

RORSCHACH VALIDATION. I: A RATIONALE.

By J. D. KEEHN.

London University Postgraduate Scholar at Institute of Psychiatry, Maudsley Hospital.

THE Rorschach ink-blot test is so well known amongst present-day psychologists that any attempt to describe its composition and the aims and ambitions of its protagonists would be presumptuous. However, the underlying assumptions involved in the use of the test are neither so widely understood nor so universally accepted. It is for this reason that there have been so many and so various attempts to establish the validity of the Rorschach.

The vast amount of clinical material offered in support of Rorschach's original hypotheses has already been summarized by Hertz (1942) and by Bell (1948), and is growing constantly. So much material is available that it is almost impossible for any single worker to obtain a complete and comprehensive acquaintance with it. This fact in itself would seem to offer a *prima facie* case in support of the justification of the continued use of the Rorschach and interpretations from it, for in all this wealth of material hardly a dissenting word can be found.

In her historical summing up of validating procedures used in the Rorschach, Hertz (1941) draws attention to the initial rather chaotic introduction of the Rorschach, and the great advances which have arisen out of the publication of the various norms and standardization tables available in the major text-books on the subject. The actual use of the norms in the differential diagnoses between various clinical groups obtainable by the use of the Rorschach furnishes a good deal of the material appearing in Bell's bibliography (1948) mentioned above.

However, both Hertz (1941) and Cronbach (1949) have drawn attention to the statistical inadequacies of many of these studies, insufficient numbers of cases, tending often to highly selected samples, frequent misuse of statistical devices and misinterpretations of the results being the most common pitfalls. Cronbach (1949), for instance, points out at some length that many authors have accepted occasional significant "t" ratios between various groups at their face value, instead of taking into account the number of significant ratios that could arise by chance alone in that particular study. Welcome attention to this point has recently been paid by Meyer (1951), who, obtaining one significant ratio in twelve, instead of pointing to this as psychologically important, recognized that this could be just a chance occurrence.

Inattention to statistical and experimental procedures is not surprising when one remembers that the Rorschach has been mainly used by practical clinicians. The demands of experimental rigour have always in the clinic been secondary to those of clinical expediency. Remembering, too, that the clinician can take no account of adequate selection principles but must take his cases as they come, one can see how easy it is for him to build up a biased view, a bias that will later affect his own interpretations. Adding to this the further natural tendency only to write up and report successful cases, and to disregard those not fulfilling expectations, it is possible to see how gradually and insidiously an erroneous picture could be built up. In order, then, to decide whether the present Rorschach picture is an erroneous one or is in fact fully justified, one must turn to an examination of the more deliberately experimental attacks upon the problem.

The variety of techniques available for studying the validity of psychological tests has been previously classified by several writers—Sargent (1945), Hertz (1941), McFarlane (1942), and Ainsworth (1951). Extracting from these studies the following major categories,

1. Correspondence with other criteria,
2. Predictive success,
3. Experimental modification,
4. Internal consistency,

let us examine some of the typical studies included in them and consider why such studies should have arisen.

One of the earliest and, on the surface, most obvious ways of testing the validity of a holistic test such as the Rorschach is known as the matching method. Devised and quantified by Vernon (1936), this type of study whereby Rorschach records are matched with case-history notes by impartial and independent judges enjoyed a great deal of success, studies by Vernon (1935*b*), Krugman (1942), Troup (1938) and Munroe (1942) giving contingency coefficients in the region of .8. However, more recently the value of this method has been brought into question.

Firstly, general acceptance is accorded to the point that even small and specific mannerisms appearing on a test record could, in the hands of skilled judges, be sufficient to form the basis of a correct match. For instance, difference in language, both of content and expression, could enable almost anyone to say which of two records was that of a normal university graduate and that of a builder's labourer. As these external facts would appear on the case-notes it is clear that matching could be done quite independently of the more relevant Rorschach indices. Were very large numbers of pairs matched simultaneously, it might be argued that these external clues could only play a very minor role, but as, on the average, the judge has only five protocols and sets of notes to match at any one time, this point cannot so easily be explained away. Further, even without the help of extra-test cues, such degree of agreement would depend largely upon the sample chosen, for unless the cases were selected entirely at random there might be a very natural tendency to include only those cases exhibiting so-called typical Rorschach protocols, with the natural result that any judge aware of the expected signs will achieve a high degree of correct matching.

There are other objections to this method, notably the range of skill of the judges themselves, which would enable some judges to make more use of the relevant clues than others, and the unreliability of the case-notes as pointed out by McFarlane (1942). Both these objections fall into a rather different class from the first ones mentioned, for whereas they could explain the superficially high degree of correct matching in the hands of competent judges, these would tend to produce the opposite. Were all unskilled judges used it would be expected that the degree of matching would fall to chance level if they ignored the protocols altogether, or less if some other personal, irrelevant criteria were introduced. Similarly, if the case-notes are unreliable, no better than chance matching would be expected. Thus the degree of agreement between Rorschach protocols and case-histories could range from almost complete agreement to wellnigh complete disagreement due to these variables alone. So the method is unreliable as well as invalid.

A variation of this technique is that usually attributed to Rosenzweig (1935). Generally known as the method of blind analysis, it is based upon direct comparison between Rorschach test protocols and clinical findings. In principle the subject is clinically examined by one investigator and given the Rorschach by a second. This Rorschach protocol is then scored and interpreted by yet a third investigator who had previously had no contact with the subject.

The study by Benjamin and Ebaugh (1938) may be cited as an example of this technique. They used 46 patients who were clinically examined by one author and tested on the Rorschach by the other. However, the following quotation from their paper admits of a very serious deviation:

"in 34 . . . cases included in the validity series proper, the test was administered personally by the interpreter. Although no remarks were exchanged beyond the usual brief instructions and subsequent review of answers, it was, of course, possible to receive diagnostic impressions which may, in turn, have influenced the interpretations."

In other words, they recognize the possibility that they may have ruined the design of the experiment and really only correlated together two sets of clinical opinions, the Rorschach test being used solely as a method of "interview" by one of the clinicians.

Of the remaining 12 patients with whom the experiment was conducted properly and whose protocols were interpreted "blind," they claim *complete agreement* between the interpretation and final clinical diagnosis in all but one. However, it must be remembered that by "complete agreement" is meant simply that the authors express satisfaction that the words used in one description have the same

meaning as those used in the other. That such satisfaction would be evinced by other observers cannot, of course, be guaranteed. This point must be borne in mind when evaluating all of these studies.

In an addendum to this article Piotrowski makes the point that the Rorschach is primarily a test for obtaining personality descriptions, and is not meant to be used in lieu of a diagnosis—a point more recently made by Ross (1950). Piotrowski (1938) goes on to argue that as the validity of a diagnosis would necessarily be lower than that of a description, the diagnosis being made on the basis of the description, then taking the Rorschach for what it is meant to be, it gives an even more valid personality description than the results of Benjamin and Ebaugh suggest.

Further similar studies on single cases were carried out by Miale, Clapp and Kaplan (1938) and by Piotrowski (1937–8). Qualitative comparisons of interview and test interpretations showed fairly high degrees of agreement, and the authors expressed their confidence in the validity of the test. Using the same technique, Brussel and Hitch (1942) reported essential agreement between Rorschach and clinical interviews in 98 per cent. of their 50 cases. Michael and Buhler (1945), using cases said to be more difficult to diagnose, obtained 70 per cent. agreement, and the same degree of agreement was obtained by Garfield (1947), who also used cases which had proved difficult to diagnose.

However, in many of these cases the authors frankly admit that diagnoses were made with the aid of the Rorschach protocols, Garfield (1947) even pointing to this as favourable to the Rorschach in that it must have had support from outside sources. Again, investigators were not always careful to ensure that contact between patient and interpreter was avoided, thus introducing yet another source of error. In the studies where these factors had been adequately controlled the number of cases have been too few to allow of important generalizations.

In all these instances the infallibility of the psychiatric diagnosis is implicit. Even where such unreliability is recognized, it is explained away as being a combined staff decision and so, presumably, reliable, Garfield (1947), but Ash (1949) has emphasized the danger of using this criterion by showing the surprisingly low degree of agreement between psychiatrists even when patients are jointly interviewed, thus giving each psychiatrist the benefit of questions asked by his colleagues.

To add to the confusion a rather different technique, concerned more with the reliability than the validity of the test has also been labelled "blind analysis." This, instead of being concerned with matching protocols and case-histories, is used to investigate the interpretations by different workers of the same protocol. One such study by Hertz and Rubenstein (1939) cites interpretations by Hertz, Beck and Klopfer as being almost identical in the case of a single subject. That such agreement must necessarily obtain for any test to be of value is too obvious to warrant discussion here. Such agreement is not always the case with Rorschach interpretations, but this is only a practical drawback and may conceivably be overcome with care, experience, and objectification.

Many writers nowadays, Hertz (1943), McFarlane (1942), Sargent (1945), Ainsworth (1951), point to the value of prediction as the most hopeful method of validating the Rorschach. However, there are certain inherent difficulties even here. As Ainsworth has already pointed out, the hypotheses relating to the Rorschach test are all concerned with inner dynamic processes, and as predictions can only be made in terms of observable behaviour these hypotheses would have to be re-stated in more behavioural terms. This is itself no easy task. Secondly, some criteria about which the prediction could be made would have to be set up, and as there are an infinite number of such criteria, it is unlikely that anything concerning the general nature of the test can ever be discovered this way. Using the criteria of recovery of schizophrenics after being given insulin, Piotrowski (1941), by studying Rorschach protocols, was able to predict almost completely successfully those who would improve and those who would not, being satisfied with his prediction in all but 7 out of 60 cases.

Of the two main studies where predictions of normal behaviour have been made either wholly or partially on the basis of the Rorschach, one shows it to be superior and the other shows it to be inferior to various objective tests. Munroe (1941, 1942, 1943), using standard administration combined with a partially objectified scoring system, in comparing the Rorschach with the Bernreuter, contends that the Rorschach is superior in predicting adjustment in college. As against this, Kelley and Fiske (1950) in predicting success in the V.A. Training Programme state that—

"The most efficient clinical predictions, in terms of both validity and economy of data, are those based only on the materials contained in the credentials file and in the objective test profiles. The addition of autobiographical and projective test data appears to have contributed little or nothing to the validities of the assessment rating. . . . In fact, the predictions based on the credentials and objective tests are better than those made at the end of the programme on the basis of all tests, procedures and observations."

So one competent authority finds the Rorschach better than objective tests for prediction of college success and another finds it worse, and, in fact, says that information gained from it tends to lower the validity of predictions. It may be that in one case more skill was used in interpreting the protocols, or it may be that the partial objectification of Munroe's scoring system accounts for her greater success, but either way we have no more information than we can rely upon about the validity of the Rorschach.

So in all of the studies in these categories no one is found that is not open to sufficient criticism to vitiate its conclusions.

What if these specific criticisms were overcome and crucial experiments were designed and faithfully carried out? This would not seem impossible, for the experiments involving the use, for instance, of the Rosenzweig "blind" analysis technique have been criticized in performance alone, and not in principle.

The principle is that of the whole versus the part.

So far we have concerned ourselves only with those aspects of the problem which have adhered to the concept of the Rorschach as a whole. The complete overall view of the test is taken and compared, in one way or another, with the personality as a whole. Indeed, this would appear to be the only way. Piotrowski (1937) has drawn a parallel between the Rorschach signs, the clues upon which interpretations are based, and a set of mathematical equations. Just as no equation has any meaning by itself, so no single Rorschach sign has any significance unless it is regarded within the whole framework of the test responses. The point is further emphasized by Piotrowski's (1938) statement that—

"since isolated aspects of personality have no meaning if considered separately from the total personality to which they belong, it is, of course, fundamentally wrong to attempt to validate isolated components of the Rorschach record with bits of behaviour. This cannot be too strongly emphasized."

Fortunately, this statement depends upon a questionable premise, for were it true the whole idea of personality study as a *science* would collapse. If it were correct to deny the possibility of isolating aspects of personality, then it would follow that the abstraction of common elements from diverse personalities would also be impossible, for they would have different meanings within their various contexts. In this case classifications of personalities would clearly be impossible and the basic requirements for scientific progress could not be met. Every single personality would be itself complete and unique, and no laws evolved concerning one could be generalized to apply to any other personality. This in itself may be the true state of affairs and would be no argument against the clinical use of the Rorschach for, as has been already demonstrated, it claims to do no more than give a personality description. If, however, this is the way things are, then we must give up all attempts to study personality scientifically and return to the ways of the arm-chair philosopher.

If we do this we find ourselves, as do the current Rorschach workers, talking in terms of "intelligence," "neuroticism," "emotion" and so on—talking, in fact, in the trait language of the scientist. So we find ourselves then in the awkward position of denying the possibility of scientific advancement in the field of personality, and at the same time using the terms and concepts which the scientist conceptualizes. The point, then, is that it is not theoretically impossible to defend the position that "isolated aspects of personality have no meaning," but that if this position is held two consequences must follow:

- (1) a new language of personality description must be developed;
- (2) the hopes of a science of personality must be abandoned.

Although adherents to this ideographic type of personality theory may admit of the second of these consequences, they most certainly do not respect the first, for nowhere do we find evidence of their renunciation of trait language.

Rightly or wrongly, then, it would seem better to depart from this axiom, and since the rest of Piotrowski's statement is made to follow from it, it too cannot be regarded as proven.

Other general arguments against the ideographic approach to personality could be raised, but would lead us too far afield. It is sufficient to turn to Ainsworth's (1951) remark that the—

“holistic approach bears testimony to the uniqueness of personality organization, and generally indicates that the projective technique “probably has something,” rather than providing evidence of the validity of the hypotheses upon which the interpretation is based. Correct matching is conceivable with only a few of the various hypotheses involved being valid, the rest being incorrect or inadequate.”

The point becomes, then, using Ainsworth's terminology, that in the practical use of the Rorschach test certain hypotheses are advanced to account for certain apparently deviant responses, and just as all scientific hypotheses must be isolated and separately tested, so too must be these hypotheses about the Rorschach. In emphasizing that many of the Rorschach assumptions have yet to be validated, Wittenborn (1949a) too makes just this point.

Consideration of the more recent experimental examinations of the validity of the Rorschach does show a greater emphasis upon individual signs than upon the test result as a whole. Thus Williams (1947) and Baker and Harris (1949) designed experiments to test specifically the value of the so-called indicators of stress. In Williams's study, the protocol was taken under conditions calculated to arouse stress at the time, whereas in the experiment by Baker and Harris tendency to breakdown had been measured previously against an independent criterion. In neither case were the Rorschach signs found to be very indicative of stress. Other “sign” studies have been carried out by Eichler (1951) on anxiety, Holtzman (1950) on shyness, Hunter (1939) and Altus and Thompson (1949) on intelligence, and so on.\* Lane (1948) did a novel study using hypnosis to test the validity of the movement responses. That this sort of analysis is likely to be the one that most workers will concern themselves with in the future seems most probable. Reviews by Benton (1950) and by Eysenck (1950) of recent work in this field lend support to this view.

Despite Piotrowski's objections mentioned above and apart from the arguments advanced herein against him, there is certainly authoritative encouragement for this line of attack. Rorschach himself says (1924):

“I consider it as being quite out of the question that even with great practice and experience, a certain and trustworthy interpretation could be obtained for the experimental record alone, *without calculation.*” (Italics mine).

By this he undoubtedly points to the immense importance of the individual signs and, as the validity of any calculations made from these signs must of necessity be no more valid than the least valid of all the signs entering into such calculations, Rorschach would surely have been most anxious to establish the validity of each of his various signs. This must not be taken to mean that each separate scoring factor must be taken individually and *in vacuo* and separately tested. This technique has already been criticized, and rightly so, on the grounds that it is irrelevant, for if, say, the Rorschach hypothesis is that factors A, B, C, D together give an indication of personality factor X, then it is not legitimate to expect each of these variables to do so alone. Certainly one would not expect this from any other test. No one, for instance, would say that any one of the sub-tests of the Wechsler in itself gave an accurate score of intelligence, and nobody would think of isolating such a sub-test and generalizing results obtained from it to the test as a whole. Many such studies have been done on the Rorschach, and have met with much criticism from those who are familiar with the test. However, there is no reason why such combinations of signs as are hypothesized to be fruitful should not be effected in the manner suggested and the validity of the composite score examined experimentally. The problem, then, becomes one of finding the clues in the Rorschach protocols, either singly or in combination, which give the greatest insights into the personalities of the individuals producing those protocols.

\* A more detailed discussion of such “sign” studies with particular reference to the colour category will be given in the second part of this paper.

A step forward in this direction can be made by examining those studies which have investigated the internal composition of the test. Originally, examining a test for internal consistency meant correlating the scores from individual items in the test with the result of the whole test. In this way it was determined which items contributed highly to the value of the test and which were of less value or even irrelevant. Sargent (1945), in drawing attention to this, criticizes the somewhat different technique used with projective tests, whereby two or more such tests are given to a group of subjects to determine the amount of agreement between the tests. However, it is not necessary to depart from the original conception and, in fact, a return to it is actually desirable. This is especially so in the case of the Rorschach, where the importance of composite scores is so often emphasized. Where, for instance, it is argued that the number of well-organized whole responses, the number of good forms, the type of content and so on must be taken into account in assessing intellectual level from a Rorschach protocol, it is implied that these types of response have something in common. This would suggest a factor analytic study of the Rorschach signs to see if, in fact, the signs to be combined do have a common factor. Such studies have been carried out by a number of workers, including Hsü (1947), Adcock (1951), Sen (1950), and Cox (1951), with the object of extracting factors from the test and identifying those factors with those obtained from other tests.

That such studies can be of great practical value can be seen by reference to an article by Hughes (1950). In this study some twenty-two Rorschach signs of neuroticism, schizophrenia and organic defect appearing on protocols of one hundred such patients were inter-correlated and the resultant matrix factor-analysed. By using a simple method of weighting the rotated factor loading Hughes was able to classify correctly some 82 per cent. of diagnosed organics, whilst only 1 per cent. of the non-organics were misclassified. Many of the signs most highly saturated with this factor were, in fact, those commonly recognized to be indicative of organic defect. However, the factor analytic method was, in this case, able to do more than just validate the empirical hypotheses, for by attributing degrees of importance to the signs corresponding to their factor saturations, Hughes was able to effect a better classification than could be obtained by treating each of the relevant signs as of equal importance.

More important, however, are the studies of Wittenborn (1949*a*, 1949*b*, 1950*a*, 1950*b*), who goes even beyond the point of extracting factors from the signs. He points out (1949*a*) that in deriving scores by adding responses in the usual way, such as totalling all the C's, M's, F's, etc., two assumptions are involved, namely:

- “ 1. That all the responses falling in a given category are similar in some behavioural aspect.
2. The psychological significance of responses falling in a given category is different in some respect from responses placed in other categories.”

The first is a truly basic question and one which had previously been completely overlooked. That, for instance, all the responses scored by location as W should be combined had never been questioned by anyone, even after item analysis had become a firmly established technique in the field of psychometrics. It is just this first assumption which is put to the test when an item analysis is carried out. The second assumption cannot normally be said to apply to the majority of tests, for usually the construction of a test involved item selection so as to reduce measurement to that of one trait only. In the case of the Rorschach, where scores of different aspects of personality are taken at the same time, it becomes essential to separate these scores into their most useful categories. Such separation has, of course, been done but, as Wittenborn (1949*a*) points out, the basis of this separation has never been adequately checked and remains only an assumption.

In first checking these assumptions Wittenborn (1949*a*) gave a group Rorschach to 247 students. Altogether presence or absence of 7 human movement and 6 colour-form and form-colour responses were noted and tetrachoric correlations carried out for all possible combinations. The expected result, that the intra-correlations between the 7 movement and the 6 colour responses respectively would be larger than the inter-correlations of the movement and colour responses, was not confirmed. Using this and other data, Wittenborn (1949*b*) factor analysed his results and extracted six factors, but reported no obvious connection with the usual method of grouping Rorschach responses. However, it would be unfair to condemn the standard

Rorschach on this evidence, as the procedure was far from normal in that both group administration and check list scoring were used.

In a later, similar experiment using standard Rorschach procedure on some 240 students and psychiatric patients, Wittenborn (1950a) was able to produce evidence confirming the usual Rorschach groupings, at least with respect to colour and human movement. He concluded that—

“ The practice of combining the number of colour responses into a total score which is interpreted differently from the human movement total score appears to be justified. Moreover, since the colour scores are related with each other, it is quite possible that the total colour response scores could bear an important degree of relationship with some other response score, e.g., a measure of some practically important feature of personality.”

So support is found for at least one of the normal scoring procedures, although no generalizations as to the behavioural concomitants can be drawn from these data.

To recapitulate, the first experimental analyses of the validity of the Rorschach attempted to correlate the overall picture obtained from the test with various types of overall personality pictures. This, having been found open to criticism, was followed by studies of the individual isolated signs. However, these in themselves have been criticized as being divorced from relevant Rorschach hypotheses. The next step, that of examining the internal construction of the test to see if the scores which are alleged to go together really do measure something in common, is just assuming prominence. The final step, then, is to examine closely the working concepts of the test, see which responses really do go together, and to correlate the emergent composite scores against predicted external criteria.

This sort of development has arisen out of the peculiar position accorded to so-called projective techniques, of which the Rorschach is one, within the whole framework of psychological testing. This position has to a large degree depended upon the ideographic outlook which already, in this paper, has been shown to be dangerously fallible. It is worth while now to turn to a more positive programme and see if and how the Rorschach and similar techniques can be fitted into a more scientifically hopeful nomothetic system.

One of the merits attributed to the Rorschach almost universally is stated by Hertz (1935) as follows :

“ Intelligence tests are criticized because they take little account of the concomitant emotional state of the subject. Personality tests are criticized in that they merely probe intelligence rather than the affective life of the individual. The Rorschach test, if reliable, would be unique in that it would at the same time reveal general intelligence level and expose the emotional traits, thus avoiding the errors ascribed to most ‘ personality ’ tests.”

From this it would seem quite legitimate to conclude that the Rorschach is both an intelligence test and a personality test rolled into one, and that there is no reason to expect more from a Rorschach than from, say, a Binet and a Bernreuter combined.

However, what the statement probably means is that the Rorschach, instead of measuring some specific hypothesized trait or other, in some way measures an interactional factor which by its very nature tells us more about a patient than do any of these other tests.

If in objective tests of intelligence and personality all “ elements ” of personality are operative, it would seem reasonable to expect that suitable modifications of the scoring of such tests would give us just as much information of the total personality as does the Rorschach. In fact Ainsworth (1951) has admitted that an experienced examiner giving an intelligence test can report much more than a bare I.Q.

These reports and generalizations, she suggests, are made intuitively and are only learned through experience, which is just about the same sort of statement as is made about the Rorschach.

Suppose we now look a little more closely into the *modus operandi* of the Rorschach. The basic assumption underlying Rorschach procedure must be that a complete personality description can be made from a combination of overt motor behaviour and verbal responses to unstructured visual stimulation. This particular principle is easily seen to be but one aspect of the more general statement that anything that

an organism may do at any moment is, in some way or another, indicative of what that organism will do at any other moment.

Among others, Wolffe (1949) recognizes not only this general principle, but also the most important extension of it, namely that although every single act and instance of an organism's life is predictive, some acts are more so than others. The important point, then, is to discover just those occurrences which will account for the largest amount of the variation between the behaviour of organisms.

This principle is precisely that followed by the original factor analysts who set themselves the task of finding the smallest number of variables to account for the largest amount of variation in mental phenomena.

Having admitted, then, that every act, both active and passive, of an individual can be diagnostic, the next step is to single out those acts likely to be the most useful. Lombroso, Kretschmer and others decided to concentrate upon physical characteristics (which may be regarded as passive instances of behaviour), and Kretschmer in particular has built up a very imposing psychological system on this basis.

There have been studies relating the voice and personality (Allport and Cantril, 1934), and artistic style and personality (Vernon, 1933). Many attempts have been made to study the diagnostic clues afforded by characteristics of an individual's handwriting (Binet, 1906; Cantril and Rand, 1934; Allport and Vernon, 1933; and more recently by Eysenck, 1948). All these studies, together with that of Vernon (1935a) who studied the relation of facial expression to personality, and many other similar investigations have, to a small extent, been able to show a relation between the variable under investigation and personality, but always the degree of correlation has been low and pragmatically sterile.

It is not difficult to see how the Rorschach test fits into this category, as also does any other form of mental test. As in the Bernreuter, the M.M.P.I., and others, the behaviour is motor and overt, so in the Rorschach the type of behaviour elicited is visual-verbal. The patient looks at the cards, forms his associations and transfers them into verbal response. The question is, then, can we infer from these verbal responses and other associated behaviour more than we can infer from the responses to all the other variety of personality tests. Can we, in fact, say anything about the patient's behaviour in general from his behaviour in this particular situation.

This is not the first time that the position of the Rorschach as a projective technique has been called to question. Of the previous writers, although both Schachtel (1942) and Bellak (1944) are mainly concerned with the ambiguity of the definition of the term "projective," Eysenck (1950) specifically attacks the "projective"-"psychometric" test dichotomy, showing them to be quantitatively (along certain dimensions) but not qualitatively different.

The great advantage of the Rorschach in this respect is that in being a relatively unstructured test it allows of a greater variety of behaviour than do those tests specifically designed to examine pre-selected traits. This advantage is at the same time its greatest drawback, for in allowing a larger variety of responses it becomes encumbered with a greater complexity of scoring. The classical solution to this dilemma has been to ignore as far as possible these scoring details and proceed upon some intuitive overall impression gained from them. In some hands this procedure has appeared to produce profitable results, but the pitfalls involved in such practice have been shown greatly to outweigh its advantages.

Here, then, is the point. We have at our disposal a test from which it is possible to get simultaneously several composite scores pertinent to the patient's personality. Those who have had most occasion to employ this test have so far been satisfied to intuit from these data the kind of personality with which they are dealing. The hypothesis advanced here is that this is not the most advantageous use of the test, nor is it in line with scientific methodology. One of the aims of scientific research is to describe a given phenomenon, i.e., personality, in as few a number of variables as possible. Hence the object of psychological testing should be to measure that variable or those variables which either singly or in combination will account for the largest amount of variation between individuals.

The procedure, then, is clear. The individual scores and hypotheses concerning the Rorschach must be isolated and put to the test. Those that prove satisfactory should be retained, and those not fulfilling their original promise must be eliminated. In this way clues upon which diagnostic judgments are made can be examined and their value assessed, and at the same time those aspects of the test accounting for the greatest differences in observed personality pictures can be isolated.



The author wishes to acknowledge his indebtedness to Dr. H. J. Eysenck for his helpful criticism and suggestions in the preparation of this article.

## BIBLIOGRAPHY.

- ADCOCK, C. J., "A Factorial Approach to Rorschach Interpretation," *J. Gen. Psychol.*, 1951, **44**, 261.
- AINSWORTH, M. D., "Some Problems of Validation of Projective Techniques," *Brit. J. Med. Psychol.*, 1951, **24**, 151.
- ALLPORT, G. W., and CANTRILL, H., "Judging Personality from the Voice," *J. Soc. Psychol.*, 1934, **5**, 37.
- Idem* and VERNON, P. E., *Studies in Expressive Movement*, 1933. New York: The Macmillan Co.
- ALTUS, W. D., and THOMPSON, G. M., "The Rorschach as a Measure of Intelligence," *J. Consult. Psychol.*, 1949, **13**, 341.
- ASH, P., "The Reliability of Psychiatric Diagnoses," *J. Abnorm. and Soc. Psychol.*, 1949, **44**, 272.
- BAKER, L. M., and HARRIS, J. S., "The Validation of Rorschach Test Results against Laboratory Behavior," *J. Clin. Psychol.*, 1949, **5**, 161.
- BELL, J. E., *Projective Techniques*, 1948. New York: Longmans, Green & Co.
- BELLAK, L., "The Concept of Projection," *Psychiatry*, 1944, **7**, 353.
- BENJAMIN, J. D., and EBAUGH, F. G., "The Diagnostic Validity of the Rorschach Test," *Amer. J. Psychiat.*, 1938, **94**, 1163.
- BENTON, A. L., "The Experimental Validation of the Rorschach Test," *Brit. J. Med. Psychol.*, 1950, **23**, 45.
- BINET, A., *Les révélations de l'écriture d'après un control scientifique*, 1906. Paris: Félix Alcan.
- BRUSSEL, J. A., and HITCH, K. S., "The Rorschach Method and its Uses in Military Psychiatry," *Psychiat. Quart.*, 1942, **16**, 3.
- CANTRILL, H., and RAND, H. A., "An Additional Study of the Determination of Personal Interests by Psychological and Graphological Methods," *Charac. and Person.*, 1934, **3**, 72.
- COX, S. M., "A Factorial Study of the Rorschach Responses of Normal and Maladjusted Boys," *J. Genet. Psychol.*, 1951, **79**, 95.
- CRONBACH, L. J., "Statistical Methods Applied to Rorschach Scores: a Review," *Psychol. Bull.*, 1949, **46**, 393.
- EICHLER, R. M., "Experimental Stress and Alleged Rorschach Indices of Anxiety," *J. Abnorm. and Soc. Psychol.*, 1951, **46**, 344.
- EYSENCK, H. J., "Neuroticism and Handwriting," *ibid.*, 1948, **43**, 94.
- Idem*, "Personality Tests: 1944-1949" (ed. G. W. T. H. Fleming), *Recent Progress in Psychiatry*, 1950, 2.
- GARFIELD, S. I., "The Rorschach Test in Clinical Diagnosis," *J. Clin. Psychol.*, **3**, 1947, 375.
- HERTZ, M. R., "The Rorschach Ink-Blot Test: Historical Summary," *Psychol. Bull.*, 1935, **32**, 33.
- Idem*, "Validity of the Rorschach Method," *Amer. J. Orthopsychiat.*, 1941, **11**, 512.
- Idem*, "Rorschach: Twenty Years After," *Psychol. Bull.*, 1942, **39**, 529.
- Idem*, "The Rorschach Method: Science or Mystery," *J. Consult. Psychol.*, 1943, **7**, 67.
- Idem* and RUBENSTEIN, B. B., "A Comparison of Three 'Blind' Rorschach Analyses," *Amer. J. Orthopsychiat.*, 1939, **9**, 295.
- HOLTZMAN, W. H., "Validation Studies of the Rorschach Test: Shyness and Gregariousness in the Normal Superior Adult," *J. Clin. Psychol.*, 1950, **6**, 343.
- Hsü, E. H., "The Rorschach Responses and Factor Analysis," *J. Gen. Psychol.*, 1947, **37**, 129.
- HUGHES, R. M., "A Factor Analysis of Rorschach Diagnostic Signs," *ibid.*, 1950, **43**, 85.
- HUNTER, M., "The Practical Value of the Rorschach Test in a Psychological Clinic," *Amer. J. Orthopsychiat.*, 1939, **9**, 287.
- KELLY, E. L., and FISKE, D. W., "The Prediction of Success in the V.A. Training Programme in Clinical Psychology," *Amer. Psychol.*, 1950, **5**, 395.
- KRUGMAN, J. I., "A Clinical Validation of the Rorschach with Problem Children," *Rorschach Res. Exch.*, 1942, **6**, 61.
- LANE, B., "A Validation Test of the Rorschach Movement Interpretation," *Amer. J. Orthopsychiat.*, 1948, **18**, 292.
- McFARLANE, J. W., "Problems of Validation Inherent in Projective Methods," *ibid.*, 1942, **12**, 495.
- MEYER, B. T., "An Investigation of Color Shock in the Rorschach Test," *J. Clin. Psychol.*, 1951, **7**, 367.
- MIALE, F. R., CLAPP, H., and KAPLAN, A. H., "Clinical Validation of a Rorschach Interpretation," *Rorschach Res. Exch.*, 1938, **2**, 153.
- MICHAEL, J. C., and BUHLER, C., "Experiences with Personality Testing in a Neuro-psychiatric Department of a Public General Hospital," *Dis. nerv. Sys.*, 1945, **6**, 205.

- MUNROE, R., "Inspection Technique," *Rorschach Res. Exch.*, 1941, 5, 166.
- Idem*, "An Experiment in Large Scale Testing by a Modification of the Rorschach Method," *J. Psychol.*, 1942, 13, 229.
- Idem*, "Use of the Rorschach in College Counseling," *J. Consult. Psychol.*, 1943, 7, 89.
- PIOTROWSKI, Z. A., "The Reliability of Rorschach's Erlebnistypus," *J. Abnorm. and Soc. Psychol.*, 1937, 32, 439.
- Idem*, "Blind Analysis of a Case of Compulsion Neurosis," *Rorschach Res. Exch.*, 1937-1938, 2, 89.
- Idem*, in discussion of paper by Benjamin and Ebaugh, *Amer. J. Psychiat.*, 1938, 94, 1176.
- Idem*, "The Rorschach Method as a Prognostic Aid in the Insulin Shock Treatment of Schizophrenics," *Psychiat. Quart.*, 1941, 15, 807.
- RORSCHACH, H., and OBERHOLZER, E., "The Application of the Interpretation of Form to Psychoanalysis," *J. nerv. ment. Dis.*, 1924, 60, 225.
- ROSENZWEIG, S., "Outline of a Co-operative Project for Validating the Rorschach Test," *Amer. J. Orthopsychiat.*, 1935, 5, 121.
- ROSS, W. D., "Rorschach Interpretations and Clinical Diagnosis," *J. Proj. Tech.*, 1950, 14, 5.
- SARGENT, H., "Projective Methods: their Origin, Theory and Application in Personality Research," *Psychol. Bull.*, 1945, 42, 257.
- SCHACHTEL, E. G., Review: "The Rorschach Technique by Bruno Klopfer," *Psychiatry*, 1942, 5, 604.
- SEN, A., "A Factorial Study of the Rorschach Test," *Brit. J. Psychol.*, (Stat. Sect.), 1950, 3, 21.
- TROUP, E., "A Comparative Study by Means of the Rorschach Method of Personality Development in 20 Pairs of Identical Twins," *Genet. Psychol. Monog.*, 1938, 20, 461.
- VERNON, P. E., "Some Characteristics of a Good Judge of Personality," *J. Soc. Psychol.*, 1933, 4, 42.
- Idem*, "Can the Total Personality be Studied Objectively?" *Charac. and Person.*, 1935a, 4, 1.
- Idem*, "The Significance of the Rorschach Test," *Brit. J. Med. Psychol.*, 1935b, 15, 199.
- Idem*, "The Matching Method Applied to Investigations of Personality," *Psychol. Bull.*, 1936, 33, 149.
- WILLIAMS, M., "An Experimental Study of Intellectual Control under Stress and Associated Rorschach Factors," *J. Consult. Psychol.*, 1947, 11, 21.
- WITTENBORN, J. R., "Statistical Tests of Certain Rorschach Assumptions: Analysis of Discrete Responses," *ibid.*, 1949a, 13, 257.
- Idem*, "A Factor Analysis of Discrete Responses to the Rorschach Ink-blot," *ibid.*, 1949b, 13, 335.
- Idem*, "Statistical Tests of Certain Rorschach Assumptions: the Internal Consistency of Scoring Categories," *ibid.*, 1950a, 14, 1.
- Idem*, "A Factor Analysis of Rorschach Scoring Categories," *ibid.*, 1950b, 14, 261.
- WOLFLE, H. M., "A Fundamental Principle of Personality Measurement," *Psychol. Rev.*, 1949, 56, 273.