Reprinted from Character and Personality,
Vol. XI, No. 4, June, 1943
PSYCHOMETRIC SCALES FOR SCORING RORSCHACH TEST RESPONSES*

JOSEPH ZUBIN, ELOISE CHUTE, AND SEYMOUR VENIAR

New York State Psychiatric Institute and Hospital

All attempts at objectifying the Rorschach test meet with criticism from two opposing camps in psychology. The Rorschach workers on the one hand claim that it cannot be done, that straight-jacketing the Rorschach into a psychometric uniform would eliminate its value as a test of the total personality. The students of visual perception and the psychometricians on the other hand point out that the stimulus in the Rorschach cards is so varied and complex that it would be better to discard the use of the Rorschach cards and build anew from simpler material. While the latter proposal has merit, its adoption would preclude the utilization of the vast amount of experience already gained by Rorschach workers during the past two decades. In order to retain the Rorschach in its traditional form and yet satisfy the criticisms of the psychometrician, it is necessary to provide a scoring system which lends itself more readily to statistical treatment. By means of such a system the reliability and validity of the Rorschach test could be readily established and the usual comparison between tests facilitated.

Such an approach was recently formulated (8). The primary emphasis in this new approach was psychometrically rather than personalistically oriented, the purpose being to present an objective frame of reference for classifying Rorschach responses, rather than a system for differentiating personality organizations of different individuals. It was hoped that if an objective frame of reference were made available it would facilitate the validation of the claims of Rorschach workers regarding the alleged correlations between certain scoring categories or constellations of such categories and personality traits or total personality organization. To this end an

*The authors are indebted to Dr. S. J. Beck, Dr. Arthur L. Benton, Dr. M. M. Bolles, and Dr. Harry Helson for reading the manuscript of this article and making many valuable suggestions.
analysis was made of all the Rorschach factors, both recorded and unrecorded, which the clinician utilizes in developing the personality sketch from the raw material provided by the responses. The basis of this analysis was the recorded literature on orthodox Rorschach scoring and interpretation, the growing literature on "signs" in Rorschach records,\(^1\) the demonstrations of blind analyses by Rorschach experts and the writers' own experience in arriving at a personality picture from a combination of both objective and subjective estimates as well as from intuitive hunches. This analysis led to the development of some forty-nine different scales and classification categories. Not all of these factors are of equal importance for the analysis of a given case, nor do these factors exhaust all the possible trends or tendencies which Rorschach workers have distinguished. They are, however, all the trends which have loomed up as significant in the several hundred records which have been carefully examined during the course of the last four years.

Before presenting the results of this investigation it might be well to explain what possible gain might accrue from the application of such a large number of scoring categories and why it is necessary to introduce such complexity when the simpler scoring categories of the orthodox Rorschach seem to satisfy most present-day needs.

While psychometricians have always been skeptical of the crude elementary type of scoring that has grown up about the Rorschach, they have nevertheless tried to deal with them in statistical fashion. All such efforts have generally failed, not necessarily because of the invalidity of the Rorschach test, but because the face value of the scores does not always serve as the real basis for the interpretation, but merely as the jumping-off-point for the interpretation. Consequently, before we can hope for a significant relationship between the psychometric and orthodox approach, we must find a way of recording the additional factors which the Rorschach expert uses in preparing his interpretation. The scales presented in this study make the first attempt in that direction.

Furthermore, a growing tendency has arisen to claim that the function of the Rorschach is not to describe or predict behavior, but to describe the "personality structure" of the individual, which underlies his behavior and which may express itself in various behavior patterns depending upon the stresses and strains of the

\(^{1}\) By "signs" is meant a specific type of score or critical value of the frequency of a score which is claimed to be indicative of a personality trait or of a diagnostic category, e.g., color enumeration as a "sign" of organic disorder.
avironment. This attempt to explain the frequent absence of correlation between observed behavior and predicted behavior by means of a theoretical construct which can never go wrong, belongs more suitably to the medieval past in personality research. It is far better to follow the accepted scientific procedure of judging a test by its empirical usefulness in description and prediction of behavior. Our intelligence tests are accepted on this basis and not because they measure the hypothetical mental faculty known as intellect. Similarly, personality inventories have been invented to predict whether a given man will succeed as a soldier, and aptitude tests are used for prediction of vocational success in a given occupation. Each test has been devised for doing a specific job—and when that is the case, the test is usually found satisfactory. What is the job that the Rorschach is supposed to accomplish? Rorschach claimed that his test was diagnostic for mental disease categories. In order to make the Rorschach into a functional test in this area, we need to quantify or categorize the trends that have been delineated in Rorschach responses and then compare two contrasted groups or diagnostic categories on the basis of these trends. The differential pattern which separates psychoneurotics, for example, from normals may bear little or no relationship to the pattern which differentiates psychoneurotics from schizophrenes. Each type of differentiation may depend empirically upon different patterning of Rorschach factors. After the differential patterns characterizing many contrasted groups have been developed, we may find a connecting link in some hypothetical framework, but the proof of the utility of the test will depend not on the elegance of our theory, but on the precision of our differentiation and prediction. In order to make possible such a series of analyses, it is necessary to provide a universal framework for scoring Rorschach responses which will contain well-defined categories, quantified wherever possible, categorized where quantification still eludes us, which will have an objectivity in scoring comparable to those obtained in other psychometric scales.

The Problems of Classification

The earlier paper cited above developed a tentative series of rating scales for classifying the characteristics of Rorschach responses, the reliability of which was subsequently tested. Each of the authors independently scored a series of responses, and the degree of similarity among the judges in the individual ratings for each response was computed for each scale. These reliability coefficients
were rather high for some scales and somewhat low for others. The scales having a low reliability were investigated to determine the cause of their unreliability, and they were then modified in an attempt to eliminate the discrepancies. Some differences between judges persisted even after a careful revision, because of unavoidable random error in ratings, and because of the individual predilections of judges who may be inclined, for example, to weight the shading factor or the movement factor more than some other factors. This predilection of the scorer has already been noted by Rorschach himself (3) and is regarded by him as a reflection of the personality of the scorer. Another source of discrepancy and one which accounted for a large proportion of the disagreement among judges was the ambiguity in the response itself. Some responses do not permit unequivocal analysis even when considerable effort is devoted to their clarification by the subject. The purpose of this paper is to present the revised scales and point out their advantages as well as their limitations. The result of the experiment in determining the reliability of these scales is reported elsewhere (4).

In approaching the problem of classifying the Rorschach responses we must note, first of all, that the stimulus areas in which the different responses are located are not uniform in their attributes (color, shading, contours, etc.). The ten ink-blots forming the Rorschach series differ radically from each other in the distribution of these attributes, and considerable variation exists between the constituent areas of the same blot. It is, therefore, important to note the objective characteristics of each area in which a response is located.

A second point to be noted is that the responses that arise within the same stimulus area often differ considerably in their content and general characteristics. We must, therefore, study on the one hand the objective characteristics of the stimulus area that gave rise to the response and on the other hand, the particular type of interpretation which the subject finds for the stimulus area. An example will make this clear.

Let us select an area of one of the cards to which many people respond—the red spot at the bottom of Card II. The basic attributes of this area are color, form, and shading. The presence of these three attributes can be established by a consensus of opinion of observers or by some physical measures appropriate to each of them. Having determined the presence of these attributes, we can analyze each response to this area in accordance with the relative contribu-
tion of each of these attributes to the formation of the percept. Some responses like "butterfly" are found to be primarily influenced by the outline of the area and to a much lesser extent by color and shading. Other responses like "belching red flames" depend more upon color than upon the other two attributes. Still other responses like "splash of water" depend primarily on shading. The relative influence of each attribute can thus vary from zero to a maximum. This differential selection of attributes in the formation of the response is the warp and woof of Rorschach interpretation. It is the individual differences in this selection process on which Rorschach typology is based.

But the individual differences in response to the same stimulus area do not end in the differential selection of attributes. Even when the relative importance of the three attributes is the same for two responses, the interpretation which the subject puts upon them is not always the same. Thus, "a belching flame" and "sunset in back of the mountains" may represent similar constellations so far as the basic attributes go, but the first has a dynamic quality which the second lacks. Consequently, there are at least three different levels on which the Rorschach responses may be classified. They may be classified, first, according to the objective attributes of the stimulus area in which the response is located; second, according to the relative importance of each of the objective attributes in the formation of the response; and third, according to the manner in which the attributes were interpreted.

This distinction between the attributes of the stimulus area and the interpretation based upon them is a rather fundamental distinction, not made by the orthodox Rorschach procedure and, therefore, in need of further explanation. Perhaps the simplest way to clarify the distinction is through an analogy. If we were to expose the letters NOTE as a unit, and ask for the meaning of this word we would probably get considerable variation in the responses of a group of subjects since it has at least seven different possible meanings. If, however, we were to interpose a reduction screen which would narrow the view to one letter at a time, we should get considerable uniformity in the responses, most of the subjects giving the names of the letters as they appear. Similarly, when we expose a Rorschach card to a group of people we note considerable variation in their responses. If, however, we interpose a reduction screen exposing only a small section at a time, the responses become much more uniform and are largely in terms of the simple properties of
the card—color, line, and shading. These may be regarded as the framework from which the structured response is made in the same way that the letters constitute the scaffold of the word-meaning. These simple properties of the blot will be designated as the determinants, while the meaningful responses arising from them will be designated as interpretations. We do not consider the perception of determinants as more basic than the perception of the interpreted object. On the contrary, the perception of the object usually arises without the mediation of the perception of the determinants. It is only by subsequent analysis that we can find the presence of the determinants in the stimulus area. But this does not lessen their importance in the consideration of the formation of the response.

The distinction between determinants and interpretations may be viewed statistically as one based on the degree of agreement of competent observers. Thus, the majority of observers will agree when questioned that a certain stimulus area has color, line, contours, and shading. It is true that some observers may overlook these determinants, but they can be readily made to acknowledge their presence. These factors may then be regarded as the basic framework upon which the response is built. They may be regarded as more objective in the sense that they may be detected by purely mechanical means, and need not necessarily depend upon the testimony of observers. These determinants show less variation from observer to observer than some of the other categories to be discussed. The second category, that of interpretations, includes such factors as perception of varying degrees of movement, structure, and surface appearance. These are not as palpable as the first category of determinants, and they are designated as interpretations because they indicate the manner in which the basic determinants are interpreted to form the completed percept. There is no one-to-one correspondence between the two categories. Thus form, which is usually the basis for the interpretation of dimensional structure, may also serve as the basis for the interpretation of movement and of surface appearance. Similarly, color may not yield an interpretation involving surface appearance only, but it may also serve as the basis for the interpretation of movement or dimensional structure. Shading, too, can serve as the basis for the interpretation of either surface appearance, or dimensional structure, or movement.

There are five determinants and three methods of interpretation. Despite these rather fundamental innovations, the new scoring
method proposed here leans heavily on the standard scoring procedure which the Rorschach workers now use. In order to allow for the innovations, it must differ from it in several essential characteristics. Whereas the orthodox procedure is satisfied with recording the major location or determinant of the response, the present method attempts to evaluate the relative importance of each possible factor in the resulting response. Thus, the response of bat to Card I is usually scored as W (whole) F+ (form) A (animal) P (popular) by the orthodox Rorschach workers. In the present scale the degree of importance of each of the possible factors affecting the response is included. By careful questioning, the importance of the details (wings, etc.), the role of the black appearance, whether or not the bat is alive, and many other factors may be noted. A rather lengthy inquiry is required for this purpose, and a careful weighing of the evidence for and against each factor must be made.

The reason for this departure from the orthodox procedure may not be immediately apparent. The orthodox Rorschach scorer is in reality not satisfied with simply recording the major location or determinant. In practice he, too, notes the secondary locations or determinants but keeps them in mind without always formally recording them. In the final analysis of the record these unrecorded data and impressions are brought to the fore and integrated into the interpretation. In fact, it is this practice, which goes under the general name of "considering the total response" (rather than any isolated element in it), that makes it so difficult for the psychometrically trained psychologist to follow the presentations of Rorschach experts. It is hardly necessary to point out the danger of "forgetting" the factors that do not jibe with the apparent general trend in the emerging personality interpretation. By providing a method for recording the secondary as well as the primary scores, the danger of remembering only the "congruent" material will be lessened. The interpreter may still be free to sift his material, but he can then do it by outright selection rather than by unconscious forgetting. Another argument in favor of including the secondary as well as the primary scores arises from the fact that when the number of responses is small, the secondary scores help to provide sufficient data to make the results more reliable. A further point of departure is the attempt to provide concise objective definitions for each of the factors. Indeed, without such clear-cut definitions, the attainment of unanimity and uniformity regarding scoring would be impossible.
DESCRIPTION OF SCALES

We shall now present each of the scales with their definitions and subdivisions. An application of these scales to one Rorschach test response is given in the appendix.

1. Locale (W, D, Dr)

The stimulus area—the area of the blot in which the response was located—may be either the entire card, a well-defined portion of it, or some small unusual portion of it. Rorschach workers have already studied the geography of the blots and have published norms for classifying responses occurring in each portion of the blot as either a (W) whole, (D) detail, or (Dr) rare detail response. Of the various systems that are now available, we have selected that of Beck (1) because of the rather large number of scored responses provided by his book.

An examination of Rorschach responses indicates that only a very small number of D and Dr responses are confined to that portion of the blot which is now scored as their locale. The rest of the card often contributes a background quality which is indispensable for the emergence of the response. On the other hand, whole responses differ in the degree to which they utilize the constituent D's and Dr's of their stimulus area. Thus, although a response may be primarily a W response, D's and Dr's may be important in its formation to a varying degree. Similarly, responses that are primarily D's may owe their existence in part to the presence of Dr's as well as to the totality of the blot (W).

In rating a response for the degree of importance of W, D, and Dr in its formation, a given locale receives a rating of 1 when it serves as a background to the primary locale. When it is more than background, but not of major assistance, it is rated 2; when it is of major assistance in the formation of the response, it is rated 3. When it is the primary locale, it is rated 4. Frequently a negative approach is an aid in determining the importance of a locale. The question is asked whether the response would have occurred had the given section been omitted.

It is important to note that a response may take in the entire card and still W may not be the primary response. Furthermore, responses located in whole or in part in the white space are to be rated on the W, D, Dr continuum. They are also rated for the importance of the white space in the formation of the response in another scale (Scale 4).
I-3. Importance (W, D, Dr)

0. Of no importance whatsoever. Totally ignored.
1. Of some minor importance, but verbal evidence of its importance given neither in the inquiry nor in the original response. (Acted as a sort of background in a neutral way.)
2. Of a secondary (as opposed to primary) importance as evidenced by inquiry or inferred directly from original response (perceptibly helped in the determination).
3. A most important locale in determining the response as evidenced by mention in the original response.
4. The most important locale.

If two locales tie for "most important," both should be rated 4.

4. White Space (S)

The stimulus area may consist of either filled-in space or white space or a combination of the two. In addition to determining the importance of each locale (W, D, Dr) in the stimulus area, the importance of white space (S) must also be gauged. The white space rating is independent of the W, D, Dr ratings. Excluding the borders, except when they constitute a significant part of the stimulus area, the white space is rated in accordance with its importance. When the response would have been impossible if there were no white space as such, the rating for S is 4. Gradational ratings are given for white spaces of less importance.

0. Of no importance whatsoever. Totally ignored.
1. Of some minor importance but verbal evidence of its importance given neither in the inquiry nor in the original response.
2. Of a secondary importance (as opposed to primary importance) as evidenced by the inquiry or inferred directly from the original response.
3. A most important locale in determining the response as evidenced by mention in the original response.
4. The most important locale.

5. Type of White Space (Ts)

In addition to its importance in producing the response, white space may be used either as figure or as ground. The usual approach to the cards is to regard the white areas in the cards as ground and to locate the figure in the filled-in areas. This tendency does not characterize all individuals. Some locate the figure in the white space and consider the filled-in area as the ground. This bipolar tendency is gauged by the type scale for white space.

x. White space is neither figure nor ground. (Not a part of the stimulus area used in the response.)
a. White space serves as vague background (always ground).
b. White space in conjunction with other portions of the stimulus area serves as figure.
c. White space alone serves as figure.

6. Inclusiveness of Primary Locale (Inc)

After rating the relative importance of W, D, and Dr in a given response and noting what contribution, if any, the white space made toward the response, we may turn our attention to the completeness with which the primary locale was utilized. Some responses include nearly all the subdivisions and general attributes of the primary locale. Other responses fail to take into consideration either major or minor portions of the primary locale. Sometimes even rare detail responses do not include the entire geographical area in which they are located. Since edge details such as profiles represent a complete cutting off of the inside of the stimulus area and a use of the contour only, they are given the lowest rating on inclusiveness. Next in rank is the inside detail, where an island is selected within a large or small detail. The highest rating is given to responses that utilize the entire area in which the primary response is located.

0. Only edge detail used (rest cut off or background).
1. Only inside detail used (rest cut off or background).
2. Some selection, a “D” cut off.
3. Some selection, only a “Dr” cut off.
4. Entire area used (nothing cut off).

II. Organization

The extent to which the subject structures the vague material presented in the ink-blot is called organization. There are two aspects to organization—mode and type.

7. Mode of organization (Mz)

In this paper a modification of the Z score of Beck is used. As a result of a study of the modes of organization, Beck formulated four types of organization. Since these have been discussed in full previously (8), they will not be repeated here. The scale for mode of organization will be found in the afore-mentioned article.

8. Type of organization (Tz)

Sometimes the response appears full blown without yielding any evidence of how it was organized. At other times the successive stages in the formation of the response are quite evident. The following scale attempts to describe the process.
x. Response characterized by a crude undifferentiated structure involving only one undifferentiated area of the blot (x is also used where the importance score for Mz is 0 or 1).

1. Response begun by some suggestive detail and laboriously built up by slowly adding adjacent or nonadjacent areas (the SW response of Rorschach).

2. Response quite apparently built up by successive association to different parts of the blot.

3. Some slight evidence of the elaboration of a response from successive associations to different parts of the blot.

4. Sudden emergence of the completed response.

III. Determinants

(F) The form determinant refers to the lines and contours that occur in the stimulus area.

(Sh) The shading determinant refers to the presence of variation in the brightness of the stimulus area. Thus an area which is uniformly and homogeneously black, white, or red, etc., has no shading. But if there is considerable variation in the brightness of the area, lighter and darker portions alternating, the shading factor is present.

(C) The chromatic determinant refers to the presence of color (with hue) in the stimulus area.

(Ac) The achromatic determinant refers to the presence of white, gray, or black in the stimulus area.

(Po) The position determinant refers to the utilization of certain topographical positions as factors in the response. For example: “They are kidneys because they are on each side of the picture.” Such responses are rare in normals.

9-13. Importance Scales

F, Sh, C, Ac, Po

x. Determinant impossible (e.g., the chromatic determinant for all achromatic cards).

0. Determinant possibly present, but not used at all.

1. Determinant probably used (as a sort of background factor) but verbalized in neither the original response nor the inquiry.

2. Determinant of secondary importance as contrasted with primary importance.

3. Determinant of considerable importance.

4. Primary determinant.

After gauging the importance of each determinant, we are in a position to know which of the determinants in a given stimulus area were utilized and the degree to which they were utilized. How “accurately” they were utilized is gauged by the quality scales.

The quality of a given determinant refers to its degree of conformity with the corresponding attribute of the reported object as it would be seen in natural life. For example, the quality of the
form involved in a response refers to the degree of conformity that exists between the lines and contours perceived on the card and the lines and contours of the "real" object supposedly represented in the response. Thus, in gauging the quality of the form determinant in the response "bat" on Card I, we attempt to decide how well the contours of the stimulus area actually describe the contours of a bat. Since this is a highly subjective matter, it is well to consider the judgment of a large group of observers as a standard for goodness of quality. If only the subject under examination observes the contours of a bat in a given stimulus area, and competent judges fail to see a bat in the given stimulus area even when it is carefully pointed out to them by the subject, it would be an idiosyncratic response, and the quality is rated 0. The proportion of raters who accept the response as good may be taken as an index of the quality of the response. Since no well-defined norms are available as yet, the judgment of competent raters must suffice for the time being.

14-18. Quality Scales

x. Quality is impossible or not rated. The x rating is used when the importance rating is 0 or 1.

(0'. Enumeration of determinants.)

0. Very poor correspondence.
1. Ill-defined correspondence.
(2'. Conventionalized use of determinants.)
2. Fair correspondence.
3. Good correspondence.
4. Excellent correspondence.

In case the subject merely enumerates colors or indicates light and dark portions or points out lines and contours without further integration or interpretation, the quality of the response is scored 0'. It is alleged that enumeration of colors is symptomatic of certain specific personality disorders, and in order to enable the scorer to separate out the two types of zeros, those for which the quality is generally poor, and those which represent enumeration, the latter are primed.

When a determinant is used in some conventional manner, such as color for separating countries on a map, or lines for indicating directions, or shading for topographical features, the quality is scored as fair, and is primed to distinguish it from other ratings of 2.
Good correspondence in color is attained when the color in the stimulus area corresponds to the natural color (daylight color of an object for an average observer) of the "real" object fairly well, but not perfectly.

**IV. Interpretation**

The determinants which a subject utilizes can be interpreted in various ways to form the final percept. For the purposes of classification we shall distribute interpretation into three classes: Dynamic Static (Dy); Dimensional Structure (St); and Surface Appearance (Su).

(Dy) In the dynamic static class we include the interpretations in which the determinants were interpreted as objects in varying degrees of movement.

(St) In the dimensional structure class we include the interpretations in which the determinants were interpreted as objects possessing structure and dimension.

(Su) In the surface appearance class we include the interpretations in which the determinants were interpreted as objects characterized by a particular surface appearance, such as colored or fleecy, etc.

The degree of importance of each of these three interpretations in a given response is evaluated by the accompanying scale.

**19-21. Importance Scales**

Dy, St, Su

0. Not important in the interpretation.
1. In the background, but explicitly stated neither in the original response nor in the inquiry.
2. Of secondary importance as evidenced by the inquiry or inferred directly from the original response. Perceptibly helped in interpreting the determinant.
3. A most important element in the interpretation as evidenced by mention in the original response.
4. The most important category in the interpretation.

**22-24. Type Scales**

(When importance is 0, type is x.)

**22. Dynamic Static (Dy)**

x. Total absence of movement.
a. Static.
   Standing, sitting, spreading (as wings of animals without any apparent movement).
b. Being moved.
   Being moved by external force, being dropped, revolved, etc.
c. Tension.
   Bending, rising actively outstretched, poised or some other dynamic
   pose, smiling or other expression.

d. Active movement.
   Shaking hands, playing patty cake, talking, gossiping, stirring some-
   thing, dancing, and other dynamic activities.

23. Dimensional Structure (St)

x. Total absence of dimensional structure as a factor in the perception.

a. “Uni-dimensional” objects—where one dimension is most important.
   String, church spire, etc.

b. Two-dimensional object—where two dimensions are most important.
   Shadows, profiles, silhouettes, etc.

c. Three-dimensional scene, giving vista effect.

d. Three-dimensional objects (without vista).

e. Three-dimensional object seen in two dimensions. X-ray, architect’s
   plans, geometrical projections, flat pictures.

24. Surface Appearance (Su)

   There are at least five different types of classification for surface
   appearance. In the first type belong those responses in which the
   determinants are interpreted as giving rise to a surface characterized
   by light-refracting properties ranging from transparency to opaque-
   ness, and including as middle steps, lustrous, luminous, and diffuse
   surfaces. If we could establish a relationship between the trans-
   parent-opaque continuum and some personality trait, it might be
   well to gauge the actual placement of each response on this scale.
   But in the absence of such correlation we shall include this whole
   range of interpretation under the category of “surface light proper-
   ties.”

   The second category consists of texture responses ranging from
   those in which the surface gives the impression of smooth softness
   to hard roughness. The intervening steps include other types of
   texture, such as slimy and scratchy, as well as varying degrees of
   the types mentioned. Here again there is no need to make a fine
   classification in the absence of any basis for claims of correlation
   with personality.

   The third category consists of surfaces characterized by chro-
   matic or achromatic colors. In this category belong those responses
   in which the interpretation is color. The determinant thus inter-
   preted may or may not be color. For example, shading is some-
   times interpreted as color.

   The fourth category consists of responses in which the deter-
minant (usually, but not necessarily, shading) is interpreted as surface shading.

The fifth category is a miscellaneous one, and includes such responses as cold surface, warm surface, etc. Eventually this category may be further classified.

x. When importance is 0.
  a. Surface light properties.
  b. Surface texture properties.
  c. Surface color properties.
  d. Surface shading properties.
  e. Miscellaneous.

If more than one type is found in the same response, the more important category should be noted, and the others discarded.

V. Content

There is a generally accepted belief on the part of Rorschach workers that the content of the response—the type of thing seen—is more subject to the accidental influence of experience than are the determinants and the interpretation which were described previously. Consequently, content as such does not play as important a role as the other supposedly more basic factors in reflecting the personality. The primary categories utilized by the Rorschach workers are human, animal, and inanimate figures, natural phenomena, art, abstract concepts, anatomy, and sex organs. An examination of the use that is made of these categories by Rorschach workers in personality interpretation indicates that they form the basis for judging the presence of the following tendencies in the subject: perception of human (H) or animal (A) figures; perception of whole figures rather than details or parts of figures (W-P); perception of animate as opposed to inanimate figures (An-I); perceptions with sex reference (Sex); and perception of abstract rather than concrete figures (Abs-C).

The tendencies described above are the most common in Rorschach responses, and though there may be other tendencies which some workers look for, only the above-mentioned factors will be dealt with here.

In order to categorize a given response according to its content on the content-importance scale, the scorer must decide the relative importance of the following factors: human, animal, plant, inanimate, and abstract. An importance scale which is the same as that used to gauge the determinants, is applied to each response. In
classifying the responses on the basis of this scale we ask the question: To what category does the response primarily belong? Is it primarily human (H), animal (A), plant (Pl), inanimate (In), or is it primarily abstract (Abs), i.e., in the realm of ideas and impressions? Some responses are simply human figures, and would be rated 4 (the maximum) for factor H (human) and 0 for all the other factors. Similarly, responses may belong to one of the other categories (A, Pl, In, Abs). When a response includes both human and animal elements, the relative importance of the two must be gauged, and similarly for other combinations of content categories.

The importance scale is as follows:

25-29. Importance of Content Scale

H (human); A (animal); Pl (plant); In (inanimate); Abs (abstract)
0. Not important.
1. In the background but not definitely seen either at the time of the original response or during the inquiry. For example, “a man’s coat” would receive a rating of H-1, In-4.
2. Of secondary importance as evidenced by inquiry or inferred directly from the original response.
3. A most important element as evidenced by mention in the original response.
4. The most important content category.

After the importance of each of these factors has been gauged, type scales may be applied to certain of the tendencies indicated. The following tendencies are rated on type scales: Whole-part (W-P); Animate-inanimate (An-I); Abstract-concrete (Abs-C); Sex reference (Sex).

The scales are as follows:

30-33. Content Type Scales

W-P, An-I, Abs-C, Sex

30. Whole-Part Type Scale (W-P)

a. Only part of a figure in a response.
b. Part of a figure mentioned with whole serving as background.
c. Part and whole mentioned with part more clearly defined.
d. Part and whole mentioned but whole most important.
e. Only whole figure mentioned in response.

In deciding whether a given response represents a whole figure, it is important to take into consideration the intent of the subject.
Even when a drawing of the response is available, it is not always easy to decide this question. For example, the bear on Card II is not present in its entirety on the card—but when a subject calls it a bear, he is credited with a whole figure. However, if he specifies the head of a bear, he is credited with giving only part of a figure. Similarly, if he gives "leg of a chair" as a response to the upper part of Card IV, it is scored as part of a figure. In all instances if the subject specifies a part of something, regardless of the drawing, the response is scored as part of a figure.

31. Animate-Inanimate Type Scale (An-I)
   a. Completely or primarily inanimate objects.
   b. Dead, either preserved or deteriorating humans, animals, or plants.
   c. Plants or plant-like objects.
   d. Animals or animal-like objects or beings.
   e. Humans or human-like objects or beings.

32. Abstract-Concrete Type Scale (Abs-C)
   a. Concrete objects, animals, or persons.
   b. Art, photographic representations, or symbols.
   c. Natural phenomena (storms, wind, fire, etc.).
   d. Abstract forces or concepts (struggle, peace, etc.).

33. Sex Reference Type Scale (Sex)
   a. No apparent sex reference.
   b. Doubtful sex reference.
   c. Secondary sex characteristics, or secondary sex activity.
   d. Primary sex organs, or primary sex activity.

34. Definiteness of Content of Response (Def)

Some responses are definite and precise insofar as their contents have definite shapes, contours, texture, or color. Examples of such responses are: Morris chair, Statue of Moses, American Flag. On the other hand, some responses are very indefinite and vague insofar as their content is amorphous without definite texture, color, or shape. Examples of such responses are clouds (indefinite in shape), flowers (with no mention of color or variety), insects, trees, water scenes, landscapes, etc. Between these two extremes we have formulated two intermediate categories yielding a scale which ranges from specific, definite, doubtful to indefinite content.²

²In the orthodox Rorschach scoring the lower categories in the scale are noted as F-, C-, etc. A minus can arise from two sources, lack of correspondence between determinants and interpretation content, or vagueness of content. The above scale has been introduced in order to differentiate between these two sources of minus scores.
0. Vague, hazy, formless, and lacking in specific content, e.g., reflection in water, amoeba, etc.
1. Indefinite, e.g., trees, forest, landscape, etc.
2. Of doubtful definiteness, e.g., generic names, fish-like animals, fowl, etc.
3. Of considerable definiteness, e.g., chair, velvet evening dress, maple leaf, etc.
4. Very definite, e.g., Michelangelo’s Moses, Columbus Circle, etc.

35. Communalilty Scale (Com)

The most popular or common responses are listed by Klopfer (2), and a list of these responses is given in (8). The degree of originality of other responses can be estimated only approximately and will depend upon the extent of the experience of the examiner with Rorschach responses. Until a statistical study of the frequency of the various responses or types of responses is made available, the degree of originality exhibited by each response may be tentatively rated on the following scales.

0. Original.
1. Semi-original (original with popular elements).
2. Semipopular (popular with original elements).
3. Popular.

It should be noted that the distinction usually made between good and poor original responses is not made in the above scale. The scale simply attempts to gauge the degree of communality exhibited by the subject in his Rorschach responses in the same way that the Kent-Rosanoff scale attempts to gauge communality in free association. The quality of the response is gauged elsewhere under quality scales.

36. Degree of Self Reference (SR)

Some responses are purely impersonal with no overt indication of any personal reference. Other responses are clearly self referred, the subject stating that it reminds him of a specific situation or object in his childhood or later life. The degree of self-reference is gauged on the accompanying scale.

0. Impersonal response.
1. Slight degree.
2. Some degree.
3. Considerable degree.
37. Gender of the Observed Figure (Ge)

We have already suggested a scale for determining the importance of sex reference (symbolism or representation) in a given response. The present scale attempts to gauge the gender preference for the animal or human figures reported by the subject. The same figure may be interpreted by one individual as a male and by another as a female.

x. Response bears no reference to gender.
a. Reference to either male or female figure or part of anatomy, without specifying gender.
b. Reference to male—either figure or part of anatomy.
c. Reference to female—either figure or part of anatomy.

VI. Psychological Characteristics of the Response

The categories described above follow the general pattern laid down by Rorschach in classifying responses according to location, determinants, and content. In addition to these categories there are certain other psychological features that characterize the responses of a given subject. Some subjects show considerable perseveration in their responses; others show marked variation in reaction time. There may also be variations in the affect and mood that accompany a given response, in the degree or ornamentation of the response, and in the degree of congruence between the different elements in the response. For each of these factors a tentative scale has been devised for gauging its importance.

38. Reaction Time (RT)

It has been claimed that the variation in reaction time results from attempts at the repression of responses or from confusion resulting when several responses occur simultaneously. Usually only the reaction time of the first response to each card is noted. The time required for producing the successive responses to the same card is rather difficult to gauge because one can never be sure of the beginning and end points for each response. Despite this difficulty it is well to gauge approximately the beginning of each new response and to consider the time intervening between the beginning of one response and its successor as the reaction time for the succeeding response. Since the reaction time scale is rather rough, the errors introduced by this method should not obscure the results unduly. The distribution of reaction times is divided into four quarters at the quartile points, and the rating for responses in each quarter is shown in scale 38.
0. First (longest) quarter in reaction time.
1. Second quarter.
2. Third quarter.
3. Fourth (shortest) quarter.

This scale can be used only for inter-response comparisons, e.g., for comparing reaction times of chromatic and achromatic cards, or similar purposes.

39. Perseveration Tendency (Pe)

Perseveration may make itself manifest either in exact repetition of responses, or in the repetition of a certain type of response such as human anatomy, or in dwelling upon a series of responses, each tending to have the same form. The degree of perseveration present in the responses of a subject may be gauged by the accompanying scale.

0. Response totally independent of previous responses.
1. Some general similarity either in form or content to a previous response.
2. Similarity in form but not in content.
3. Similarity in content but not in form.
4. Exact, or nearly exact, repetition of a previous response.

40. Ornamentation Tendency (Or)

Some subjects tend to adorn or ornament their responses. They either weave a story about what they see or they add little curlicues or other adornments to the blot when they are asked in the inquiry to draw what they see. Thus, one subject added a pair of spectacles to the well-known response of two old ladies on Card VII. This tendency to ornament the response is gauged in the following scale.

0. Precise, unadorned response.
1. Slight tendency toward adorning and confabulating.
2. Some tendency.
3. Marked tendency.
4. Clear-cut example of ornamentation.

41. Mood (Mo)

Very often the subject indicates the mood that accompanied the percept by such terms as gloomy and ominous. The following scale can be used in recording the mood of the response:

a. Very pleasant.
b. Pleasant.
c. Indifferent.
d. Unpleasant.
e. Very unpleasant.
42. Affect of the Subject (Af)

Subjects may show incongruity between the mood of the response and their own affect—for example, laughing while describing two men being burned in oil—Card IV. This response would be rated Mo-e, Af-a.

The affect may be classified on the following scale:

a. Very pleasant.
b. Pleasant.
c. Indifferent.
d. Unpleasant.
e. Very unpleasant.

43. Congruity of Response (Cg)

By congruity one means the degree to which the several elements of a given response hang together. If the response is a simple object or figure consisting of only one primary element such as bear, human figure, and so forth, the response is said to be congruent on an elementary level, for there is no possibility of any conflict between its elements. In more complex responses the factor of congruence or incongruence may be present. For example, “black bear” or “white bear” are congruent responses, but “grass bear” is an incongruent response since no grass bears exist in nature. Rorschach has obtained such a response from a schizophrenic patient who telescoped two percepts into one—green grass and the form of a bear. Rorschach (3) called this a contamination response. Thus the scale of congruity ranges from contamination on the one hand to completely congruent responses on the other.

0. Completely incongruent response. Rorschach’s contamination response, where two percepts from the same stimulus are telescoped in an unusual manner, e.g., grass bear. (Usually found only in schizophrenic patients.)

1. Incongruent response. (Rorschach’s confabulatory response, which, although possible, contains incongruous elements, e.g., Card II—two men shaking hands and a drop of blood falling between them.)

2. Flight of fancy of the uncontrolled type—teetering between whimsicality and lack of congruence.

3. Whimsical but acceptable response. Make believe, e.g., animated cartoons of impossible acts.


44. Succession (Suc)

Succession is usually measured roughly by noting the order of appearance of W, D, and Dr responses on each card. Several cate-
categories of types of succession have been given by Rorschach, Skalweit, Beck, and Klopfer, all of whose systems are qualitative rather than quantitative in nature. Since succession as defined by Rorschach lends itself directly to measurement, a quantitative scale has been devised for measuring it and is explained in greater detail elsewhere (7). Here it is sufficient to indicate that a coefficient $\phi$ can be obtained from the scores which indicates the degree of regularity of the succession.

VII. Rating of Subject's Attitude Toward Response

Certain attitudes are exhibited by the subject toward his own responses which have become important in Rorschach interpretations. These attitudes are evaluated in the following scales.

45. Self-Estimate of the Adequacy of the Response (SEA)

0. Self-condemnatory. Points out major deficiencies or general inadequacy of response, but cannot help giving response (impotent).
1. Insecure. Anxious about response; wants to know whether it is acceptable.
3. Whole-hearted acceptance. Simple acceptance of the response as adequate to the task.

46. Attitude toward Reality of Response (RR)

a. Painfully conscious that the percept is not real.
b. "As if."
c. Interpretation.
d. Percept considered as almost real. Some reference to its interpretation quality.
e. Crude acceptance of percept as real—"recognizing or guessing what it is."

47. Rating the Point of View of the Subject (PVS)

In determining the position from which the object was seen, it is well to ask: "Where are you when you see this? Where do you see it from? Where are you likely to see it?" The answer to this question also often helps in determining the other determinants of the response. Thus, if the subject sees an aeroplane in Card I, and does not see it above him, the chances are that the aeroplane is not in motion, and that it is the picture or diagram of an aeroplane.

(a. Subject being looked at, e.g., eyes peering.)
b. Subject looking straight ahead.
c. Subject looking up.
d. Subject looking down.
e. Subject moving, e.g., in a train.
f. Subject bending over, or in another unusual position.

48. Turning and Handling of Card (TC)
0. Does not take card, leaving it on table.
1. Handles card tentatively on the table, without turning.
2. Takes card and replaces card meticulously, with little or no turning.
3. Asks permission to turn cards.
4. Turns cards at will.

49. Subject's Readiness to Respond (SRR)
0. Complete refusal.
1. Card descriptions (ink-blots, colors, etc.).
2. Vague comments or abstract remarks (gay, nice colors, etc.).
4. Easy or fairly free flow of responses.
5. Extremely free flow of responses.

ANALYSIS OF THE RESULTS

Thus far, the only treatment of the data that has been attempted is the computation of median weights for each of the continuous rating scales and per cent distributions for each of the categories in the noncontinuous scales. Comparisons were then made between the median weights in each of the natural groupings of the scales. For example, in the Locale family of scales (W, D, Dr) the median weight for W, D, and Dr is examined to see in which area the subject shows the strongest tendency and in which he is not very strong. Similarly, the Determinant family of scales (F, Sh, C, Ac, Po) is examined to see the relative strengths of the different factors. The Interpretation family and the Content family can be similarly treated. Thus far, a comparison of the results of our scaling methods and those of the orthodox Rorschach method reveals some similarities but also considerable differences. This is not unexpected, since our scales do not always correspond exactly to the orthodox Rorschach factors that they attempt to approximate. Before conclusive results on these similarities and differences can be reported, many more records will have to be examined. It is suggestive, however, that the following points of comparison have already emerged. The W:M ratio seems to be approximated by our Mz:Dy ratio and the M:C by our Dy:Su ratio. The degree of similarity remains to be investigated.

* Scales 48 and 49 may be used as overall measures.
One way of avoiding the treatment of isolated variables is to treat the above family groupings of scales in patterns. Thus, the Locale family can give rise to about 120 different patterns of the type \( W_4 \, D_3 \, D_r \), etc. After the median weights are computed, the nearest whole number can be substituted for the actual median \( (W_{4.7} \, D_{3.3} \, D_{r.1}) \) becoming \( (W_4 \, D_3 \, D_{r.1}) \). Not all of the 120 different patterns arise in actual practice. The patterns that are diagnostic for the identification of one or two contrasted groups can then be easily discovered by making pattern analyses of the frequencies of each pattern in the two groups by a method described elsewhere (5).

**Summary**

A series of psychometric scales have been provided for scoring Rorschach responses. The purpose of these scales is to provide an objective framework for the rating of Rorschach factors over all the responses. The objectivity (scoring reliability) of these scales has been determined and found sufficiently high to permit their use. While this scoring system is in no way offered as a substitute for the orthodox system in clinical use, it does provide an objective framework on the basis of which the reliability and validity of the results can be determined statistically. When sufficient data are collected on contrasted groups of normals and various categories of abnormals, the claims of Rorschach workers may be subjected to scientific testing and the virtues as well as the limitations of the test objectively determined.

**Bibliography**

SCORING RORSCHACH TEST RESPONSES

APPENDIX

A Sample of the Scoring of a Rorschach Test Response on the Psychometric Scales

Card 1.

Response: 1. 7” Resembles a butterfly.
Inquiry: 1. (Where?) Here’s the body and the wings. There are various types of them. It’s the general layout of the thing. Butterflies are colored. (Is this colored?) No. (What’s this?) Head. (What’s this?) Colors, I suppose. (Why is this a butterfly?) General layout. Color effect on wings. (What color?) Gray, light, and dark in spots.

Scoring:

I. Locale
   (1) W 4 (2) D 2 (3) Dr 1 (4) S 2 (5) Ts b (6) Incl 4

II. Organization
   (7) Ms 1 (8) Ts 4

III. Determinants
   A. Importance
      (9) F 4 (10) Sh 2 (11) C x (12) Ac 1 (13) Po 0
   B. Quality
      (14) F 2 (15) Sh 2 (16) C x (17) Ac x (18) Po x

IV. Interpretation
   A. Importance
      (19) Dy 0 (20) St 4 (21) Su 2
   B. Type
      (22) Dy x (23) St d (24) Su c

V. Content
   (25) H 0 (26) A 4 (27) Pl 0 (28) In 0 (29) Abs 0
   (30) W-P d (31) An-I d (32) Abs-C a (33) Set a (34) Def 2
   (35) Com 4 (36) SR 0 (37) Go 0

VI. Psychological characteristics of the response
   (38) Rt - (39) Pr 0 (40) Or 2 (41) Mo c (42) Af c
   (43) Cg 4 (44) Sw -

VII. Rating of the subject’s attitude toward response
   (45) SEA 2 (46) RR c (47) PVS b (48) TC 2 (49) SRR 3

* The numbers in parentheses refer to the number of the scale.