Part Two: Discussion I

By JOSEPH ZUBIN, Ph.D.

Reprinted from
RELATION OF PSYCHOLOGICAL TESTS TO PSYCHIATRY 1950
CRUNE & STRATTON, NEW YORK
Made in United States of America

Zubin, J. Discussion I. Part Two. In P. H. Hoch & J. Zubin (Eds.),
Relation of psychological tests to psychiatry. New York:
Crune & Stratton, 1950, 96-104.
A cursory review of the development of psychiatry indicates that long before the introduction of quantitative testing procedures by Binet, psychiatrists had used informal testing procedures to great advantage. Bleuler, for example, would often require his students to determine whether a patient was "intelligent" or "unintelligent," but such determinations were based on common sense methods only. With the introduction of intelligence tests, the psychiatrist was provided with a more stable basis for evaluating intelligence. It is, therefore, of great interest to see what the team of psychiatrist-psychologist represented by Drs. Dietheilm and Knehr have to say about the use of psychological tests in their clinic. They point out that psychological tests at the present time are used in cross-sectional studies primarily and therefore can be of value only in the field of diagnosis and only of incidental value in the field of therapy. This is, however, a criticism of the users of tests rather than of the tests themselves. There is no reason why suitable tests cannot be applied from time to time from the beginning of therapy, during the process of therapy, and after the end of therapy to determine the status of the patient with regard to important aspects of his personality as well as important details of his condition. Actually, such longitudinal studies of patients with mental disorder became possible with the introduction of fever treatment for general paresis, and have been applied since to most of the other available therapies. To be sure, the number of experiments in which such monitoring of the therapeutic process was made by means of tests are still few in number. The use of such tests for the evaluation of current somatotherapy is still new and there are many problems to be solved before the tests can be applied to the individual case with precision and accuracy. The patient cited by the authors who showed an actual gain in performance after a lobotomy for intractable phantom limb pain, is a case in point. Reduction of pain and anxiety had led to many an

* Psychiatric Institute, Columbia University.
improvement in psychotic patients undergoing psychosurgery. There is no reason why the intractable pain case should not show a similar rise. This gain, of course, is only apparent since, by removal of tissue, one can hardly hope to increase the level of performance. The assumption must be made that the initial performance was considerably below the actual capacity of the individual, and when the morbid pain was removed, a return to the pre-morbid level occurred.

Since the presence or absence of deterioration is of great consequence to diagnosis and treatment, many proposals have been made for measuring the presence of deterioration in patients by means of test-score patterns. Perhaps the most trenchant criticism of such patterns is that they can be no more reliable than the individual subtests on which they are based, in the same way that no chain can be stronger than its weakest link. The only way to establish more valid patterns for differentiating various degrees of deterioration is to develop more reliable tests in the first place. Another factor which interferes with measurements of patients’ progress is the fact that on repetition of a test, practice effects enter, which may themselves be influenced by the therapeutic process. For example, although it was possible to demonstrate that no losses occurred following operation in the Columbia-Greystone I patients, it was clear that a definite although not reliable lag in benefit from practice occurred in the operative cases. Just how to control this differential practice effect becomes a very important issue. It is apparently not due to lack of learning ability since experiments directed towards the specific testing of learning ability, not only failed to reveal losses, but instead, produced actual gains. Perhaps this failure to gain from practice may be attributed to the apparent loss of familiarity for learned material which many patients exhibit after the shock therapies and also after psychosurgery.

The use of the Rorschach which is by far the most popular test in the clinic today is subjected to severe scrutiny by the authors. It is the belief of most workers that the rather striking results of blind analysis of the Rorschach are no longer of either heuristic or scientific values. Instead, only open-eyed interpretations of Rorschach protocols including all the information that one can get from interviews and other tests is far more useful for an understanding of the patient. Though the usefulness of projective technics in the field of psycho-
pathology is limited, their use with normal human beings or in comparative culture studies is still less developed, since there are no objective criteria for validating the many personality characteristics which these tests allegedly reveal. This is unfortunate, since it sets off projective technics as a field apart from the rest of psychology without much communication with the rest of the science from which it sprang. It is to be hoped that as the theoretical formulation of the structure of personality and its variables are discovered, the proper place of projective technics in normal personality analysis will become more defined.

Regarding the use of the Minnesota Multiphasic Personality Inventory, and similar inventories, it is surprising to note that these tests are as useful as they have come to be. It had once been thought that the primary usefulness of such tests was for screening purposes only. With the increased usefulness of such tests in the clinic as aids in diagnosis and in evaluation of therapy, it becomes important to realize that they suffer from a basic though remediable deficiency. At the present time, scores on these tests credit only deviant responses; that is, a score is given only for those items which differentiate a given psychopathological group from the normal group. Items which the allegedly deviant individuals answer in a way similar to normals, are not counted. Furthermore, the construction of the test by the usual methods of item analysis eliminates such items entirely. This is unfortunate since both the liabilities as well as the assets of an individual are important in determining his status with regard to psychopathology or normality. The pressure of certain normal trends may offset or keep in check an abnormal trend. We already have available methods of pattern analysis which give the total response pattern of individuals. It would be well to develop these methods further to the end that a balanced indicator of the individual’s status would be available in which both his liabilities as well as his assets would be included.

It has sometimes been argued that psychological tests cannot help basically in evaluating disease processes or determining diagnosis because the disease by its very nature is dynamic, never stationary, whereas tests can only catch the momentary status of the individual and hence lack the dynamic quality which would make them useful in measuring a dynamic entity. This argument is valid insofar as
tests sometimes fail to vary comitantly with alterations in the disease status. Nevertheless, for any given moment, even the dynamic disease process is stationary, and a static representation of the traits and behavior of the individual at the moment in which he is tested may be of value in determining the subsequent course of the disease. For this reason, it is quite possible that despite the static nature of test scores, they may still be useful in the evaluation of the patient.

The importance of anxiety in producing deviations in performance in such tests as the Vigotsky and other conceptual thinking tests is, according to the authors, independent of the type of disease. They wonder, however, whether the influence of emotion on thinking in general is comparable in the schizophrenic and in the neurotic. Since it is possible to induce deficiencies in behavior under states of cerebral anoxia in normal individuals, it becomes even more apparent that underlying thinking deviation is probably a biochemical imbalance in the brain. That this biochemical imbalance is independent of specific tissue function is demonstrated quite clearly by the Columbia-Greystone I study in which removal of tissue actually enhanced conceptual thinking in some patients. The problem, as the authors state, requires much further analysis and will yield a clearer resolution when the biochemical functioning of the brain in neurotics, schizophrenics and normals becomes clear.

Regarding the problem of deterioration, it is important to realize that many individuals suspected of deterioration may be showing only apparent deterioration. It is necessary to differentiate between the senescing process which characterizes everyone with advancing age, and the deteriorating process, which characterizes only those individuals who suffer an accelerated senescence because of some organic or unknown cause. The problem is still further complicated by the fact that the individual curve of growth and decline is quite different from person to person, and only a longitudinal study of the single individual can reveal whether he is growing or declining at the proper biological rate, or whether he is exhibiting a sudden onset of deterioration over and above that to which he is entitled from the sheer process of senescence. This whole problem needs much further investigation. The studies which have already appeared indicate that there are many functions in which the senescent not only holds his own against the younger man, but actually excels him. On the other hand, there
are some functions in which the senescent individual seems to be below the younger man. Until proper age norms are developed the problem of deterioration will remain unsolved. Furthermore, the entire notion of deterioration in relation to schizophrenia has been disturbingly reopened by the recent results of psychosurgery. In the Columbia-Greystone studies, patients who show what is ordinarily called deterioration improve either after operation or after "total push" procedures. On the other hand, patients who do better on the tests and show less deterioration, or none at all, fail to benefit from psychosurgery or "total push." It seems as if the patients who performed better on tests are the ones who are actually "deteriorated" from the point of view of their illness. Individuals who are still struggling with their psychoses have a chance for improvement, whereas those who can perform well on the tests despite their psychosis may already have reached the point of no return in their chances for improvement. Another suggestion that has emanated from these psychosurgery findings is that conceptual thinking is more severely disturbed by emotional or conflicting elements than is perceptual observation of the immediate environment. It may be that patients who are conceptually confused but perceptually clear, are the ones who may eventually show an improvement, while those who are, on the other hand, conceptually clear but perceptually confused, have passed beyond the point of reversibility, at least under present methods of treatment and therapy.

The paper by Drs. Piotrowski and Lewis promises to become a landmark in prognostic work with psychological tests. It must be pointed out, however, that the study is not prognostic in the real sense of the word. True prognostic studies in which prediction is made at time of testing or examination are very rare in psychiatry. Most so-called prognostic studies are in reality retrospective studies in which the initial test scores or traits of the patient are analyzed in the light of subsequent outcome. It would be well to coin a new term for such studies—perhaps "hysterognostic" would be suitable—in order to differentiate them from studies in which the outcome is actually predicted at the time of testing, and not postdicted at the time of outcome. Two problems arise to vex the prognostician. First, the time interval for which the prediction is made. This is discussed by the authors and it is pointed out that their prediction for the
seven-year follow-up, based on their three-year hysterognostic study, was not very successful. (A statistical analysis of the actual outcome after seven years indicates that the prediction was not statistically significant.) The second point that needs to be borne in mind is that the test performance and the predictions based upon it depend upon the stage in the disease process where the test is applied. Heretofore, most prognostic studies have tacitly assumed that the better the score, the better the prognosis. Generally, this holds true, but several studies have appeared in the literature in which the reverse held true, poor test performance being prognostic of good outcome, while good test performance was prognostic of poor outcome. This was especially marked in the studies of O’Connell and Penrose (1) and Schnack, Shakow and Lively (2), who found a negative correlation between test performance and outcome for metrazol patients and a positive correlation for insulin therapy. More recently, the hysterognostic studies of the Columbia-Greystone I patients have indicated that in their case, too, a negative correlation existed between test performance and outcome. The hypothesis that suggests itself is that for recent mild cases such as those who are selected for insulin therapy, a positive correlation is to be expected, while for more chronic cases such as those who are selected for psychosurgery, a negative correlation may be found. Why poor performance should be related to good outcome remains an unsolved question. Perhaps the authors’ suggestion that the disparity between basic capacity and apparent ability is the measure by which prognosis can be achieved, may explain this dilemma. It may be that at a certain stage in the disease process, ability is depressed beneath capacity, but as the disease process continues, ability may return and approximate more the level of capacity exhibited by the patient. Another way of looking at it is that the disease process itself may set up emotional disturbance of such intensity that the patient is incapable of dealing with the problems and tasks which he would otherwise be capable of dealing with. It is quite possible that the patient who had a criterion score of minus 2, and nevertheless improved, belonged to the chronic group who are characterized by low scores but good prognosis. It is clear, then, that for a prognosis to be successful it is necessary to know the stage in the disease process at which the patient has arrived, and the interval for which the prediction is made. Thus, in subsequent validation
studies, it will be necessary to study groups similar to the present group not only in age, sex, diagnosis, etc., but also with regard to stage of disease process. Unfortunately, mere duration of hospitalization or interval since first symptoms appeared is not always trustworthy as an indication of the duration of illness. Furthermore, mere duration itself is not sufficient, since the disease process may advance at different rates in different patients.

The need for a cross validation study on similar patients becomes clear when it is realized that chance alone may often produce a statistically significant difference. Thus, if 100 indices had been compared in this study, one would expect 5 of them to meet the level of statistical significance even when no real relationship existed. The authors do not state how many comparisons they made before arriving at their final prognostic criterion. Only by cross validation can the truly significant be separated from the apparently significant.

Coming now to the actual indices developed by the authors for prognosis, it becomes quite clear that only by scoring rigorously according to the authors' system can any comparability be attained. Unfortunately, the degree of reliability of the authors' scoring system, while no doubt high for their immediate colleagues, has never been established in other clinics or other laboratories, and the discrepancies in scoring sometimes are quite high. Some type of standardization in this respect is urgently needed.

The authors have demonstrated that neither tradition nor clinical practice need bind one to any conventional approach when such questions as the prognostic validity of a test is raised. They, therefore, examined the Rorschach test from old as well as new angles and emerged with new approaches which seem to be more suitable than the old ones. It is necessary, however, to indicate that their new indices or signs derive from the old even though they may be superior to them. For example, inductive perception is probably a legitimate derivative of the DW or DdD scoring of Rorschach protocols, and adds to such scoring a rationale which makes it quite clear and more specific. Regarding sign no. 7, which relates the sum of the movement and color responses to the differences between these two types of responses, it is difficult to understand why such an innovation was necessary. Simple arithmetic would indicate that the new index is nothing more than the old well-known ratio for the Erlebnistypus. If
you let \( r \) represent this ratio, then the new sign proposed by the authors is identical with \( (r + 1)/(r - 1) \). Whenever such a direct relationship exists between the old and the new, the burden of proof for introducing a new index ought to fall on those who introduce it. There may be ample reasons for the needed change. Perhaps the difficulty of dealing with zero values for \( M \) or for \( \Sigma C \) lies behind this introduction. But this improvement is obtained at the cost of producing a ratio of infinity when \( M = \Sigma C \). It would be well for the authors to point out the basis for the change.

Another needed clarification ought to be made with regard to the change in differentiating power of each of the fifteen signs from the 3 year follow-up to the 7 year follow-up. Only the changes in per cent differentiation between the good and the poor groups are given. The statistical significance of these changes is not referred to. One wonders whether a change from 15 per cent to 21 per cent is sufficiently important to mention, without mentioning its significance.

In summary, the authors are to be congratulated on the splendid attempt at providing us with new indicators of the possible course of the disease in milder forms of schizophrenia. This is the group of patients that are now inundating our clinics and hospitals and it is very important to be able to separate them from those who have progressed much further in the disease process and for whom different types of therapies are essential. It is to be hoped that others will provide the necessary corroborative evidence to establish the validity of the proposed procedures.

Dr. Tompkins attempts to find relationships between such factors as the \( M \) response on the Rorschach and psychometric measures of intelligence. The chief purpose of this integration is to discover the reasons why some individuals, though possessed of a relatively high I.Q., fail to continue to develop their learning abilities or to achieve success in life while others maintain a continuous interest in learning new things throughout life. This problem is especially important in psychopathology since it is possible that the factor which determines continued growth and development through learning ability may be crucial for such psychopathological conditions as psychopathic personality on the one hand, and character neurosis of the inhibited variety on the other. The author then postulates that individuals of approximately the same intelligence or same history of past achieve-
ments differ in their learning ability depending on whether or not they had developed an “inner life.” The degree of inner life is to be measured by the M response. The degree of learning ability on the other hand is to be tested by an orientation test in which learning ability is involved.

The author concludes from the reliable correlation between the orientation test and the per cent of M responses in the Rorschach record that there is a relationship between inner life and learning ability. This is a suggestive hypothesis but it is quite apparent that the existence of the correlation between these two factors does not necessarily indicate the presence of an inner life in the individual as the primary factor. The correlation between the orientation test which is dependent upon perceiving a direction of the prow of a ship in two successive positions and the proportion of movement responses may both reflect a perceptual factor which may be entirely divorced from such considerations as inner life. It is also interesting to note that all attempts to provide a rationale for the meaning of M in terms of inner life or creativity have failed to produce positive results. This has held true not only of the M response on the Rorschach but of attempts at measuring the capacity for movement through such tests as the Levy movement blots.

REFERENCES