EVENT-RELATED POTENTIAL AND BEHAVIORAL METHODOLOGY IN PSYCHIATRIC RESEARCH

JOSEPH ZUBIN*, SAMUEL SUTTON**, 1, AND STUART R. STEINHAUER*

*Biometrics Research, 151R, VA Medical Center, Highland Drive, Pittsburgh, PA 15206, and University of Pittsburgh School of Medicine; **Department of Psychophysiology, New York State Psychiatric Institute, 722 W. 188 St., New York, NY 10032, and Columbia University

INTRODUCTION

The more data-oriented studies to which this chapter serves as a prelude contain a mine of innovative findings dealing with neurophysiological aspects of psychiatric conditions. Our function here is to provide an introduction which attempts to sketch out, in relation to psychiatric research, the role of event-related potentials (ERPs) as a bridge between brain function and behavior. To this end we shall present our own loose framework -- our strategy and tactics -- in the hope that this can serve as a kind of structure into which specific findings can be integrated.

Ever since the birth of the concept of dementia praecox and its subsequent transformation into schizophrenia, the underlying basis for determining the presence of this disorder has inched in the contrast between a sound sensorium and an unsound mind, i.e., the presence of thought disorder, ambivalence, loose association, etc., in a person whose sensory and perceptual capacities are intact. It is as if the engine were intact but its functioning was at fault. This dictum went unchallenged for almost a century -- it initially dissuaded experimental psychologists from examining the psychophysiological, sensory and perceptual functioning of schizophrenics, and it permitted clinicians to search almost exclusively in the domain of broad conceptual behavior for the tell-tale marks of schizophrenia.

Clinical psychological tests have for a long time been used for the differentiation of schizophrenics from other mental patients and normals, but the validity as well as the reliability of these tests has never been well established. In a survey of clinical psychological tests for diagnosis and prognosis carried out in the early 50's, we were forced to give up any hope that the available projective and personality tests would be of value (Windle, 1952; Zubin et al., 1953).

1 Dr. Samuel Sutton died suddenly on March 29, 1986, after putting the last touches to this paper. He was the discoverer of P300 in its various manifestations, a pioneer in the application of basic research methodology to the problems of psychopathology covering the spectrum ranging from physiological to sensory, perceptual, and cognitive functioning in the mentally ill and normal individuals. He was formerly the Deputy Director of the Biometrics Research Unit in the Department of Mental Hygiene of the State of New York and, until his death, Chief of the Department of Psychophysiology of the New York State Psychiatric Institute and Professor of Medical Psychology in the Department of Psychiatry, Columbia University. This tutorial chapter owes much to his clear, incisive delineation of the role of ERPs in psychopathology; his colleagues, as well as his students, express their great loss at his passing. His co-authors would like to dedicate this chapter to the memory of a great teacher, and a warm-hearted colleague and friend.
The gross behaviors which are noted clinically in schizophrenics and which constitute the basis for the diagnosis of the disorder are so molar in character that it is futile to try to resolve them into their underpinnings relative to causal models. While persecutory ideas, for example, can be related to the patient's ecological niche — the Russian patient is chased by OGPU, the American patient by the FBI — it would be very difficult to form hypotheses with respect to the origin of persecutory ideas without making a more intensive examination of the single individual, his prior history, etc. To try to determine, for example, a neurophysiological basis for persecutory ideas is beyond our capacities.

Instead we turned from the clinically observed gross behaviors of schizophrenics to more classic categories of human responses to stimulation — such as sensory, perceptual, psychomotor, and conceptual — and we have tried to determine whether the pattern of response across these categories differentiates patients from normals. We developed a Mendeleev-like table classifying responses by category of stimulus (Zubin, 1968).

While crude global tests and techniques have always yielded differences between schizophrenics and normals, these differences could usually be attributed to such general factors as lack of attention, motivation, difficulty of task, etc. The more sophisticated and well-controlled the techniques, the less the differences — when adequate consideration is given to many of the non-specific sources of the poorer performance.

However, techniques have begun to emerge which seem to find differences between schizophrenic and normals, which are not likely to disappear under the most exacting controlled conditions. It has become quite clear that the sensory and perceptual functioning of the schizophrenic is not as free from deviations from the normal as had been thought. To be sure, the differences are not as dramatic as those found with the more global, uncontrolled techniques.

The finding of such sensory and perceptual differences became possible because of two developments. First, the global diagnoses of schizophrenia, depression and personality disorders, based upon free-floating interview techniques, have been replaced in many laboratories by more objective systematic interviews which support the clinical diagnoses with a dimensional profile that provides a more formal description of the patient's underlying psychopathology. Armed with this information, we can obtain more homogeneous subgroups which can be compared with other homogeneous subgroups for any laboratory techniques that are applied. For example, we can now contrast schizophrenics who show thought disorder with schizophrenics who do not, or with depressives who show thought disorder; or depressives who show high delusional loadings with schizophrenics and depressives who do not show such loadings, etc.

A second advance has come from progress in sensory and perceptual techniques which can utilize brief stimuli that elicit responses within the first 1,000 milliseconds following stimulation (Zubin & Kietzman, 1966). Such techniques, especially in conjunction with trial by trial own-control experimental designs, may be somewhat freer than the
older techniques from the influence of motivation, prior learning and experience. Among these are certain types of reaction time experiments, temporal energy integration, interaction of successive auditory or visual stimuli, to cite only a few. With such techniques, it is found that the sensorium of the schizophrenic is not as similar to the sensorium of the normal as had been thought. There are specific functions in which schizophrenics differ from normals as well as from non-schizophrenic psychiatric patients.

The utilization of such objective indicators should in turn ultimately be of assistance in the search for better classification of the mentally ill into more homogeneous groups. By aiding in the specification of subgroups that are as homogeneous as possible by means of both laboratory techniques and improved clinical diagnoses, we can contribute to the search for the etiology of mental disorders, for more specific therapies, and for preventive measures.

In the attempt to determine the possible etiological sources of schizophrenic behavior we have postulated the following classification of models: (1) ecological, (2) developmental, (3) learning, (4) genetic, (5) internal environment, (6) neuroanatomical, and (7) neurophysiological or brain function model (Zubin & Kietzman, 1966; Zubin et al., 1985).

Let us first indicate our conviction that not one of these seven models by itself can provide the necessary and sufficient basis for a theory of schizophrenia. It is rather the interaction of these causal factors that may be regarded as both necessary and sufficient. Nevertheless, because of our abysmal ignorance regarding the causes of schizophrenia, it is strategic to try to determine how far each of these models can take us by itself, before we invoke its interactions with the others. In this chapter, we are concerned only with the neurophysiological model.

I THE NEUROPHYSIOLOGICAL MODEL

The rationale for applying a neurophysiological approach to mental disorders is the hope of finding some kind of indicator of differences in central nervous system functioning between different groups of patients, or between patients in a given category and normals. This hope is linked to the overall goals of obtaining objective measures which may assist with clinical and diagnostic problems, hopefully cross-culturally, as well as clues which may be relevant to the question of etiology for that illness. Kety (1959), in commenting about the subtlety and complexity of undertaking biochemical research in mental illness, noted that the kinds of artifacts which may arise in atheoretically oriented biochemical analyses of blood, urine and other body tissues make it highly desirable that some clues regarding potentially significant research directions be supplied to the biochemist from other fields -- particularly the field of psychology.

In recent decades, experimental psychology and psychiatry have made active efforts in the direction of objectifying differences in behavioral performance between psychiatric patients and normals. The nature of the clinical impressions that the investigator may undertake to objectify need not concern us here. They are many and
varied and rarely have any compelling arguments been advanced for picking one above another as being a key or root symptom. Thus various investigators have attempted to objectify the general observation of perceptual disorder; others, the common observation of psychomotor retardation, and so on. We have argued in the past for the importance of studying the temporal processing of stimuli as being an important dimension to study (Zubin & Kietzman, 1966). We have further argued that phenomena which occur relatively rapidly in time, in the first 1000 msec or less, might be somewhat less susceptible to contamination by cultural and conceptual variables. Of course even this criterion must be applied with a great deal of caution since no matter how short the span of time one might be dealing with, cultural or conceptual contamination may arise in determining the attitude or point of view that the subject brings to the presentation of the stimuli and nature of the task that is to be performed.

It might seem paradoxical that one would undertake to measure behavioral performance if one is concerned with obtaining some reflection of central nervous system functioning in a particular illness. Why not obtain measures which are more directly neurophysiological? When, for example, we obtain differences in event-related potentials (ERP) between schizophrenics and normals, are we not by that token demonstrating organic differences in the central nervous system between the two groups? A précis of our answer to this question, which we will expand below, is as follows: First, the interpretability of ERP findings depends heavily on the interpretability of the behavioral designs which generate the ERP data; and second, we need approaches that will yield information beyond the mere fact of group differences in ERPs, in order to make meaningful inferences from ERP findings.

ERPs are exquisitely sensitive to the psychological meaning of the stimuli (Sutton & Tueting, 1975), and even to the moment to moment alterations in the psychological meaning of identical stimuli (Squires et al., 1976). A dramatic example of such properties of the ERP is the fact that we obtain what is apparently the same P300 component whether it is the presentation of the stimulus, or the absence of the stimulus, that conveys sufficient information (Ruchkin et al., 1981; Sutton et al., 1967). As a result of the sensitivity of the ERP to psychological variables, ERP findings in relation to psychiatric status become, at least at first blush, no more interpretable in relation to differences between psychiatric patients and normals than behavioral findings.

The more accurately ERPs reflect the psychological state of the subject, the more likely it is that a confused patient will yield a "confused," or at least very different, ERP. We had a very vivid demonstration of this some years ago when we asked naive subjects who knew nothing about ERPs to compare sets of three wave forms and tell us which wave form was most different from the other two. In each trial, we presented one wave form apiece from a schizophrenic, a psychotic depressive and a normal. The schizophrenic wave forms were typically so unusual in their conformation, so relatively "messy," that our completely naive subjects had no difficulty in identifying the wave forms of schizophrenics as the most different (see Fig. 7 in Shagass et al., 1978). It would seem then that here we have very dramatic evidence of the operation of the "crooked" molecules that may be involved in determining schizophrenic thought. But, of
course, this is not necessarily so! As we noted above, the ERP is enormously sensitive to the general psychological state of the subject as well as to the interpretation of the task and to the role of the stimulus in relation to the task. The complexities of the task may be too much for the patient. The patient may be less cooperative, less attentive, concerned with emotional problems, or listening to hallucinated voices, and may care very little about the apparently pointless task the experimenter presents. Paradoxically, we may also obtain uninterpretable results because the patient is too motivated, too anxious, and too concerned with the implications of poor performance. Our ERP findings, therefore, may be no more than a demonstration that physiology faithfully reflects the fact that the schizophrenic approaches the experimental situation and the task differently from the normal. If we could induce a similar state of distraction and uncertainty or indecision in normals, their wave forms would also appear aberrant.

How then can attention to the behavioral aspects of our experimental designs aid us in obtaining a valid indication that it is the physiology or anatomy or biochemistry of the nervous system that is at fault in obtaining particular differences between some group of patients and other groups of patients and normals? It is certainly true that all the problems of obtaining reliable and meaningful measurement that the patient brings into the psychological laboratory are also present in the physiological laboratory. In fact, we have just argued that it is quite possible and even likely that many of the physiological differences were obtained as a result of differences in psychological state.

Let us turn now to some approaches to resolving these difficulties.

Shagass (1976), in his consideration of the same set of problems, distinguishes between "psychophysiology" and "pathophysiology." He notes that from the psychophysiological point of view, it is no surprise that the schizophrenic patient who we know has trouble paying attention generates a smaller CNV -- which is a correlate of attentive readiness -- than the normal. On the other hand, from the pathophysiological point of view, it would be of great interest if the schizophrenic patient were to generate abnormally large CNVs while inattentive. Such a finding would constitute a more convincing physiological sign of deviance rather than a normal physiological reflection of a deviant psychological state (e.g., inattention). A finding which would qualify as pathophysiological is Shagass' report of larger, less variable, and earlier exogenous components in the first 100 milliseconds of the somatosensory evoked response in overtly psychotic chronic schizophrenic patients -- patients in whom the later more endogenous components are smaller in amplitude and highly variable. Shagass interprets this finding of less variability in the early components as evidence of impaired modulation or filtering of sensory and perceptual input, perhaps as a result of underactivation of the mesencephalic reticular formation.

The finding of such apparent anomalies is highly intriguing and holds real promise. Our own thrust has been in a somewhat different direction. We have tended to emphasize the fact that many of the fundamental problems of obtaining reliable and meaningful data in psychiatric patients are the same whether we are measuring behavior alone, or whether we are also obtaining ERP correlates.
Before turning to a discussion of our own strategies for obtaining reliable and meaningful data in psychiatric patients, we would like to discuss a critical issue, partly implicit in Shagass' strategy, namely, the ways in which ERPs can be used to yield new information.

When one obtains, for example, a slower reaction time in schizophrenic patients than in normals, it is difficult from the behavioral data to infer the reason for the slower reaction time. Among many possible reasons, the schizophrenics may have been less prepared for the stimuli; or they may have had greater difficulty in discriminating between different stimuli; or they may have been less confident in their decisions; or their decision times may have been longer; or their response execution time may have been longer. Bearing on such issues is the fact that significant progress is being made in basic research with normals in our understanding of the functional role of various components of the ERP. For example, the terminal CNV preceding the stimulus is related to perceptual anticipation; the amplitude of the N1 region reflects the degree of selective attention to a particular channel of sensory input; the latency of N2 is related to decision time; P3e amplitude is related to the immediate meaning of the stimulus; P3b (often referred to as P300) is related to contextual evaluation of the stimulus (e.g., in order to give a larger P3b to a "rare" event, one needs the larger context which defines that event as rare); and Slow Wave amplitude appears to reflect the amount of further processing required by the stimulus. Thus ERP data have the possibility of shedding light on processes that are hidden from view and difficult to disentangle from behavioral data alone. It is in this context that some of our ERP findings may be potentially most useful, e.g., the finding that it is the P300 component -- which is related to contextual evaluation of the stimulus -- that is apparently most reduced in schizophrenic patients (Levit et al., 1973), or the finding that some demented patients show dramatic increases in the latency of the P300 component (Goodin et al., 1978).

Baribeau-Braun et al. (1983) provided an interesting example of the use of ERP data to shed light on processes that are "hidden from view." They compared schizophrenics and normals under different conditions of selective attention and rates of stimulus presentation. The ERP components they examined were N1, which is sensitive to selective attention, and P300, which is related to stimulus evaluation. They found abnormalities for schizophrenics in both components, but because the abnormalities occur under different conditions, neither abnormality can be due to the other. The N1 abnormality occurs at slow stimulus rates, but not at fast stimulus rates; the P300 abnormality occurs at both slow and fast stimulus rates. Thus the abnormality in stimulus evaluation represented by P300 is not secondary to a deficit in selective attention since it also occurs at fast stimulus rates, when subjects are selectively attending to stimuli as demonstrated by the normality of N1.

This example illustrates three methodological points. First, the normality of one component when another is abnormal adds conviction to the finding of a specific abnormality since, if all components are abnormal, we do not know whether the schizophrenics are truly defective in the processes represented by the various components, or whether they are simply not relating to the task. Second, the normality
of one component when another is abnormal adds interpretive meaning to the abnormality. This kind of finding makes it possible to begin to identify the stage of information processing where difficulty is being experienced — for further, the apparent normality of brainstem potentials in schizophrenics whose cortical potentials are abnormal suggests the difficulty is at a higher level of processing (Pfefferbaum et al., 1980b). Finally, knowledge of the functional role of various components of the ERP can assist in identifying the process or processes involved in a deficit. Thus, in the Baribeau-Braun et al. data, the fact that the behavioral responses of schizophrenics were slower and somewhat less accurate than those of normals left open the question of what aspect of information processing was at fault. The ERP data suggest that it is not a generalized defect in selective attention that is at fault, but rather a defect at a later stimulus evaluation stage — or perhaps more generally, in the control and maintenance of a selective processing strategy.

We turn now to our strategies for obtaining reliable and valid behavioral data in patient groups. As should be clear from the foregoing discussion, and will be further illustrated below, it is our conviction that, particularly in information processing paradigms, the problems of obtaining meaningfully interpretable ERP data, and the problems of obtaining reliable and valid behavioral data, are inseparable.

II RESEARCH STRATEGIES

Strategy 1: Simplicity of the task

Well designed simple tasks have certain advantages over more complex tasks. If both the stimulus structure and the response are simple, the possibility is reduced that there will be alternate modes of accomplishing the same task. This increases the likelihood that deviance can be interpreted to reflect central nervous system function directly (see Brindley, 1960 for a discussion of the logic of making inferences with respect to the central nervous system from behavioral tasks). In comparisons of patients and normals, one also wants to avoid the possibility that differences in performance arise from alternate modes of doing the same task. (However, the advantages of "simplicity" must be balanced against the necessity that the task should fully engage the subject — see Strategy 3, below).

We have often used reaction time, and simple reaction time at that, to measure differences in performance of patients to different kinds of stimuli. In other words, in a particular experimental design we may present different stimuli in random order successively, but the subject need make no conscious discrimination among the stimuli. Regardless of which stimulus occurs, subjects have only to lift their finger as rapidly as possible. The rationale, which is well established in the psychophysical literature and confirmed in our laboratory, is that different stimuli will give different reaction times although no conscious discrimination between them is required (Kietzman & Sutton, 1977). Presumably, automatic information processing mechanisms are at play here. Such automatic processing is reminiscent of the ERP finding of a "mismatch negativity" for rare stimuli which deviate physically from the background of frequent stimuli (Nätänen et al., 1980). The mismatch negativity is obtained even when the subject is engaged in another task, and the stimuli to which ERPs are obtained are task irrelevant.
In a reaction time study (Bruder et al., 1975), we presented either a stimulus package consisting of a 25 dB click by itself or a stimulus package consisting of a 25 dB click followed 15 msecs later by a 10 dB click. If the reaction time to these two stimulus packages is different, e.g., if the reaction time to the double click is faster than the reaction time to the single click, we have direct evidence that the central nervous system appreciated the difference between these two stimulus packages.

The simplicity of the task, both in terms of stimulus structure and in terms of response, is also useful for the examination of ERP correlates. While significant strides have been made in delineating and specifying the functional role of a number of ERP components, there remain many problems to be solved. Overlap of components in time and space (Sutton & Ruchkin, 1984) continues to create difficulties which can only be dealt with effectively under optimal conditions: where relatively few components are involved, and where, in some time region and/or electrode site, overlap is minimal. With respect to the functional role of components, we do not have a final parsimonious formulation even for those components that we know a great deal about, while for other components, our specification of functional role is in an even more preliminary stage. Complex tasks tend to bring many components into play and this is not desirable when one is attempting to compare different groups of subjects. For example, in a recent study (Friedman et al., in preparation) in which children and adults matched pictures under a variety of rules for categorization, we found only two out of six components which appeared to be homologous across the two groups. Apparently, the non-homologous components reflected differences in the way children and adults processed the same task, even though both groups performed the task at a high level of accuracy. In contrast, simple tasks generally involve fewer components and this kind of problem is much reduced.

Returning to reaction time measures, the reaction time of schizophrenic patients, and to some extent of most hospitalized psychotic patients, is much slower than normals to all stimuli. This brings into play our second strategy.

**Strategy 2: Own control design**

Because of the differences in intersubject variability on almost any measures between patients and normals, there are problems to be coped with in both experimental design and statistical inference. While statisticians have often been critical of our reluctance to use classical analysis of variance designs which require that each cell contain measures independent of any other cell (different individuals), we have always argued that this is not a feasible or desirable strategy in comparing the performance of mental patients with the performance of normals. It is well known that the intersubject variability among psychotic patients is quite high, even when we limit our sample to a particular diagnostic category. This arises presumably from the nature of specific illnesses, the nature of the hospital setting, and probably the fact that different patients are suffering from different illnesses although we may with our current limited knowledge be classifying them erroneously as being in the same group. Therefore, it becomes an exercise in futility to have different groups of subjects for each experimental condition. Rather we use repeated measures designs in which both patients and normals are tested.
under all conditions, despite the non-independence of the ensuing measures. The inflation of the degrees of freedom inherent in such designs requires the use of corrections in analyses of variance (Jennings & Wood, 1976), or alternatively, the use of multivariate analyses of variance which bypass this problem. Whenever possible, we also randomize our experimental conditions, preferably on a trial by trial basis, so that subjects cannot with any ease assume differing attitudes when a given experimental condition is presented, e.g., “this is a challenging condition, I will try harder,” or “this condition is too difficult, I give up on this one,” or “this is an easy condition, I don’t have to try so hard” and so on.

In ERP research, own control designs are almost the rule, even in studies with normal subjects. This probably arose from the fact that it was quite obvious from the very beginning of human ERP research that wave forms obtained under the same experimental condition were clearly different from subject to subject. On the other hand, with careful attention to experimental design, they were quite reliable within subjects. We were so impressed with this stability of the wave forms within the same subject that we likened it to the subject’s signature. Yet despite reliable intersubject differences, differences between experimental conditions are also reliable, at least in normals. In schizophrenic patients, the dimension of the state of the illness, as well as probable non-homogeneity of the diagnostic grouping, tends to compound intersubject differences and makes own control designs even more necessary. We have observed, for example, that ERP wave forms tend to change in the direction of appearing more normal as a function of clinical improvement (Levit et al., 1973), a finding which might have been occluded had we not compared wave forms within each individual.

One highly reliable finding within the context of applying an own control design has been that reaction time of schizophrenics is more retarded than that of normals by shift of sensory modality of the stimulus in successive trials (Mannuzza et al., 1984; Rist & Thurm, 1984; Spring, 1980; Sutton et al., 1961; Sutton & Zubin, 1965; Waldbaum et al., 1975). This disproportionate retardation of reaction time for crossmodal as compared to ipsimodal stimuli in schizophrenics is found even when the stimulus of the previous trial occurred several seconds before. Rist and Thurm (1984) demonstrated that this difference between schizophrenics and normals can be obtained for sufficiently different stimuli in the same sensory modality, suggesting that the key variable is change, or “shift” vs. “repetition,” rather than crossmodal vs. ipsimodal. Furthermore, the greater degree of reaction time retardation for shift in schizophrenics is not due to a greater degree of mismatch between expectancy and the stimulus actually presented. A greater retardation for shift as opposed to repetition is found for schizophrenics even when they are informed prior to each trial what the sensory modality of the stimulus would be (Waldbaum et al., 1975; Spring, 1980).

In experiments using ERP measures, the P300 component is also shown to be sensitive to sequence of stimuli, and this sensitivity to sequence also differs for schizophrenics and normals (Levit et al., 1973; Verleger & Cohen, 1978). One interesting paradox though: While the reaction time of schizophrenic patients is more retarded than that of normals by shift of sensory modality, it is the P300 of normals which is more
increased in amplitude by shift of sensory modality (Sutton et al., 1978). Since one of the
determinants of the amplitude of the P300 component is that it is larger when more
uncertainty is reduced, i.e., larger the more "surprising" the stimulus, the ERP data imply
that it is the normals who are more "surprised" by shift of sensory modality. This
inference was supported when we obtained guessing data. Normals predict repetitions
significantly more often than they predict shifts, whereas the pattern for schizophrenics is
closer to 50:50 (Waldbaum et al., 1975).

In the above designs, repetition and shift are equally probable or improbable.
Somewhat different effects are obtained when one manipulates the relative probability of
shift and repetition. In Steinhauer and Zubin's (1982) study, following relatively rare low
tones, a high tone was always presented, whereas following relatively frequent high
tones, a high tone was presented in only two-thirds of the trials. As would be
expected, normal subjects show larger P300 amplitude to the less predictable high tones
(ones that follow high tones) than to the fully predictable high tones (ones that follow low
tones). The schizophrenics, in addition to generally lower P300 amplitude, reverse this
direction of findings for normals, giving a larger P300 to high tones which follow low
tones (predictability = 1.0) than to high tones which follow high tones (predictability = .67).
It is as if the schizophrenics are more influenced by shift than by probability.
Alternatively, it is as if, like normals, they are sensitive to shift, but, unlike normals, they
are not consistently sensitive to probability differences. Inspection of the data of
Duncan-Johnson et al. (1984) indicates that in a 50:50 random program, for normals,
P300 amplitude increases when repetitions exceed three, presumably reflecting the fact
that, in a random program, long strings of repetitions are subjectively more improbable.
Schizophrenics on the other hand do not show this increase in P300 amplitude as the
number of repetitions exceeds three.

Despite a number of empirical findings for shift vs. repetition in both reaction time
and ERP studies, a general understanding of the meaning of the obtained differences
between schizophrenics and normals continues to elude us. Callaway and Nagdi (1982)
have made an interesting attempt to integrate the shift vs. repetition findings in
schizophrenics into a more general formulation that schizophrenics are deficient in limited
channel capacity processing, i.e., more conscious, or voluntary, or "controlled," serial
processing, but not deficient in more "automatic" parallel processing. Since the shift vs.
repetition experiments were not designed with such an axis of comparison in mind, new
research is needed to test the relevance of this proposal.

Strategy 3: Full demand tasks and the role of behavioral evidence
in demonstrating that patients are fully engaged

Partly in conflict with the simplicity requirement is the requirement that
experimental tasks engage the subjects fully. Full engagement in the task is an important
way of assuring the minimization of what we call subject option (Sutton, 1969; Sutton,
1973). It is inadvisable, for example, to compare ERPs for schizophrenics and normals to
stimuli for which the instruction is simply to "attend," or to "ignore" the stimuli. In such
designs, there is no experimental control over the degree to which the attention of
different subjects or diagnostic groups will drift in and out during the course of the task,
with consequent effects on the recorded ERPs. In such situations subjects may invent games to play with the stimuli to reduce boredom, and some schizophrenic subjects may hallucinate intermittently. On the other hand, when the task demands full engrossment of the subject -- for example, the detection of the presence or absence of a threshold-level click -- it is more likely that subjects are processing the stimuli in the way the experimenter requires.

Of course, it is essential to have some method for evaluating the engagement of the subjects. For example, in one study we showed that for a reaction time measure, schizophrenic patients showed a shorter critical duration for integration of visual energy than did normals (Collins et al., 1978). Our evidence that the findings were not due to greater variability of the reaction time of schizophrenics arose from the inclusion of stimuli at more than one level of light energy. The schizophrenics, despite longer and more variable reaction time than the normals, did not differ from normals in the degree to which an increase in energy reduced their reaction time (Collins, 1972). This provided direct evidence that the schizophrenics were engaged in the task, and indirect support for our finding of interest which related to the different way they handled stimuli of the same physical energy, distributed differently over time.

Strategy 4: Controlling the criterion variable

An important source of misinterpretation of the meaning of differences in performance between schizophrenics and normals arises when the criterion variable is not controlled or measured. In the classical psychophysical method of limits, as the intensity of a stimulus is slowly increased from some level below threshold, subjects are asked to report when they detect the signal. Particularly in a difficult task, subjects vary in their willingness to take risks in making decisions. Therefore, when there is no control for criterion differences, subjects who need to be sure before they are willing to report that they detect a signal will inevitably score as less sensitive than subjects who are more willing to take risks. Thus, for example, if schizophrenics are more cautious than normals, the method of limits will define them as less sensitive than normals.

An important advance which has been made by Signal Detection Theory is the demonstration that two independent measures are obtained in threshold situations with this method, one of which, called d prime, is a measure of sensitivity, and the other, called beta, is a measure of criterion. It turns out that it is indeed the greater cautiousness of schizophrenics that accounts for the earlier reports that their critical flicker fusion frequency was less sensitive than that of normals. What W. C. Clark et al. (1967) have shown is that, rather than being deficient in sensitivity, their psychiatric subjects (almost all of whom were schizophrenics) want to see a clear unmistakable flicker before they are willing to say that the light is flickering (also see Herskovic et al., in press, for findings in depressives). When measures of sensitivity are obtained which are independent of criterion, it turns out that there is no difference between schizophrenics and normals in flicker sensitivity. The W. C. Clark et al. study has cut the underpinnings from a great deal of research comparing the sensitivity of schizophrenics and normals which was done using classical psychophysical techniques. It is not that there are no differences in the sensitivity of schizophrenics on the many variables that have been measured, but rather
that we do not know whether those findings which have been reported as differences in sensitivity actually represent differences in sensitivity or differences in the degree of cautiousness. While differences in criterion are also informative, they are less likely to represent evidence of differences in the central nervous system of schizophrenics, but are rather more likely to represent a reflection of more global state or trait variables, for example, the way the illness impinges on the personality or the natural premorbid proclivities of the individual.

Presumably, being cautious in making a decision reflects a lack of confidence, and this brings us to the repeated finding in ERP research that the amplitude of the P300 component in schizophrenics is much smaller than that of normals. It has been demonstrated in research on normals, that the amplitude of P300 is highly sensitive to the confidence that an uncertainty has been resolved (K. Squires et al., 1973; Ruchkin & Sutton, 1978) -- the greater the confidence, the larger the amplitude. Does this mean that the smaller P300 in schizophrenics is another instance of their greater cautiousness, i.e., their lack of confidence? Clearly, new research comparing schizophrenics and normals is needed in which ERPs are recorded in the same trials as behavioral performance, and in which a variety of approaches are used to obtain independent estimates of decision confidence.

**Strategy 5: Accuracy indicators**

In Signal Detection Theory, in any given observation interval, a signal (let us say a very low intensity click) is presented, or a "blank" (no signal) is presented. The subject must report either "yes" (there was a signal), or "no" (there was not a signal). Thus, four contingencies are possible: when a signal has been presented, there are Hits (correct "yes" responses) and Misses (incorrect "no" responses); when a blank has been presented, there are False Alarms (incorrect "yes" responses) and Correct Rejections (correct "no" responses). Stated conceptually, the ratio of correct responses (both Hits and Correct Rejections) to the total number of trials represents a criterion-free estimate of sensitivity, whereas the ratio of "yes" responses (both Hits and False Alarms) to the total number of trials represents a sensitivity-free estimate of criterion. This method clearly has accuracy indicators: We know how to weigh the subject's saying "yes" when a stimulus was presented because we know the degree of readiness to say "yes" when a stimulus was not presented. If, for example, the subject says "yes" equally frequently regardless of whether the signal is present or absent, we can conclude that the subject is merely guessing. A high proportion of correct responses permits us to conclude that discrimination of signals and blanks is above chance levels. If the proportion of "yes" responses is low, we can conclude that the subject is cautious, whereas a high proportion of "yes" responses permits us to conclude the subject's criterion is "liberal" or "lax."

In contrast to yes/no, or single observation interval, Signal Detection methodology which permits differences in criterion, but nevertheless yields a criterion-free measure of sensitivity, multiple interval forced-choice methods eliminate the criterion from consideration by requiring the subject to say "yes" in every trial. For example, in three interval forced-choice, in which there are three successive observation intervals in each
trial, two of the intervals contain blanks, while one of the three contains a signal. Which of the three observation intervals will contain the signal is programmed at random. By specifying which observation interval is "most different" from the other two, the subject is in effect responding that the signal occurred in that observation interval. Since one of the observation intervals must be specified in each trial, there is no way for the subject to express a cautious or lax criterion.

Both the single interval signal detection method and the three interval forced-choice method have the advantage of having accuracy indicators. Unlike the method of limits in which a signal is presented in every trial, both the single interval method and the three interval method have both signals and blanks. In the single interval method, at random, a fixed (often 50) percent of the trials are signals, and the remainder are blanks. For the subject's decisions to be correct, "yes" must correspond to the presentation of a signal and "no" must correspond to the presentation of a blank. In the three interval forced-choice method, the signal is presented at random in the first, second or third observation interval. For the subject's decisions to be correct, the observation interval specified as "most different" must be the observation interval in which a signal was actually presented.

Thus, two sorts of controls are immediately available. First, when the subject's accuracy of detection is above chance, we are assured that the subject is doing the task. Secondly, in ERP research, for the single interval method, we can examine each of the four categories separately, and in this way reduce the confound of different perceptions entering into the same ERP average. For example, when the subject's overall accuracy level is above chance, an ERP average of hits is more likely to represent homogeneous perceptions for the majority of trials in the average, though some proportion will correspond to chance decisions. Similarly, for the three interval forced-choice method, when the subject's accuracy is above chance levels, an ERP average of observation intervals which the subject identified correctly as containing a signal is more likely to correspond to homogeneous perceptions for the majority of observation intervals in the average, though, again, some proportion will correspond to chance decisions. For both methods, the requirement of confidence ratings by the subject can even further increase the uniformity of trials in each average. For example, in one study we had subjects identify which of two observation intervals contained a very low intensity click. Only for high confidence decisions was P300 clearly larger for the observation interval in which the click was presented than in the observation interval in which the click was absent (Sutton et al., 1982).

In a behavioral study, in which we used a three interval temporal forced-choice method, we were able to demonstrate that hospitalized psychiatric patients had less sensitive absolute thresholds than normal subjects for auditory clicks (Bruder et al., 1975). While this difference between patients and normals was on the order of 3 dB, when we subsequently separated the patients by diagnostic classification it turns out that one group of patients, those with affective psychoses (the majority of whom were manic-depressive), were less sensitive than the normals by 6 dB, whereas the patients classified as schizophrenic did not differ significantly from the normals.
Using a Signal Detection Theory paradigm, ERPs have been recorded in normals while they were engaged in the detection of the presence or absence of very low intensity signals (see for example, Hillyard et al., 1971; Parasuraman et al., 1982; Paul & Sutton, 1972; Squires et al., 1973). Hits yield the largest average P300, while average P300s to the other three categories — correct rejections, false alarms and misses — are substantially smaller. For correct rejections and false alarms there is no physical stimulus for time-locking the P300, and it is only the time estimation of the subject — the absence of a stimulus at the point in time after the warning stimulus when the stimulus should have been there if it was going to be present — that time-locks the appearance of the P300. For these categories of trials, it has been generally assumed that the average P300 is small because there is no stimulus present for the ERP to time-lock on. Consequently, the individual trial P300s are jittered in time and averaging of these individual trial P300s results in a flattened and small average P300. However, for the misses, a stimulus for time-locking P300 is physically present, though the subject did not detect it. If the availability of a stimulus on which to time-lock is the critical factor, it is surprising that the P300 in the average for misses (stimulus is present) is not larger than the P300 in the average for false alarms (stimulus is absent). This suggests that it is the occurrence of the decision, and not the occurrence of the stimulus, which elicits P300. Thus, in the misses, where the stimulus is present but not detected, it would be the variation in decision latency which is responsible for the P300 time jitter that flattens the average P300.

Most clearly for the hits where P300 is relatively large, this component increases in amplitude both with greater accuracy of detection and with higher confidence in the decision. These findings taken together suggest that confidence in the decision is the key determinant of P300 amplitude in such paradigms, since the very slightly higher intensities which permit the subject to be more accurate would also result in more confident detections. The critical role of confidence in determining P300 amplitude in a signal detection design also bears on the relative amplitude of P300 for the various decision categories. In addition to trial to trial jitter in the time of occurrence of decisions, any differences in degree of confidence among the different decision categories would bear on the relative amplitude of P300 for these categories. Furthermore, since less confident decisions are likely to be more delayed, and delayed decisions are likely to be more variable in latency, confidence and time jitter of the decision probably interact in determining the amplitude of P300.

We are not aware of any ERP studies in psychiatric patients of detection at absolute threshold intensity levels. However, the availability of behavioral findings, particularly for patients with affective psychoses, and the potential clarity of interpretation possible when signal detection or three interval forced-choice methods are used, would suggest that such studies might be illuminating.

**Strategy B: Better patient performance**

While the use of forced-choice methodology controls the criterion problem very well and also provides some safeguard against falsely attributing to patients less sensitivity than they actually possess, it does not completely solve our problems in
obtaining meaningful data from patients. For example, one can make the argument that the patients who had poorer thresholds than normals were simply less attentive or less motivated than normals and, therefore, they made more errors in their forced-choice decisions. Presumably, according to this argument, had the patients been in a state which had permitted them to be fully attentive and cooperative they might have obtained just as good thresholds as the normals. Still other problems which are more statistical in nature arise in the interpretation of experiments in which one of the groups being compared may be assumed to display generalized deficit (Chapman & Chapman, 1973). One satisfactory antidote to these various problems is as follows: to demonstrate in a particular task that patients were able to perform better than normals (Underwood and Shaughnessy, 1975; Sutton, 1973). It is not possible to attribute better performance to poorer attention and cooperation on the part of patients. While several attempts to meet this rigorous criterion have been made, none have yielded robust, reliable, and completely unequivocal results.

In the Bruder et al. (1975) study discussed above, we presented randomly either a click which was 25 dB above the subject's threshold, or the identical click followed 15 milliseconds later by a click which was 10 dB above the subject's threshold. Subjects were required to lift their finger from a key as rapidly as possible to the presentation of either stimulus package. We found that for normals the average reaction time to the paired clicks (the 25 followed by the 10) was only very slightly faster than the average reaction time to the 25 dB click by itself. However, for patients with affective psychoses, there was a much larger improvement in reaction time for the paired clicks over the reaction time to the single click. Here then, is evidence that these patients are benefiting more from the presence of the additional 10 dB click than were the normals. The reaction time of the patients with affective psychoses is more facilitated by the presence of the 10 dB click; they are thus more "sensitive" to its presence, and in that sense are performing better. Schizophrenic patients did not show any more reduction in reaction time than did normals due to the presence of the 10 dB click in the stimulus package.

A variety of post-hoc analyses were addressed at the possibility that a more trivial interpretation would explain the finding in the patients with affective psychoses. For example, could the fact that these patients had the longest reaction times somehow produce the finding of greater reaction time facilitation? We were able to discount this interpretation to some degree by showing that five of the schizophrenics who had reaction times as long as those of the patients with affective psychoses did not show the facilitation. Since the intensity of the stimuli were determined for each subject by presenting the stimuli at the same intensity above each subject's threshold, could the facilitation for the patients with affective psychoses have been brought about by the higher thresholds for this group? Again, this interpretation is made less probable by the lack of a significant correlation for these patients between thresholds and degree of reaction time facilitation. One final possible source of artifact could not be ruled out. If the reaction time for normals was close to the physiological limit, then it is possible that the presence of the 10 dB click did not facilitate their reaction time because their reaction time to the 25 dB click by itself was close to the fastest reaction time they
could produce.

In the Collins et al. (1978) study, there were two 2-millisecond flashes of light, which in one stimulus package were separated by 2 milliseconds and in another stimulus package were separated by 0 milliseconds — in effect, for the second case, it was a 4 millisecond flash of light. Note that here the two stimulus packages have the same total amount of light energy. However, in the one case the light energy is distributed over 6 milliseconds, and in the other case the same amount of light energy is distributed over 4 milliseconds. As is known from the Bunsen-Roscoe law, if the stimulus duration lies within the critical duration period, then the response to these two packages of light will be identical because when stimuli are within the critical duration period, the summated energy determines the level of response, i.e., none of the energy gets "lost." If, however, the longer stimulus package exceeds critical duration for the response being measured, some of the energy is in effect lost in relation to that response, and differences would be obtained for the two equal energy packages.

The two equal energy stimulus packages were presented at random and subjects were instructed to lift their finger as rapidly as possible whichever stimulus package appeared. For our normals these two packages are apparently perceptually identical. In a verbal response paradigm, normal subjects were unable to distinguish between equal energy packages in which the light is distributed over 4 milliseconds or 6 milliseconds. Similarly, normals show no difference in reaction time between these two packages of light. However, schizophrenic patients did show significantly longer reaction time to the 6 millisecond package of light than to the 4 millisecond package of light. From these findings, we infer that for normals, critical duration for reaction time is longer than 6 milliseconds, i.e., they are able to use the total energy of the 6 millisecond package in generating their reaction time response, whereas for schizophrenics, critical duration for reaction time is shorter than 6 milliseconds, i.e., they are not able to utilize the total energy of the 6 millisecond package in generating their reaction time responses. The reaction time of schizophrenics, therefore, showed evidence of the difference in packaging, whereas the reaction time of normals did not.

What do we mean when we say that the schizophrenics are performing better than normals? What we mean in the critical duration experiment is that they can make a discrimination, by the differential speed of their motor response, that normals are apparently unable to make. If you like, the schizophrenics are more sensitive to the way light is distributed over time than normals are. It should be noted, however, that what we have done here is to turn what may well be a defect into an advantage. It is fairly likely that the visual system, particularly at close to threshold intensities, is "designed" to integrate energy over time, perhaps because it provides an evolutionary advantage in the detection of predators. Schizophrenic patients, presumably because of some difference in their nervous system, do not integrate energy over as long a period of time as do normals. We have taken advantage of this by creating a design in which that very inability results in their demonstrating a discrimination that normals cannot make.

Again a caveat, however. The effect is apparently a delicate one, and in one
attempt to replicate it, we were unsuccessful. However, in another study using auditory stimuli we obtained evidence for shorter critical durations for hallucinating psychotic patients who were diagnosed by hospital psychiatrists as schizophrenic (Bakoff et al., 1980).

Another example of the "better" performance of schizophrenics is the finding by Place and Gilmore (1980) that, under some perceptual conditions, schizophrenics achieve greater accuracy in counting tachistoscopically presented stimuli. In this task as well, the design turns what is probably a handicap of the schizophrenic patients into better than normal performance. The task is to count the stimuli, and for this purpose, the perceptual organization of the display could hinder counting performance if time is taken to process the display globally. If for schizophrenics, the perceptual organization of the display is less compelling than for normals, the schizophrenics would, by being less distracted by the perceptual organization of the display, be able to achieve better than normal accuracy in counting the elements of the display.

In a recent replication by Orlowski (1986), in addition to accuracy, reaction time, i.e., the speed of counting the number of elements in the display, was measured. In this study, the superior accuracy finding for schizophrenics was not obtained. However, Orlowski did obtain faster reaction time for the schizophrenics than for the normals. Orlowski was able to rule out the possibility that the schizophrenics produced faster reaction times by being more "impulsive," i.e., by responding to the speed requirement by making less effort to discriminate accurately. She obtained faster reaction times for the schizophrenics even when she limited her analyses to those subjects who performed clearly above chance levels, and for those subjects, even when she limited her analyses to trials in which correct discriminations were made. Clearly there is a real effect present and we are currently revising the experimental design in order to get the better performance of schizophrenics to evince itself in a more robust manner.

No ERP studies are available in relation to tasks where patients show better behavioral performance in the sense described here. Partly, this is due to the fact that none of the better performance findings are robust and sufficiently reliable to warrant such an undertaking. If one could get the behavioral findings to the level where they are robust and reliable, this would provide a desirable paradigm for exploration of ERP correlates. We would then obtain information on the nature of differences in information processing between patients and normals in a paradigm in which we were certain that the behavioral difference between groups was in the phenomenon under investigation and not a spin-off of general deficit.

Roth et al. (1984) describe a kind of "better" performance in schizophrenics in relation to ERP measures. In their task, Roth and Cannon (1972) presented frequent and rare auditory stimuli at random. Subjects were instructed to "ignore all stimuli as much as possible." Schizophrenics showed small P300s to the rare stimuli throughout the experiment, whereas the normals showed relatively large P300s to the same rare stimuli at the beginning of the experiment, and their P300s did not decrease to the level of the schizophrenic P300s till the end of the experiment. Large P300s to the rare stimuli
imply that probability differences were being registered. In this sense, the schizophrenics whose P300s to rare stimuli were small throughout the experiment followed the instructions to "ignore all stimuli as much as possible" more successfully than did the normals, and in this sense, displayed "better" performance.

However, this kind of "better" performance is probably a spinoff of the general deficit syndrome. Schizophrenics show smaller P300s than normals in almost any task, and whatever the instructions. For example, in the Levit et al. (1973) experiment where smaller than normal P300s were also obtained, the stimuli provided feedback on guessing outcome. Here, smaller P300s implied that patients were less involved in the task, i.e., that they performed more poorly than normals. Tueting (unpublished data) ran schizophrenics and normals in the paradigm used by Levit et al. with the addition of the instruction to report after each trial the identity of the feedback stimulus and whether their guess was right or wrong. One third of the schizophrenics performed at chance levels on the accuracy of their reports, whereas all the normals performed at essentially 100% accuracy.

In contrast to the Roth and Cannon (1972) type of "better" performance, in the Place and Gilmore design, both schizophrenics and normals were instructed to count the stimuli in the display. Not being distracted by the perceptual organization of the display was implicit in the instructions. Schizophrenics were more successful in counting accurately because they were less distracted by the perceptual organization of the display (Place & Gilmore, 1980), and in responding more rapidly because they did not lose time performing irrelevant global processing (Orlowski, 1986). In the Collins et al. design, the reaction time of schizophrenics showed evidence that they were detecting a difference in energy packaging which was not found for the reaction time of normals. In this sense, they were "discriminating" better than the normals. In the Roth and Cannon type of "better performance," if the schizophrenics were less attentive and less motivated, they would be less likely to register probability differences and, by that token, be more "successful" in following the instruction to ignore the stimuli -- as Roth (1984) points out, "in a Zen Buddhist way."

In summary, the value of designs in which schizophrenics perform "better" than normals arises from how successfully the better performance assures us that they differ from normals in the specific function being examined, i.e., when their better performance cannot be explained away as arising from other sources, such as generalized deficit. From this point of view, the finding that schizophrenics are more successful than normals in ignoring probability differences, even when this is the experimental instruction, does not permit the unequivocal assignment of deviance to a specific function. That such a finding is implicit in their being less attentive and more disengaged means that we have gained little in the way of new information. Similar difficulties are what make non-interpretable, in the sense we are requiring, the Braff et al. (1978) finding that in schizophrenics the blink response was less disrupted than in normals by a weak prestimulus tone. In contrast, discrimination by schizophrenics of energy packages that normals cannot discriminate, or better accuracy or speed of counting by schizophrenics, cannot be dismissed as another reflection of generalized deficit.
Strategy 7: Iterative approaches

All the above strategies are addressed to approaching behavioral aspects of experimental design in ways which will allow more unambiguous interpretation of findings of differences between psychiatric patients and normals. However, we are faced with at least as serious a problem in defining the nature of the independent variable, namely the group to which the subject belongs. We can, of course, be arbitrary about this and call any one a normal who is recruited from the United States Employment Service, and any one a patient who is resident in a mental hospital. However, some years ago when we applied scaled interview results to obtain a more objective description of all subjects, we found that as many as half of our “normals” did not clear as normal on our rating scales, and similarly about half of our patients who had been diagnosed as schizophrenic by the hospital psychiatrist did not clear as schizophrenic on our rating scales. When we limited our analyses to the schizophrenics and the normals who were assessed as such by our rating scales, we found that, despite the reduction in size of the two samples, the reaction time differences between groups were both larger and less variable (Waldbaum et al., 1975). It is clear that, for proper evaluation of differences on objective measures, we require better classification of our subjects.

Even within a diagnostic group, we do not know in advance which group of schizophrenics will have shorter critical durations, and which will not differ from normals. Nor do we know in advance which schizophrenics will show greater retardation due to shift of sensory modality. If such intragroup differences are reliable, they may provide important new information, i.e., an approach to refining our diagnostic classifications. The use of structured interview rating scales permits us to examine the characteristics of the patients who differ most clearly from normals on our laboratory procedure. Here we are interchanging our independent and dependent variables and saying that the individuals who perform in a particular way on our behavioral measure are a particular kind of patient. We know that we are violating the rules of statistical inference in which the hypothesis must be made in advance and that, therefore, our post hoc conclusions become more tentative, but the gain justifies this loss. What we must do when taking this approach is to treat our rather formal experiments as pilot experiments and take the stance that though we did obtain a statistically significant difference between group X and group Y, when we turn around and redefine our groups post hoc, this cannot be taken as valid statistical inference. The experiment must be replicated, this time with the finding of the previous experiment used as a hypothesis which is made in advance. In still further construct validating operations we can make efforts to find out what other properties characterize the patients singled out by our laboratory procedure. For example, if they are more retarded by shift of sensory modality, are they also the patients who perform better than normals in counting tachistoscopically presented stimuli under certain perceptual conditions?

Furthermore, we can make similar inroads into our rating scale categories. In the case of the critical duration findings, they were strongest in schizophrenic patients who showed evidence of thought disorder (Collins et al., 1978). This impels us to attempt to further refine our definition of thought disorder, perhaps even to develop a more objective measure of thought disorder through item analysis of existing scales, or through
the development of a new scale. After clustering patients into subgroups in accordance with their performance on laboratory tests and interview status, we could proceed to see whether these subgroups differ on other characteristics such as response to specific treatments, duration of hospitalization and type of outcome on follow-up. For example, we are finding that those depressed patients who show hemispheric lateralization opposite to normals on one of our tasks are the patients who respond to treatment with a tricyclic drug, whereas those depressed patients whose lateralization is most like the normals are the patients who do not respond to treatment with a tricyclic drug (Bruder, unpublished data).

What we have described is the direction of a continuing effort. Our ultimate objective is to be able with some precision to define subgroups of patients in such a way that we can have some indication which ones are, or are not, characterized by some kind of central nervous system deviation and to develop some insight into the nature of the deviation. At some point, we may be able to go to the biochemist and ask him what else he can tell us about the specific subgroups of patients we have found.

We should note that at that point our ERP research would also become more powerful. By bringing to bear what else is known about the ERP components which most reflect the group differences in our behavioral measures, we could both sharpen up our behavioral experimental designs and gain more insight into the nature or meaning of the group differences.

III SIGNIFICANCE OF NEUROPHYSIOLOGICAL FINDINGS

The differences between schizophrenics and normals on laboratory measures present us with an interesting paradox. The behaviors which bring the schizophrenic to our attention are gross behavioral deviations whose connection with the subtle laboratory findings such as critical duration for energy integration, higher auditory thresholds, and cross-modality retardation in reaction time are at best difficult to trace. What possible connection might exist between our findings and the schizophrenic disorder?

The total disorder, in all of its manifestations, is so encrusted with so much varying life experience that it defies analysis; all one can do is describe it. We have had 33 centuries of such descriptions. The earliest truly detailed description dates back to the Charaka Samhita text of Hindu Ayurvedic medicine in the first century A.D. (Haldipur, 1984). Despite the long history of awareness of the disorder, however, a fundamental understanding still eludes us. Careful, well-defined sociocultural, communicational, familial and other analyses have contributed to the understanding of some macro level aspects of the disorder; complementary micro level analyses -- e.g., the biochemical and neurophysiological work of recent decades -- are clearly also necessary to attempt to get an integrated theory of the disorder.

Our own approach is to try to deal with behavior that is relatively free of prior experience and cultural influence. When we limit ourselves to laboratory tasks using brief stimuli, we are attempting to obtain differentials between schizophrenics and normals that reflect a basic substrate of brain functioning in the processing of information. Such
differences may yield the culture-free indicators we so desperately need when studying
subcultural groups or when making crosscultural comparisons. By depending on
interviewing methods or clinical tests alone, or on techniques highly dependent on prior
reinforcement history, such as higher mental functions, we cannot escape the cultural bias
inherent in such techniques. It is hoped that culture-free indicators may help us detect
deviations that would be either falsely occluded or spuriously introduced when culture
dependent techniques are used.

A more speculative line of inference would be to propose that the small deviations
measured by subtle laboratory tasks are more than indicators of central nervous system
malfunction and serve in a causal role in relation to the development of some types of
schizophrenia. Perhaps deviations in simple micro processes which characterize the
schizophrenic, e.g., the inability to shift as readily from one sensory modality to another,
shorter critical duration and the weakness of perceptual organization, constitute the real
building blocks of schizophrenia. Conceivably, deviations in such lower level processes
would interfere with the smooth flow of more global behavior such as the filtering and
coherent sequencing of information processing and motor activity, or with the appropriate
allocation of attentional resources, resulting in ongoing stress which may create
difficulties in personal relations. In other words, schizophrenia may be “caused” by these
losses of pawns in the game of life -- the rest of the symptomatology which we
include under schizophrenia may in this sense be an “epiphomenon.” This formulation
represents one form of the hypothesis that schizophrenia occurs in a vulnerable individual
when subjected to sufficient stressors (Zubin & Spring, 1977). We are aware that it is a
highly speculative formulation, and we present it only as a possibility to be explored.

The relation between ERPs and the various psychophysical tasks which constitute
our experimental procedures requires some comment. ERPs represent the electrical
activity of the brain while a task is being performed. They reflect the activity of large
groupings of neurons which are engaged in the processing under examination in that task.
Differences in ERPs between schizophrenics and normals which are concomitant with
differences in behavior provide us with information on how the brain processes these
tasks. As noted above, we may be able to utilize our knowledge of the functional role
of various ERP components to provide insight into processes which are hidden from
view. By monitoring the ERPs corresponding to task performance we can perhaps find
out just where, when and how the information processing engaged in by patients deviates
from the corresponding processing in normals.

By emphasizing the methodological nuances of behavioral aspects of the research
we have been discussing, we may seem to have implied that the ERP itself can be
interpreted straightforwardly. However, as we have developed at length elsewhere
(Sutton & Ruchkin, 1984), interpretation of ERP data has become significantly more
complex as our knowledge has advanced. One source of the problem is that ERP
research has uncovered many more components, multiplying like the particles of nuclear
physics. Unfortunately, in scalp recordings, many ERP components overlap in space and
time. Thus it is no longer clear which and how many of the half dozen or so
components in the P300 epoch were being dealt with in earlier studies. For example,
there is the overlap of P300 and a slower positive component known as Slow Wave (Ruchkin & Sutton, 1983). Since all of these components serve quite different functional roles, misassignment of differences between patients and normals to the wrong component or components can result in serious errors of interpretation. The difficulties are further compounded by the fact that there are no straightforward methods for dealing with component overlap. In a sense, the task of unravelling the multiplicity of components of the ERP, and their functional roles, may be likened to that of identifying the loci and functional roles of the individual genes of the genome. What this means is that although component overlap can in many instances be disentangled, and proper assignments made, quite sophisticated approaches have to be taken to accomplish this end.

Interpretation of ERP data is even more complex when comparing findings for various groups of subjects. A common finding is that the "same" component has a somewhat different scalp distribution in different groups. For example, P300 has been reported to be more frontal in older subjects (Pfefferbaum et al., 1980a). The first question that needs to be dealt with is: is this apparent difference in scalp distribution due to the component under consideration, or due to overlap of the component under consideration with another component? For example, it turns out that the apparent differences in scalp distribution of P300 as a function of age are due to increased amplitude of overlapping Slow Wave in the older subjects (Pfefferbaum et al., 1984).

If the answer to the first question can dispense with the overlap issue, the next question is: Do the differences in scalp distribution between groups mean that we are obtaining two different components in the response of the two groups to the same task? For example, Knight (1984) found that while normals and patients with prefrontal lesions responded to infrequent target stimuli with the classical processing parietal–maximal P300 (P3b), the two groups responded quite differently to unexpected novel stimuli which were task irrelevant. For the normals, the unexpected novel stimuli yielded a fronto–central P300, whereas for the patients with prefrontal lesions, the same unexpected novel stimuli yielded the parietal–maximal P300 (P3b) which is obtained to infrequent target stimuli. Therefore, what the data demonstrate in this instance is that the same stimuli may elicit different components in different groups.

The burgeoning developments brought about by scanning techniques (CAT scans, PET scans, NMR), BEAM type imaging, cerebral blood flow imaging, and laterality findings should help us in mapping schizophrenic deviations in brain structure and function, and provide information which may help guide ERP research. Similarly, the recent breakthroughs in molecular genetics, the revival of the psychosomatic field via the discoveries in the immune system and its relation to stress ought to be harnessed to ERP research wherever possible. The search for "markers" which the vulnerability hypothesis has set in motion (Zubin & Steinhauser, 1981; Zubin, 1985) ought to be greatly facilitated as we begin to see the role that the underlying brain activity, as exhibited by the ERP, plays in the tasks eliciting the markers.

The issue of the correspondence between ERPs and behavior leads into more
fundamental, and by that token more speculative, questions. The ERPs we record represent an estimate of macropotentials, which in turn represent synchrony of activity in large populations of neurons. At first glance, it would appear that such activity must of necessity be epiphenomenal -- the muffled roar of the crowd which cannot tell us whether a home run has been hit or a political leader is being cheered. Or we may be dealing with phenomena which are analogous to the illusion provided by two synchronous clocks, one which designates hourly time visually and the other by ringing a bell (see Boorstin, 1981, p. 71). To the naive observer it may appear that the clock with the visual signal causes the second clock to ring, or perhaps the reverse. Uttal (1965) distinguished between "signs" and "codes." A sign represents an indicator of the system under examination, for example, the brain, but if it cannot be shown to affect any behavior, it is an epiphenomenon. A code is a set of transformation rules which represent a body of information, which in the case of the brain, must influence behavior. What is at issue here is whether our ERPs are signs or codes.

Over the last two decades, so many sensory, perceptual and cognitive phenomena have yielded lawful ERP correlates that the very richness of the findings provides at least suggestive evidence that we are not dealing with epiphenomena when we study ERP correlates. It seems improbable on the face of it that what our measures are tapping are epiphenomenal, rather than integral to underlying mechanisms, and yet yield so much lawful data. Even the absence of a significant stimulus, when that absence is informational, yields a P300 whose amplitude is related to the amount of uncertainty resolved by the absence! Yet, in principle, no amount of such evidence would constitute proof that our ERPs essentially constitute codes (albeit somewhat distorted by being recorded at scalp); that they constitute the actual way in which the brain does the processing. For example, on a priori grounds -- since the pupil is not part of the brain -- the fact that pupillary dilation mimics many P300-behavior relationships would not be taken to mean that pupillary dilation constitutes the brain's code of the information being processed.

One approach to the sign vs. code issue utilizes a design in which a behavior and an ERP measure, though found to be associated under some conditions, are found to be dissociated under other experimental conditions. For certain patterns of differential association, such a finding would suggest that the brain activity that the ERP measure reflects is not a code. D. L. Clark et al. (1969) reported what was ostensibly such a case. ERP components were found to be present at threshold intensities during a non-drug session. Administration of an anesthetic abolished these components at threshold intensities, though accurate behavioral responses could still be obtained at these intensities. However, the study was flawed because behavioral and ERP thresholds were measured in separate sessions (Donchin & Sutton, 1970; Paul & Sutton, 1973). But the question addressed by Clark et al. of course remains open.

The idea of dissociation under some circumstances between ERPs and behavior has been taken up more recently to aid in dissecting the way in which information is processed (Donchin, 1984), rather than to test the sign/code distinction. The reasoning which underlies this approach is delineated as follows: In a variety of experiments it has
been shown that the latency of P300 reflects the time it takes to process a stimulus. More complex decisions, for example, deciding whether the presented word is a female name or a male name, yield a longer latency P300 than deciding whether the presented word is "Nancy" as opposed to any other female name (Kutas et al., 1977). However, increased complexity of response execution, for example, responding with the right hand when the word "left" is presented, does not alter the latency of P300 (McCarthy & Donchin, 1979). These properties of P300 latency were used to explore the mechanism of interference in the Stroop Test. In this procedure, reaction time for color names which are printed in discordant ink colors, i.e., colors other than those which their names imply -- is compared with reaction time for color names which are printed in ink colors which correspond to the color names. When there is non-correspondence between color names and ink colors, reaction time is longer than when there is correspondence, but P300 latency is found to be the same (Duncan-Johnson, 1981). From this dissociation between P300 latency and reaction time, it is inferred that the mechanism of interference in the Stroop test is in the processes of response execution and not in the processes of stimulus evaluation.

In such applications, dissociation between ERPs and behavior does not necessarily imply that ERPs are epiphenomena. Rather, Donchin (1984) sees such contrasts between association and dissociation as a way in which "the true dimensions of human information processing might be thrown into sharper relief" (p. 118). Presumably, by focussing on different components, or recording at different areas of scalp, ERP correlates of the reaction times which are dissociated from P300 latency may well be found -- unless the relevant brain processes cannot be accessed by our usual ERP methodology.

The notion that what ERPs measure may constitute codes has startling implications. In conflict with the thrust of the overwhelming body of data and theory of modern neurophysiology which has shown correlates of information at the single neuron and subneuron level, the possibility that what ERPs measure constitute codes means that information can be coded as macropotentials. E.R. John (1967, 1976) is one of the few investigators who has consistently maintained and marshalled the arguments for such a view. He points to the fact that "the pattern of electrical oscillation recorded from a region corresponds to the integrated membrane potentials of vast numbers of neurons" and proposes that "information is the time course of coherence ... in an extensive population of neurons ... information is a statistical property of the neural aggregate" (John, 1967, Chapter XVI).

John's formulations were concerned primarily with the brain mechanisms mediating memory. The issue we are raising is even more general, namely, the mechanisms that mediate the various information processing activities for which we obtain lawful ERP correlates. It should be clear though that there is no implication here that information is carried exclusively in macropotentials and not also at the level of single neurons and their circuitry of interconnection. By analogy, that one understands the meaning of a sentence taken as a whole does not imply that information is not also carried in the sounds of the syllables constituting the words and in the meaning of the words constituting that sentence. Extending this analogy further, one notes that syllables, words and sentences
involve different levels of "chunking" of information, in a sense different levels of integration. Similarly, the roles attributed to endogenous components of the ERP are processes at a relatively high level of integration, e.g., perceptual anticipation for the terminal CNV, contextual categorization of the stimulus for P300, or semantic incongruity for N400. It seems to us quite possible, that for the brain mechanisms corresponding to such high levels of processing, information may be coded as macropotentials.

It is not clear how the epistemological role of ERPs bears on the conduct of ERP research in psychiatric populations. Nonetheless, while it would certainly be useful if ERPs could cast light on the nature of the defect in schizophrenia, whether they represent signs or codes, evidence that they represent codes would surely make ERP findings in schizophrenia more theoretically meaningful.

ACKNOWLEDGEMENTS

We are indebted to Muriel Hammer for a critical reading and substantive editing of this chapter and to Daniel S. Ruchkin, David Friedman and Mitchell L. Kietzman for valuable suggestions. Preparation of the chapter was supported in part by the Medical Research Service of the Veterans Administration and USPHS Grant #AA05909.
REFERENCES


Electroencephalography and Clinical Neurophysiology, 59, 85-103.


