JOSEPH ZUBIN, Ph.D.

The major interests of Dr. Joseph Zubin, Chief of Psychiatric Research, New York State Dept. of Mental Hygiene, have been interviewing technique, brain function and biometric methods in psychopathology.

Only those psychotherapists who are unashamedly mystical or sentimental will be easily able to dismiss Dr. Eysenck's charges. The material that he refers to in this paper is well known to the psychiatric community in America, but the forthright and unequivocal conclusion that he derives from them has the quality of a revelation. There is, however, a great difference between recognizing the problems of validating psychotherapeutic outcome and insisting that there is sufficient evidence to establish the null hypothesis.

Some of the problems in evaluating outcome of psychotherapy are shared with clinical research in general, while others are unique to psychotherapy. The adequacy of controls is a universal problem. The ideal situation of matched identical twins for a treatment and nontreatment grouping is no more available to pharmacologists than to psychotherapists. Many workers agree, however, that by using age of onset, duration of disease, sex and diagnosis as the minimum essentials for comparability of two patients, a workable control situation can be established. Obviously the first 3 parameters are as available for the evaluation of psychotherapy as for that of any other treatment, but it is the last, i.e., diagnosis, that raises particular methodological
problems. As long as the essential nature and cause of mental disease are unknown, and there continues to be disagreement among qualified persons concerning the broadest designations of mental disease, diagnosis will be imperfect and unreliable. But disputes about diagnosis can be circumvented for the purpose of a particular research project by a rigorous definition of terms or by the application of clearly outlined symptom complexes for matching of patients.

Recent work in our own laboratory has provided tools which make the life of the evaluator more bearable. By means of systematic interviews in which molecular aspects of behavior can be recorded as present or absent, an objective evaluation of the traits and behaviors which characterize the patient at the time of his admission, at release and on follow-up can be readily obtained. These instruments are: 1. The Mental Status Schedule (Spitzer, R. L., et al. 1963) for the use of psychiatrists; 2. The Structured Clinical Interview (Burdock, E. I., & Hardesty, A. S., in press) for use of clinical psychologists; 3. The Ward Behavior Inventory (Burdock, E. I., et al. 1960) for the use of nurses and attendants. Two of these instruments have already demonstrated their usefulness in the evaluation of outcome of drug treatment and of intensive milieu therapy. (Cole, J. O. Drug treatment in psychiatry. In P. H. Hoch & J. Zubin (Eds.) Goldberg, S., et al. 1963).

At the present time these tools are being evaluated from the point of view of two criteria: 1. severity of illness and 2. prognosis under specified treatment such as psychoanalysis, behavior therapy or no therapy. It is clear that if we can specify both of these criteria and standardize our instruments against them we can arrive at a much better basis for evaluating the outcome of any therapy. Two individuals who have similar prognoses should yield similar outcomes under two equivalent therapies or differing outcome under non-equivalent therapies.

An interesting conflict often arises in the mind of the clinician in prognostic studies. When the prognosis is very poor, the clinician usually redoubles his efforts under the assumption that the case he is treating has the one in 1,000 chance of making a recovery. This leads in some situations to undue self-adulation in success and to breast-beating in failure. In others, the failures are soon forgotten and only the successes are remembered. By providing prognostic baselines in a regular manner, the attention of the research clinician can be focused on those with good prognosis who failed and those with poor prognosis who succeeded, while the good as well as the poor prognoses which were borne out by experience can be laid aside. The disappointing failures and the unexpected successes can do more in furthering our understanding and improving our prognoses than gloating over expected successes or mourning after expected failures.

The problems of control are serious, but not necessarily insuperable. It is evident that there is no uniformity among mental health workers with respect to usage of such terms as "cured," "recovered," or "improved." But this problem is not unique to psychiatry either. These terms are after all the vocabulary of acute diseases, like appendicitis or a strangulated hernia, and are as imprecise in such chronic conditions as tuberculosis or cancer as they are in schizophrenia. Tubercle bacilli may be eradicated from the sputum and symptoms may disappear, but if the overall life expectancy has been shortened, can the patient really be called "cured"? In cancer therapy it is customary to refer to five or ten year "cures" but the arbitrariness and irrelevance of such terms to an individual patient do not compromise their use in evaluating treatments for cancer. It is clear that explicit and precise criteria for outcome are necessary in any satisfactory clinical evaluation. Freyhan's concept of "target symptom" which the therapy attempts to reduce or eliminate is an example of the type of precision that can be introduced.

The evaluation of such specific goals of therapy need not depend on ratings alone. Techniques for measuring changes in such specific characteristics as flatness of affect, overinclusive thinking, level of retardation in depression, anxiety, etc., are either already available or are in the process of development. Thus, Salzinger has developed objective measures of self-referred affect through the use of
reinforcement techniques, Payne has developed objective measures for thought disturbance in his tests of overinclusive thinking, and similar techniques for other target symptoms are in the making.

The criteria that have been used by psychiatrists too often cannot constitute the objective framework required for scientific evaluation. What we need are more definite measures of the improvement in 'comfort' of both the patient and his family. The preand post-treatment inventory of behavior suggested earlier is one promising maneuver. Further refinement will be necessary to avoid the atomistic and highly selective profiles that such scales currently tend to give. There is, moreover, a growing body of information from a variety of reliable sources that will ultimately contribute to our knowledge of spontaneous improvement rates. Such data will provide a baseline for comparative studies of outcome.

In order to make comparisons between the various types of therapy, I had earlier suggested that a center be established where a standard population of patients might be housed. The standard population would be specified with regard to age, sex, symptoms, duration of disease, age at onset, and other pertinent variables. This population of patients would serve as the proving ground for the relative efficacies of different types of treatment. In psychiatric centers throughout the country varieties of therapy could be tried out on comparable groups characterized according to the same variables. The outcomes of these therapies could then be evaluated and a definitive statement arrived at as to the efficacy of each of the treatments.

Thus far, the suggestion for a standard population made in 1950, has not caught on. Perhaps it is too much to expect physicians to keep patients 'on ice' without applying any of the apparently promising therapies. An alternative suggestion is to establish a Central Assessment Bureau to which patients could be referred before and after therapy for an assessment by the most promising instruments now available. At the present time, the evaluation of outcome is made by, to say the least, those in whom a conflict of interests is most likely to be present—the patient, his family and therapist. Such evaluations will no doubt continue to be made, but the provision of a neutral assessment agency, on a confidential basis, can bring about a most salutary influence in the long run. Such an assessment can at first be made on a purely advisory basis. As the agency continues to function over the years, actuarial material can be provided for the probable outcome of a given type of patient under a specified therapy. While such actuarial tables will never replace clinical judgment any more than life tables replace the clinician's prognosis for survival of a given case, they will provide guides for choice of therapy and give each therapist a batting average against which to compare his results. Such comparisons are practically impossible now.

It is interesting to note that in his writings Dr. Eysenck changes from therapeutic nihilism to emphasis on crucial experimental design. The change in point of view is salutary but experimental design without proper tools will be of no avail. The preparation and standardization of objective instruments and criteria for evaluation is the most urgent need of the moment. The confirmation or refutation of the efficacy of treatment is built on sand unless the tools are worthy of the effort.

REFERENCES