain disease and functional disorder. Dementia is a progressive and untreatable disorder; functional illness, on the other hand, may be reversible with adequate psychiatric treatment (Roth, 1960). Recognition clinically and effective treatment of reversible conditions is urgently required if progress is to be made with the serious social problem (Locke et al.) presently by mental illness in the elderly. Expectation of mental illness rises rapidly with increase in age, as is shown by the steep rise with age in the rate of first admissions to mental hospitals. Furthermore, of all patients in wards for the old and infirm, a third or more have been estimated to suffer from mental disorder (Howell).

A. **Organic brain syndrome** (Dementia). Intellectual deterioration in the aged presents clinically with progressive and irreversible loss of memory as the chief symptom. Some memory impairment is accepted as "natural" in the elderly. Conventional psychiatric examination of an elderly person brought for medical evaluation includes cognitive tasks aiming to test the subject’s memory functions, attention and concentration, and orientation. A clinical diagnosis of chronic brain syndrome rests on the psychiatrist’s judgment that there has been a decline in the subject’s intellectual abilities, as evaluated by clinical cognitive tests (e.g., subtracting 7 from 100 serially, repetition back by the subject of an anecdote like the "Cowboy Story"). Evidence of memory impairment is the central deficit the clinician looks for, the aged subject not knowing items of current information (e.g., who the president of the U.S.A. is) and being unable to provide a coherent account of recent or old past experiences. Characteristically recent memory is more impaired than recollection of remote events.

Examples from the subjects investigated will indicate specifically how this diagnostic category was applied:

*Dementia, mild,* Mr. I. H. aged 79: "There's an actor in the house. Impossible! He takes the paper, and when he's through with it he tears it up. At least he says he's an actor. Hanged if I know his name — who can believe him?" This subject’s ability to discuss his former profession coherently was impaired by his memory failure: I like Herbert and... what's the name of that Hungarian who died?... *Maytime.* One of my pupils is a conductor in Hollywood today — I can’t think of his name." This man could not say where he was nor what year it was.

In the gradual onset of senile chronic brain syndrome, memory
defect may be the only manifestation until overt impairment of behavior limits the subject's social competence.

_Dementia, moderate, Mrs. B.M. aged 89:_ This subject did not know the day, the month or the year. She did not know who the President was nor his predecessor. In conversation she repeated statements and anecdotes she had related only minutes earlier. She said that she had stopped going out, because she could not find her way back. In spite of the defect in memory and orientation she was content in home: “You've got to associate yourself with some organization that can take care of you.” She was modestly satisfied with her own emotional adjustment: “I have a happy faculty of putting things behind me which are apt to upset me.”

Accompanying the intellectual deterioration of senile psychosis, lability and shallowness of mood can occur together with disorganization of separate aspects of behavior. The affective disturbance may present more prominently than the cognitive impairment.

_Dementia, moderate, Mrs. S.H. aged 92:_ “My mother had black stockings, seems the blood got poisoned from it and she died (sad expression). I was a tomboy. My _heilige_ husband Jakob H – in our home was like the Kaiser is here. I never let any man come near me except Jakob H – ; I don’t bother with men, I don’t even talk to them” (angry expression). When asked if she was sociable, she described with accompanying outgoing gestures her frequent “_singen und tanzen._” She kissed the examiner. The admitting psychiatrist had labelled her hypomanic.

**B. Functional psychiatric disorder.** The other large group of mental disorders in the aged is not characterized by intellectual impairment. In these common “functional” psychiatric disorders gross structural degeneration of the brain is not considered etiological of the abnormal behavior manifestations. Instead, the psychiatric symptoms are similar to those seen in the emotional disorders of early life when there is no accompanying cerebral disease.

A survey (Hobson & Pemberton) has disclosed that varying degrees of manifest anxiety and depression were found in 25 per cent of men and 50 per cent of women in the general population. Whether subjective emotional discomfort comes to be regarded as constituting psychiatric disorders will vary with social and economic conditions, and the medical attitudes prevalent in the particular community. Of
persons over the age of 60 years admitted to one mental hospital (Roth, 1955), half had a disorder dominated by a depressive symptom complex; such affectively-disordered patients were diagnostically distinct from cases of organic brain deterioration, only two to three per cent exhibiting the clinical features of dementia.

Depressive reactions are common functional psychiatric disorders of old age. Also distinct clinically is systematized paranoid delusional psychosis, a persecutory syndrome occurring with relative preservation of the personality (senile paraphrenia). Another category of functional illness occurs in subjects who, having complained persistently since youth of mild psychoneurotic symptoms, in old age present with increase in severity of these functional symptoms. Clinical manifestations of functional psychiatric illness include: pathological degrees of anxiety, depression, restlessness, suicidal behavior, agitation, ideas of reference, paranoid delusions and hypochondriacal preoccupations. Various symptom complexes are differentiated in this age group; the lack of medical consensus in recognizing the various syndromes clinically, and the variable expression of behavior disturbance under differing cultural conditions, may explain the discrepant prevalence and hospitalization rates for old-age psychoses in different countries. There will probably be high agreement, however, among psychiatrists in clinically distinguishing the general group of functional psychiatric disorders from organic intellectual deterioration.

Mild functional illness, Mr. P.G. aged 80: Depressive reaction in a man of passive-dependent personality who drinks secretly, (was alcoholic) but within acceptable limits, had been divorced, and shows accusatory ideation directed at the environment. "Many times I got so drunk that I could hardly find my home — I was despondent and bored. I feel it here too. I feel I am the forgotten man, you see. I was a good uncle, you know. I have been here three years, and not one of them visited me. My son in California wrote, 'Pop, if you need anything let me know.' I wrote back I can use a Lincoln occasionally. In his reply he sent me a four cent stamp. For Christmas he sent me $10. You're forsaken by everything. Nobody cares if you're dead or alive. Most of our children would prefer you dead — they've got to donate something to the home. My son kicks in a $5 or $10 bill occasionally; then you can have a bottle with you but you don't have to advertise it. Years ago in Philadelphia I bought four
quarts at a time. I felt life wasn't worth living. I couldn't sleep. I worked until I was 78, then people said, 'You're too old.' Here I have some roommates I don't talk to. I was sleeping here with a man for six months — every time I see him I'd like to kick his face in. One night the temperature was 90 and I opened the window only four inches. He dressed up and went to call the police, saying, 'There's a man trying to kill me.' I came here because I was living in poor furnished rooms. I ate here and there. There was not even a bathtub. Once I was miserable and bought a bottle of sleeping pills. The cleaning woman found half my body outside the clothes closet — nobody knew how long I was lying there. I went to St. Vincent's for 10 days, and then I was brought back to the same dirty room. Then I got sick and after five months in a nursing home I overheard a social worker tell another patient about this home. Here I keep well out of the way of the staff. I once asked the cashier about my money... how can you approach them?"

In some cases doubt arises during the psychiatric examination whether a subject's anxiety or other painful effect should be classified as comprising mental disorder, or whether the psychic discomfort is within the limits of acceptable reaction to the altered circumstances of being in an institution.

*Mild depressive reaction* (doubtful), *duodenal ulcer*, Mrs. G.S. aged 81: A determined, energetic old woman discomforted by her feelings of resentment, anxiety, frequent weeping and fears on behalf of her relatives. "I would have liked to stay all my life with my children. I can't blame them for leaving me alone, can I? I have a lot of heartache. Like my niece — that bothers me all the time, it makes me cry (weeping). I brought her to this country. That is the biggest thing that aggravates me. My niece always promised that when I'm in need she'll help me. And she didn't. When her husband died he gave over $400,000 to charities. That aggravates me. I always used to worry a lot for everything. Now I am nervous all day whether my sons are happy and well. I'm always afraid their children will hurt themselves. When my small grandson was circumcised I thought..."
the man had done him harm — I took a couple of days before I settled down."

In the present investigation, the psychiatric examination was used to differentiate four groups of subjects: organic brain syndrome (dementia), functional psychiatric disorder, dementia with functional psychiatric disorder, and a normal group.

Social Adjustment in An Old Age Home

The purpose of this investigation is to discover whether psychiatric disturbance is related to social maladjustment. In order to study social adjustment it is necessary to study individuals in a milieu in which the environment impinging upon the individual is relatively constant for all.

A home for aged is an insulated community in which the norms and social structure can be assessed. Thus, it is possible to determine the demands made upon the residents of the home and to establish minimal criteria for adjustment to a social environment.

In the course of previous research conducted in this home (Tec & Granick, and Granick & Nahemow) criteria of social adjustment were established. Adjustment consists of three processes:

1. Integration, which is operationally defined as the number of interpersonal relationships in which the individual is engaged in the home. It is measured by obtaining an estimate of the number of clubs to which he belongs, the number of activities which he attends, the number of voluntary jobs for which he is responsible and the number of informal groups of which he is a member.

2. Conformity is defined as behavior in accordance with social norms. It is known from previous research in the home that two norms exist among the residents. One is that the rules and regulations of the home are to be obeyed, the other is that a "good" resident is one who is not bothersome. In order to assess degree of conformity, the following kinds of questions are asked: "Have you ever gone out of the home without asking permission or getting a pass?" and "Do you with other residents when they are wrong?" Twenty such questions constitute the conformity index.

3. Evaluation is defined as the degree to which the resident likes the home and feels a part of it. In order to ascertain whether the residents uses the home for aged as a reference group, questions are asked requiring evaluations of specific aspects of life in the home, such as food, staff, and activities.
Description of the Sample

The mental status of a group of 50 elderly people in an aged home was ascertained by psychiatric examination. They were the surviving members of 100 persons, already studied extensively (Granick & Nahemow), who had been consecutive admissions to the home between May, 1958 and August, 1959. The original sample consisted of 28 men and 72 women ranging in age from 61 to 95 years. (It is the policy of this home to exclude known psychotic persons from admission.) Fifty subjects had died or were in the infirmary. The survivors form the present sample, 11 men and 39 women ranging in age from 63 to 92 years. These remaining 50 individuals may differ psychiatrically from the members of the original sample not surviving. All accompanying sociological data to be presented is pertinent only to the 50 persons examined clinically for psychiatric impairment.

The Psychiatric Examination

Clinical assessment of each subject took about 45 minutes. The examination covered uniform ground for each subject and a previously devised protocol form was used for recording interview data. A life history was obtained and the mental status ascertained. In addition to the interview seeking to diagnose in each subject the presence or absence of chronic organic brain syndrome and functional psychiatric disorder, a number of (1) scales were administered, and (2) items checked.

(1) Scales: (a) The Mental Status Questionnaire (Kahn et al., 1960) was given to each subject; this consists of 10 questions, dealing with orientation in time and place and including information items, e.g., “Who is the president of U.S.A.?” The questionnaire tests the subject’s cognitive abilities. (b) The Anomie Scale (Robert & Rokeach, 1956) is one scale of personal adjustment; it consists of five items exploring the individual’s sense of relatedness to the social environment, e.g., “It is hardly fair to bring children into the world, the way things look for the future.” Agreement with an item obtains three points, while if the subject disagrees he is given one point; an uncertain response scores two points. A high total anomie score, therefore, indicates that the subject agrees strongly with statements expressing alienation from and mistrust of the social environment.

(2) Items Checked: The protocol form included specific queries about (a) Sleeping behavior; whether the subject reported sleeping well, fairly well, or poorly; (b) Mood; whether the subject reported
feeling in good spirits most of the time or seldom; (c) Tendency to worry, whether little or a fair amount. Such items evaluate more subjectively individual reactions than e.g. the Anomie Scale, and give the subject an opportunity to convey directly his sense of innate wellbeing or lack of it.

Assessment of Social Adjustment

Interviews were conducted independently of the psychiatric examination to assess the degree to which each person was adjusted to the home. A standard interview was used to measure the three components of social adjustment, which are integration, conformity and evaluations of the home.

Psychiatric Rating Scale

During the early stage of this research, two years prior to the clinical examination of each person by a psychiatrist, a rating scale was used by an interviewer for the individuals in this sample. This rating scale consists of 67 items designed to rate behavior which is elicited during an interview. It contains such items as: (1) exhibits concern for social approval of the interviewer; (2) shows great variation in rate of speech during the interview. The interviewer indicates whether or not the behavior was characteristic of the individual. This rating scale was utilized during the first stages of the longitudinal study. Thus, the information was obtained approximately two years prior to the clinical interview. This information was obtained independently of the psychiatric examination and the results were not available to the psychiatrist in the course of his interviews.

Analysis of the Data

The data on the 50 psychiatric evaluation protocols were punched on IBM cards, together with a number of social indices measured independently by the sociologist and the social psychologist.

Psychiatric Categories Used

The procedure of the psychiatric examination subdivided the sample into four diagnostic categories.

1. Chronic brain syndrome (dementia). Subjects judged clinically to show evidence of intellectual deterioration totaled 15. Of these, five subjects had symptoms of functional psychiatric disorder in addition to the dementing process.
Dementia with functional psychiatric disorder, Miss R.M. aged 67: Sits in a wheelchair, confused and agitated (personality disorder). Cannot recall simple information she is asked to retain. Repeatedly insists she should be playing Bingo. Anxiously pulls up her dress, perseveratively exposing her thighs. "God dammit, I'm bare, doctor. My God, I'm bare doctor."

Does not know what year she was born, the place she is now, the name of the President, etc. Information subsequent to interview: Never worked, did not marry, was dependent during adulthood on a sister: nursing staff report that subject disturbs the whole ward, is noisy and rude.

2. Functional psychiatric illness. Eighteen of the 50 subjects were diagnosed by psychiatric interview as suffering from functional psychiatric disorder. (Five of these had chronic brain syndrome in addition). The total number of disorders identified exceed 18, a number of subjects presenting more than one clinical syndrome.

(a) Minor functional illness (psychoneurosis). Twelve subjects presented psychoneurotic syndromes (mild in two cases, not disabling the subjects to a gross degree). Half these subjects showed predominantly depressive, the other half mainly anxiety symptoms.

Of the six subjects with minor depressive syndrome, some were severely disabled by the abnormal mood disorder.

Psychoneurosis (depressive and anxiety symptoms) with combined chronic brain syndrome, Mrs. E. E. aged 72: "I just eat and stay in my room and don't bother with anybody. I think they're sadistic. I'm only friendly with Miss Patterson, the nurse, a very lovely person. I think I did come here under my own steam. I felt let down. I was assigned to a room for three. How cruel some people can be! I think I made bad friends. One was crippled, just so high, tight underwear — her legs were long, like an animal. Something I'd never seen before! Especially on waking — I woke early. She raised herself on her elbow: 'Do you know Judy?' It took me a long time before I realized she was saying 'duty.' Then a woman reported me for saying 'Mind your own business.' With the result they put me in a dormitory with the most awful people. As if I was taken to another country and thrown among foreigners. I'm
so unhappy about it. I’m most unhappy here. I hate every minute of it.

“I have always felt keenly when things hurt me. I wasn’t my mother’s favorite. I was my father’s favorite, which she resented. She used to beat me. She kicked me. Many’s the time I’d fall to the floor. She injured my pelvis, distorted it, with the result that I could never have children. She wanted to take me out of public school before I graduated. Then I was incompatible with my husband. He was much older than I was. I divorced him and went back to work. I don’t know why I ever married him. My home life wasn’t happy — maybe that’s the reason why I thought, ‘Well, I’ll get married’.” Tense and anxious during the interview, she attended poorly to tests: she could not say what the place was where she was, the day, month or year, what year she was born, or who the President is. She is repetitive. The institution psychiatrist emphasize in his notes her fearfulness, disdain of other residents, anxiety symptoms and the poor social contact she made.

In six subjects anxiety symptoms predominated. In addition, one presented with hostile criticisms of persons in the environment. Psychoneurosis, Mrs. R.S. aged 88: “I’m very nervous and sensitive. I shake from top to bottom. I worry a lot, about myself and the people I’ve got to be intimate with here. They talk nonsense. We have one up here that’s a real psychologist, she analyzes everybody. She keeps the radio on — oh, that bad static! I’ll go out of my mind! She’s the nurse’s pet. Other people give the nurse $5 — I’m not doing it! The nurse treats me like a mental case. I’m having a lot of trouble with her. She claimed I hit a woman who handed me my purse. Later a woman who slapped me fell and broke her hip. It had nothing to do with me. The nurse says I’m a trouble-maker. I told her she’s a disgrace to her profession. What must I do? I went to my social worker to tell her I must have air in summer. I had a heart attack, and my roommate closes the windows. I feel I must ask the doctor to take my pulse, but I’m afraid then he’ll stop me going on the ride.”

One markedly anxious subject exerted herself in repetitive activity and thinking:
Walton et al.: Psychiatric Illness

Psychoneurosis, Mrs. I.C. aged 72: "I survived everything and I'm sorry (weeping). I sort of feel there's no place for me. I'm tired of everything. You get depressed. I have my knitting and my sewing, but no companions. When I'm sitting by myself I keep trying to fish out why this had to happen to me. You can't - so you leave it alone. I call the children every day. I worry if they're sick. Both my daughters-in-law had hysterectomies. I've got no grandchildren - if I did, I'd have had a job: baby sitter. Now I'm nothing. I hate this home. I go to Broadway and sit on benches and talk to strangers in Strauss Park. It's a miserable situation."

(b) Personality Disorder. Six subjects were diagnosed as psychiatrically disordered on the basis of abnormal personality organization of long duration, their social maladjustment long preceding their entry into the home.

1. Immature mother-dependent spinster, never worked, now with chronic brain syndrome.

2. Inadequate personality, probably schizophrenic illness during adolescence, cared-for during manhood by parents; now hyperactive, argumentative, inquisitive and quarrelsome.

3. Timid, withdrawn, sensitive, anxious woman, expressing much fear of contact with other people who may be ill, apprehensive at being interviewed and of information leaking.

4. Younger, more submissive partner in a pair of sisters showing folie a deux, a spinster who had never worked, suspicious and resentful.

5. Woman with obsessional personality, perfectionistic, preoccupied during conversation with the evils of the world, can't fall asleep for "taking an inventory of the day gone by and thinking of what is going to be." Resents interview, claiming she is "not crazy." Much hospitalized during lifetime.

6. Passive dependent spinster, socially inadequate, borderline intelligence.

(c) Major functional disorder. (psychosis). Three subjects were severely disordered with psychotic symptomatology. Of these, one had a paranoid systematized delusional psychosis (the sister who shared her room participating in the subject's false beliefs). Two had depressive psychoses.
Depressive Psychosis, Mrs. M.R. aged 80: "I don't like to live in a home. My daughter married a non-Jew, he talked her into sending me here. (Weeps). I'm very sorry I came. The food is terrible. We don't count any more, only money! The social worker is so mean. If anybody says anything against the home they're throwing them out. People are very mean: they like to dress up, pink and rouge and have permanents. Some have it good here, who just want to eat and wear rags. You should put on necklace, earrings! I'm not afraid of dying, only of suffering. When we are old and sick we shouldn't live. I'm not happy (sobbing). I hate people, I hate them! I just think, another day - and another day! I pray to God to take me quick. They don't care - if you don't like it, leave! Then they hate you. I'm dumb, stupid. I want two days to read the paper, because I'm depressed. I am going to go out and kill myself outside the home, because I don't want to make trouble. If you tell them, I'll kill myself. They're out for money. They like you to be sick. The Welfare pays more then. Dirty Jewish business! (Sobbing). They like you to break your legs."

(d) Psychophysiological (psychomatic disorder. One subject had a duodenal ulcer, and another asthma; both, in addition to this somatic symptomatology, were diagnosed on the basis of psycho-neurotic symptoms as showing minor functional illness.

3. Psychiatric normality. The psychiatric examination is an investigative method to detect specific patterns of abnormal behavior. Normality is presumed in a subject when morbid symptom patterns are not discovered: there are no positive clinical signs of mental health. Furthermore, an aged home is a terminal institution, inescapably different from the wider social environment; "normal" adjustment in the home does not necessarily indicate that a subject would be adjusted outside.

Accepting these reservations, the behavior presented by the 22 subjects in whom no signs of psychiatric morbidity were detected merits some detailed description. A number of subjects not only coped with the difficulties imposed by living in the institution, but experienced apparent personal contentment in addition. Of the twenty-two subjects who were evaluated as free of psychiatric disorder, some were satisfied and others were dissatisfied with the institutional environment.
Mrs. L.K. aged 69: "The children would not be happy if I was with them. I want to keep away from the children, and I wasn't healthy any more. The whole home is my friends. It's no use to be nasty. If I don't like it, I try to go away. I'm going away and we're friends. A lot of old aged people are cranks and they're not normal in the mind, but you've got to live with them. When I'm well I dance and run around people. For every one I got a joke and smile. I make myself useful, I make rugs. They like me a lot."

Such capacity for evading or controlling troublesome fellow-residents is frequently emphasized by subjects as a necessity for emotional comfort. One subject (Mrs. D.L. aged 72) expressed this, "I like humans - but some you avoid." The staff also appear to perceive this component of adapted behavior; the social worker commented about Mrs. J.K. aged 78, "Nice and easy adjustment. Does beautiful handwork. Stands up for her rights. No one steps on her."

Some contented subjects are evaluated by the staff as defective personalities. In the records, consulted after the psychiatric examination, the clinical psychologist attached to the home had described Mr. E.K. aged 84: "Never matured. Dependent. Obsessive compulsive. Women represent a devouring force to him." However, this subject's overt behavior during examination indicated no subjective discomfort in home conditions:

"I think the world is getting real nice for the average person. I know from experience. I do what I can to help the institution: library duty, information table duty, taking people to doctors and to hospital. At one time I had many professional men as my friends - I was known to be able to make friends and keep them."

The impression at times obtained is that many aged persons adapting without signs of mental disorder in the home, if seen earlier in life, would have been considered poor psychiatric prospects. Protracted parental dependency, failure to marry, social isolation, obsessional character traits, suspicious attitudes, social prejudice are all to be found conspicuously in "normal" subjects who convey that the conditions of the home are well suited to them. The individual may convey only relatively mild criticism of the environment, compatible with a symptom-free adaptation.

Mr. A.Z. aged 77, divorced when 47, in reply to the
question when tested for orientation, "What is this place?", replied "Germany concentration camp." "I like to choose my friends. I'd rather be alone. And I can't eat their food, it makes me nauseous – I'd rather go hungry. But there's nothing for me to worry about. I can explain my feelings without causing trouble. I don't want to be a yeser or a noer. I enjoy a conversation – you can learn from a child." He described his wheelchair as "my Cadillac"; another subject spoke with admiration and affection for him. Mrs. S.H. aged 81: "I'm particular whom I associate with. Tell me who your friends are and I'll tell you who you are. I don't like yakety. I never think of anything – if it comes, it comes. I try to get away from troubles. I wouldn't let anyone step on me, that I wouldn't. Nobody gets a chance on me."

Resistance against discomforting contacts with others may take the form of active hostile disparagement, and still not impair apparently normal mental status.

Mrs. R.R. aged 82: "I am an American. I don't associate with many here – they're most Germans, and they call me the Proud American. In German. I worry over nothing – you can't change nature. The only thing that worries me is that I hope when I get ill I don't have to suffer."

In this research, as in the community sample studied by Busse the largest single group of subjects was diagnosed as having no disorder. (1961). Table 1 shows the per cent with diagnoses of mental normality, chronic brain syndrome (dementia), functional psychiatric disorder, and chronic brain syndrome with functional disorder in the two samples.

The preponderance in these two samples of functional disorders over organic brain syndrome (senile dementia and arteriosclerotic psychoses) is regularly found also among elderly patients admitted to mental hospitals. Affective depressive disorders have been found present in 50 per cent of first admissions of aged persons to mental hospitals in one study (Roth, 1955), and in 45 per cent in another (Gibson, 1961). A recent investigation confirmed this preponderance of depressive disorder over dementia. (Kidd, 1962a)

Results

Table 2 shows the relationship between dementia and functional psychiatric disorder. No correlation was found between them.
### Table 1
**Psychiatric Status of Elderly Individuals in Two Populations of the Aged**

<table>
<thead>
<tr>
<th>Psychiatric status</th>
<th>Home for aged and infirm Hebrews, N.Y.</th>
<th>Busse's sample at Duke, N.C.*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Per cent</td>
<td>Per cent</td>
</tr>
<tr>
<td>Chronic brain syndrome (Dementia)</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>Functional psychiatric disorder</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>Chronic brain syndrome with functional psychiatric disorder</td>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>Normal mentally</td>
<td>44</td>
<td>40</td>
</tr>
<tr>
<td>Population size</td>
<td>50</td>
<td>222</td>
</tr>
</tbody>
</table>


### Table 2
**Distribution of Individuals in Diagnostic Categories**

<table>
<thead>
<tr>
<th>Functional psychiatric disorder</th>
<th>Dementia present</th>
<th>Dementia absent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>10</td>
<td>22</td>
<td>32</td>
</tr>
<tr>
<td>Present</td>
<td>5</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>35</td>
<td>50</td>
</tr>
</tbody>
</table>

This finding supports the clinical and theoretical differentiation between the two major mental illness categories of old age.

**Table 3** shows the relation between scores on the Mental Status Questionnaire (MSQ) and the psychiatric status of the residents.

It is clear that the M.S.Q. differentiates those diagnosed as having chronic brain syndrome from those not so diagnosed. Looked at from another point of view, it may be said that the extent of this relationship supports the clinical diagnosis of dementia.

**Table 4** shows the relation between scores obtained on the
TABLE 3

RELATION BETWEEN PERFORMANCE ON THE MENTAL STATUS QUESTIONNAIRE (MSQ) AND PSYCHIATRIC STATUS

<table>
<thead>
<tr>
<th>Psychiatric status</th>
<th>Performance on MSQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No error</td>
</tr>
<tr>
<td>Chronic brain syndrome with functional psychiatric disorder</td>
<td>0</td>
</tr>
<tr>
<td>Chronic brain syndrome (Dementia)</td>
<td>1</td>
</tr>
<tr>
<td>Functional psychiatric disorder</td>
<td>11</td>
</tr>
<tr>
<td>Normal mentally</td>
<td>19</td>
</tr>
</tbody>
</table>

*Chi Square for difference between the mentally normal and those with chronic brain syndrome equals 12.5; \( p < .01 \).

Psychiatric Rating Scale and the psychiatric status of the residents which was ascertained two years later.

Residents diagnosed as suffering from functional psychiatric disorder.

TABLE 4

RELATION BETWEEN PERFORMANCE ON THE PSYCHIATRIC RATING SCALE AND PSYCHIATRIC STATUS OF RESIDENTS

<table>
<thead>
<tr>
<th>Psychiatric status</th>
<th>Ratings on psychiatric rating scale</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Symptoms present (above median)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Symptoms absent (below median)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic brain syndrome with functional psychiatric disorder</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Chronic brain syndrome (dementia)</td>
<td>5</td>
<td>9†</td>
</tr>
<tr>
<td>Functional psychiatric disorder</td>
<td>8</td>
<td>11‡</td>
</tr>
<tr>
<td>Normal mentally</td>
<td>7</td>
<td>21</td>
</tr>
</tbody>
</table>

*The psychiatric rating scale was not administered to 5 residents in this sample.
†Chi Square for difference between the mentally normal and those with chronic brain syndrome = .54 \( p \) is not significant.
‡Chi Square for difference between the mentally normal and those with Functional Mental Disorder = 3.06 \( p < .10 \).
disorders had been found to have more symptoms on a psychiatric rating scale, than those who were diagnosed normal even two years before the psychiatric examination was conducted. This was not found to be true for individuals suffering from chronic brain syndrome.

Relation between Psychiatric Status and Measures of Social Adjustment

Table 5 shows the relation between the psychiatric status of the residents and their evaluations of the home.

Residents with functional psychiatric disorder are characterized by their dislike of the home. Three times as many of the residents suffering from functional mental disorder dislike many aspects of the home. Of the group of normal residents, the majority are favorably disposed toward the home. Of those residents suffering from chronic brain syndrome, the majority were also favorably disposed toward the home.

Table 6 shows the relation between psychiatric status of the residents and the extent to which they conform to norms.

Residents suffering from functional mental disorder are nonconformists or deviants. These appear to be the people who become management problems in the home. Those with chronic brain syndrome, on the other hand, are not deviants. They are at least as conforming as those who are normal.

Table 5

Relation Between Psychiatric Status of the Residents and Their Evaluations of the Home

<table>
<thead>
<tr>
<th>Psychiatric Status</th>
<th>Evaluations of the Home</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive (above median)</td>
</tr>
<tr>
<td>Chronic brain syndrome with functional disorder</td>
<td>3</td>
</tr>
<tr>
<td>Chronic brain syndrome</td>
<td>7</td>
</tr>
<tr>
<td>Functional psychiatric disorder</td>
<td>3</td>
</tr>
<tr>
<td>Normal mentally</td>
<td>14</td>
</tr>
</tbody>
</table>

*Chi Square for difference between the mentally normal and those with chronic brain syndrome = .003; p is not significant.
†Chi Square for difference between the mentally normal and those with functional psychiatric disorder = 3.88; p < .05.
TABLE 6
RELATION BETWEEN PSYCHIATRIC STATUS OF THE RESIDENTS AND CONFORMITY TO THE NORMS OF THE HOME

<table>
<thead>
<tr>
<th>Psychiatric status</th>
<th>Conform (above median)</th>
<th>Deviate (below median)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic brain syndrome with functional psychiatric disorder</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Chronic brain syndrome</td>
<td>9</td>
<td>1</td>
<td>10*</td>
</tr>
<tr>
<td>Functional psychiatric disorder</td>
<td>2</td>
<td>11</td>
<td>13†</td>
</tr>
<tr>
<td>Normal mentally</td>
<td>13</td>
<td>9</td>
<td>22</td>
</tr>
</tbody>
</table>

*Chi Square for difference between the mentally normal and those with chronic brain syndrome = 1.79; p is not significant.
†Chi Square for difference between the mentally normal and those with functional psychiatric disorder = 5.27; p < .05.

TABLE 7 shows the relation between the psychiatric status of the residents and the extent to which they are integrated in the home.

Integration, unlike the other two aspects of adjustment, appears

TABLE 7
RELATION BETWEEN PSYCHIATRIC STATUS OF THE RESIDENTS AND INTEGRATION IN THE HOME

<table>
<thead>
<tr>
<th>Psychiatric Status</th>
<th>Integrated (above median)</th>
<th>Isolated (below median)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic brain syndrome with Functional psychiatric disorder</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Chronic brain syndrome (dementia)</td>
<td>4</td>
<td>6</td>
<td>10*</td>
</tr>
<tr>
<td>Functional psychiatric disorder</td>
<td>6</td>
<td>7</td>
<td>13†</td>
</tr>
<tr>
<td>Normal mentally</td>
<td>18</td>
<td>4</td>
<td>22</td>
</tr>
</tbody>
</table>

*Chi Square for difference between the mentally normal and those with chronic brain syndrome = 3.82 p < .05.
†Chi Square for difference between the mentally normal and those with Functional Mental Disorder = 3.31 p < .10.
to be equally difficult for both residents with chronic brain syndrome and those with functional psychiatric disorders. Many more of the normal residents are well integrated in the home than those with functional psychiatric disorders, chronic brain syndrome, or both. These differences cannot be explained in terms of age or physical status.

**TABLE 8** shows the relation between psychiatric status of the residents and scores on the Anomie Scale. This scale reflects feelings of alienation from and distrust of the environment. It is a scale of personal rather than social adjustment.

**TABLE 8**

<table>
<thead>
<tr>
<th>Psychiatric Status</th>
<th>Scores on Anomie Scale</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anomic (6-9)</td>
<td>Not anomic (0-5)</td>
</tr>
<tr>
<td>Chronic brain syndrome with functional psychiatric disorder</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Chronic brain syndrome</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Functional psychiatric disorder</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Normal mentally</td>
<td>10</td>
<td>12</td>
</tr>
</tbody>
</table>

* Chi Square for difference between the mentally normal and those with chronic brain syndrome = .82 p is not significant.
† Chi Square for difference between the mentally normal and those with functional mental disorder = .33 p is not significant.

The group of normal residents evidently feel less alienated than those with functional psychiatric disorder, chronic brain syndrome, or both types of disorder. However this tendency is not statistically significant. It appears that although residents with chronic brain syndrome were capable of making a good social adjustment to the home, this does not mean that they are in no way impaired psychologically. On the contrary, **TABLE 8** suggests that there is a tendency for them to feel cut off from the world and distrust the greater society.

**Discussion**

In the home for aged studied, the mental disorders found were of two types: dementia (senile and arteriosclerotic) and functional mental disorder. A distinction between these types of disorders in
the aged is of great clinical importance. Accurate diagnosis is known to be of direct relevance to outcome. Dementia is a progressive and untreatable disorder; functional illness, on the other hand, may be reversible with adequate psychiatric treatment. In the population of elderly individuals studied, it was possible to distinguish those individuals who had senile dementia from those with functional mental disorder or a combination of both disorders; moreover, no relation was found between the occurrence of senile dementia and various functional mental disorders.

The clinical importance of differentiating these mental disorders of the senium is established. When of sufficiently gross degree to require admission to mental hospital of the sufferer, the disorders are still distinct diagnostic entities. Investigations of different mental hospital populations show uniform findings (Gibson, 1961, Kidd, 1962a, Roth, 1955). Patients with affective (depressive) disorder have the greatest likelihood of being discharged, those who have senile dementia of dying, and those who have arteriosclerotic dementia of still being in the mental hospital a year later.

Old people unable to maintain themselves independently can be institutionalized in geriatric hospitals, mental hospitals or old age homes. About fifty per cent of old people going into mental hospitals have functional psychiatric illness, i.e. potentially reversible mental illness (Gibson, 1961, Kidd, 1962a, & Roth, 1960). The discharge rate of these mentally ill old people is high (Roth, 1960).

It has been found that geriatric and mental hospitals do not serve distinct populations of the elderly sick (Kidd, 1962b, & Kahn, Goldfarb, Pollack & Peck, 1961); the physically ill do not go to geriatric and the mentally ill to mental hospitals. On the contrary, as Townsend points out, misplacement often occurs (Townsend). The social consequences are considerable, there being urgent demand for institution facilities for the aged; moreover, in individual cases the institution in which an elderly person is misplaced may not be geared to provide diagnosis and treatment of potentially reversible conditions.

In the home for aged studied, those residents suffering from senile and arteriosclerotic dementia can be differentiated from those with functional psychiatric disorder on measurable social dimensions. Those with falling off of intellectual capacity can adjust to the institution environment with a sense of subjective satisfaction, and in accordance with the expectations of the institution milieu. But those with functional psychiatric disorder are personally miserable and, in addition to this personal discomfort, do not conform to social rules or live up to the expectations and requirements of
other residents. The evidence that intellectual impairment does not throw an individual out of step with the environment is some indication that mildly demented persons are suitable for care in homes for the aged.

All mentally disordered subjects are characterized by low participation in home activities, such as clubs, games, concerts and friendship groups. Nonparticipation should alert the staff that mental disorder may be present, and diagnostic and remedial measures called for.

Demented subjects, like the normal aged, evaluate the home positively and are positively identified with it. Those with functional disorder, on the other hand, evaluate the home negatively. They are the residents who have no positive affective bonds with their present environment.

Failure to conform is associated with functional psychiatric disorder. A difficult and malcontented resident requires to be assessed psychiatrically for the presence of such functional psychiatric disorder, which may be remediable. If functional illnesses are not treated appropriately, or are fatally misdiagnosed as due to brain pathology, social impairment can persist. Subjects with functional psychiatric disorder were so uniformly maladjusted on social dimensions that the conclusion appears warranted that a home for the aged fails to meet the needs of functionally disordered subjects.

Acknowledgments

We are grateful for the cooperation and interest offered so generously by Dr. Frederic Zeman, Dr. Alvin Goldfarb and other personnel of the Home for Aged and Infirm Hebrews of New York. We are also indebted to Dr. Joseph Zubin for his invaluable assistance and Dr. Eugene Burdock for the use of the Psychiatric Rating Scales and his aid and counsel in applying them.

References


GRANICK, RUTH & LUCILLE NAHEMOW. 1961. Preadmission isolation as a factor in adjustment to an old age home. In P. H. Hoch and J. Zubin (Eds.) Psychopathology of Aging. New York: Grune & Stratton,


TOWNSEND, P. Geriatric Institutions. (to be published.)
PROBLEMS AND PROSPECTS OF THE BIOMETRIC METHOD

Joseph Zubin

One of the most important gaps in the biometric approach is the lack of suitable techniques for integrating the data emanating from the various scientific models. The reason why the integration problem is so difficult stems from the fact that the population we are dealing with, schizophrenic patients, represents a universe which is defined by psychiatrists in a rather subjective, intuitive way. We are not even sure whether it is one population or a conglomerate of many. In order to place this problem in proper perspective, I must digress for a moment to discuss the taxonomic problem in general.

One of the earmarks of progress in a discipline is a gradual shift from impressionistic description to the more objective practice of qualitative categorization and eventually to the fully quantitative procedure of measurement. In the field of medicine, for example, gross observation gave rise to the classification of all diseases in which a wasting of tissues took place, as consumption. Castiglioni, in his History of Medicine (1947) points out that Antoine-Laurent Jesse Bayle, who lived in the latter part of the 18th century, recognized six kinds of consumption, of which the commonest would be correctly designated as tuberculosis today. In 1651, Willis complained that people were not distinguishing true consumption from other diseases and were calling everything consumption. As one advance after another was made—first by the discovery by Thomas Willis in 1670 that the urine of some of the consumptives was sweetish, then by the discovery of the leprosy bacillus by Hansen in 1880, and then by the discovery by Koch in 1882 of the tubercle bacillus, etc.—each group of patients with one or more unique characteristics in common, not found in the other subgroups, was finally identified as a separate entity. Thus, the disease entity, consumption, was eventually analyzed into diabetes, tuberculosis, cancer, etc. Now each of these diseases in turn is being scrutinized for further subdivision.

The classification which brought modern nosology into being began in the 18th century soon after the ancient notion that all disease was caused by disturbances of the four bodily humors was given up. In psychiatry, the term insanity still lingers on as a vestige of the earlier unitary conception of mental disease. It was not until late in the 19th century that the modern nosology of separate mental disorders was evolved by Emil Kraepelin.

*This study was supported in part by grant M-1541 from the National Institute of Mental Health.
Kraepelin's approach was clinical, based on a concatenation of factors observed in the behavior of the patient. Although psychologists have tried to develop tests to provide objective qualitative or quantitative criteria for the different diagnostic categories, not much success has been attained thus far.

One of the disease entities which Kraepelin enunciated was dementia praecox, which Bleuler later broadened into schizophrenia. In the earlier literature, while Kraepelin's diagnostic procedures were adhered to, much more regularity was observed in such characteristics of schizophrenia as incidence rates, release rates, etc. With the broadening of the base by Bleuler, the earlier consistencies seemed to disappear. Perhaps Bleuler's attempt was a step in the wrong direction—apparently science proceeds by narrowing rather than widening concepts.

Kraepelin had hoped that by studying large populations he could discover the incipient mentally ill through the use of psychological tests. He wrote as follows:

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 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.)

If we follow Kraepelin's suggestion, we would use such well-identified groups as registrants for military draft, random samples
of birth registers, or entire populations for psychological investigation, and pick off the deviant individuals for scrutiny. At the present time, such plans are beyond us. Meantime, we are reduced to using the available diagnoses as starting points, applying a variety of psychological and other tests and using statistical technique for determining the subgroups that exist.

If schizophrenia consists of several subgroups, the problem arises how to determine the natural lines of cleavage in this disorder so as to find the underlying subgroups.

While there are techniques available for contrasting samples from two different populations on a variety of tests, viz., Fisher's discriminant function and Hotelling's $T^2$, there are no comparable techniques, to my knowledge, for dealing with the problem of separating a population into subgroups on more than one variable. The univariate problem was solved by Pearson in 1894, by the method of moments. Since there are five parameters to be determined, i.e., the two means, two variances and the proportion of mixture, five moments are sufficient to determine these parameters. The solution depends on a suitably chosen root of a ninth-degree equation. A solution based on the method of maximum likelihood also exists in the univariate case, and we anticipate the appearance of a maximum likelihood solution to the multivariate problem in the not too distant future. Maximum likelihood deals only with estimation. However, since we are concerned more with the question of classification than estimation, the method proposed below, in spite of its being less efficient than that of maximum likelihood, yet has the advantage of not requiring that the number of subgroups be specified a priori.

In summary, one of the acute problems now facing the research biometrician in the field of psychopathology arises from the fact that most diagnoses are based on subjective judgments whose reliability is often not very high. Consequently any particular group of patients who carry a specific diagnostic label is often found to contain individuals who differ from each other not only in level of performance on a given series of tests but also in the interrelationships among these performances, some individuals showing positive trends, while others show negative trends on the same set of variables. As a result, comparisons of one diagnostic group with another or with a normal control group often yield insignificant differences because of the high interindividual variability within these diagnostic groups. While this difficulty is not uncommon in medical, psychological, and social diagnoses, the frequent absence of any objective indicators for making diagnoses in psychopathology makes the heterogeneity in this field a particularly serious problem.
The statistical techniques for dealing with this problem are in a rather primitive state of development. While several methods are available for finding the best set of weights to apply to a series of variables in order to discriminate between two or more populations, these methods are not applicable to our problem since we do not have segregated populations to begin with. Instead we must find the natural lines of cleavage within the diagnosed group so as to separate it into its natural subgroups.

As a first step in this direction we provided the method of likemindedness analysis (Zubin, 1938a, 1938b), which is suitable for finding the individuals who are like-minded or like-structured on such qualitative variables as response to dichotomous items. This method has already proved its usefulness in a variety of situations and more recently has been exploited by McQuitty (1954, 1960) in several analyses.

In dealing with quantitative data, the like-mindedness method cannot be applied directly unless the quantitative variable can be dichotomized or trichotomized without loss of too much information. If loss of information is to be avoided, a new method suitable to quantitative data must be found.

In order to develop a method applicable to continuous variables, we standardized each score and determined the profile of each patient on these standard scores. Each patient’s profile was then matched with the profile of every other patient and patients who resembled each other in level of profile as well as in shape were grouped together. Certain statistical criteria for adequacy of matching were developed and utilized in this grouping. While this method is still in its early stages, it seems to yield subgroups that relate to such criteria as outcome and test performance.

The future of biometric research will depend on the degree of acceptance that objective quantitative methods will receive and on the type of training that can be provided for this field. Because the assessment of mental patients is now largely in the hands of clinically trained men, progress in providing objective measures will be slow. Nevertheless, the advance already achieved through measurement, as demonstrated in the previous papers, lead one to hope that in the long run, objective measures will gain acceptance. Perhaps the greatest stumbling block in the way of progress is the lack of proper training facilities for the biometricians of the future. To really make progress in this field, it is necessary to find individuals who, though steeped in their own discipline, are nevertheless conversant enough with progress in adjacent fields to permit the new discoveries to influence their work. This does not performe
require an interdisciplinary research by individuals cooperating on the same problem. It requires simply having people from different disciplines, each concentrating on his own area, working side by side with those in other areas, so that benefits accruing from informal contacts will be available. Just how to provide such facilities is a problem of the first magnitude — but, acquaintance of the workers in this field with the entire spectrum of psychopathology seems essential to progress in biometric research.

An example of the kind of problem on which interdisciplinary cooperation has been utilized in biometric research is the problem of prognosis. It is not the only problem toward which the various disciplines and techniques can be directed, but it will serve to illustrate one dimension along which different methods can be aligned in the biometric approach to psychopathology. I have selected prognosis because prediction is a primary problem of science. Furthermore, we have gone as far as we can go with diagnosis, and it is time prognosis came to the rescue. One need only to look into a series of case histories to realize to what a low state prognosis has dropped. The most frequent entry is "guarded." The drop in interest in diagnosis and nomenclature is at least in part due to the fact that we have shifted our release policies so that now almost two-thirds of the patients leave the hospital regardless of what therapy or what treatment they have received. If this tremendous outflow of patients occurs anyhow, of what purpose is diagnosis? Furthermore, it looks as if no matter what the diagnosis is, or what the therapy is, in a five-year followup the outcome seems to be independent of the nature of the therapy and to a great extent depends on the nature of the patient. There are at least the following levels of information which may be utilized in making prognoses of outcome of mental illness: paranatal, premorbid, morbid, course of illness, and finally, the criterion — outcome on follow-up. We have attempted to sample each of these levels in order to obtain information which may be useful in making predictions. The first step was to complete a literature survey of the characteristics in these various levels which have been claimed to have prognostic import. A literature consisting of some one thousand articles was surveyed and about a hundred or so traits were discovered that had been claimed to be related to outcome in schizophrenia. This has been reported elsewhere in summary form (Zubin, et al., 1961).

Our initial approach to prognosis (Zubin, et al., 1961) tapped the premorbid area through interviews, the morbid area through psychological tests of the psychomotor, perceptual, and conceptual variety, and the course of illness through ward behavior inventories.
The crudeness of our initial tools and the revolution in release policies which occurred in the middle of our investigation prevented our initial results from emerging to their best advantage. Nevertheless, correlations of the order of 0.40 between test performance and outcome was the rule rather than the exception and certain new sub-groupings of schizophrrenia were developed based on our test results. It became quite clear that the social-cultural milieus and premorbid characteristics of our patients loomed larger than our psychological tests in making predictions of outcome. Since they are reflections of deviant behavior, rather than the behavior itself, they are not as good indicators of deviation as observations based more directly on the behavior itself.

It might be of interest to inquire why the social-cultural, premorbid characteristics of our patients are so important in prognosis. Anyone acquainted with the current scene must admit perforce that we detect mental disorder not on the basis of anything else, but deviation in behavior from expected social-cultural norms. While it is true that the inner suffering of the patient is also important, this in itself may not always bring the patient to attention unless it affects his overt behavior, nor can we regard this inner suffering as the primary cause of the illness. Social-cultural factors not only determine the detection and emergence of the person as mentally ill but they also determine the kind of treatment he gets and the chances of his return to the community. For this reason, most of the psychological tests used in the clinic measure conceptual responses, which are highly dependent on social-cultural norms.

It is of course important not to lose sight of the biological underpinnings of mental disorder, or some of the very early environmental influences (those which can not now be distinguished from the hereditary) which may propel an individual in the direction of mental illness. How can one separate the biological factors from the social-cultural factors in the emergence of mental illness when, as a matter of fact, it is the social-cultural factors that permit the illness to emerge and be detected. One way of answering this question is to conduct cross cultural studies in mental disorder. Presumably, those factors leading toward the emergence of mental illness which are culturally determined should vary considerably from culture to culture, but those which make the person vulnerable in the psychophysiological sense should be relatively independent of culture, unless they themselves result from social-cultural pressures. The search for these culture-fair indices of mental disorder, namely, some neurophysiological responses which may characterize the vulnerable individual regardless of whether the social-cultural
pressures are enough to elicit mental illness, is indeed a worthwhile search. Perhaps some of the indices which have been discussed in this monograph, namely, pupillographic responses, cross-modal reaction time, and delayed auditory feedback may provide us with measures which are culture-fair.

Thus, biometrics can fruitfully integrate the techniques of various disciplines and focus them on a single problem area in psychopathology. This volume presented and discussed some of the methods being developed and sharpened in connection with various problems. However, as these techniques are perfected, it is hoped that they will be brought together again and again to complement one another in united research efforts.

References
