

CHAPTER II

MEASUREMENT OF PERSONALITY

Techniques of Personality Assessment

A large variety of techniques and approaches have been applied to the measurement of personality. The basis for the particular content of the test has ranged from unsubstantiated hunches about personality to predictions from reasonably well-formulated personality theories. The techniques range from highly structural multiple-choice forms to free interview situations. The different techniques were developed in different situations and today there exist a wide variety of them. The chief ones are classified in the following categories:

- | | |
|----------------------------|--------------------------|
| 1. Interview | 5. Situational Tests |
| 2. Case study Method | 6. Projective Techniques |
| 3. Rating Scales | 7. Objective Tests |
| 4. Personality Inventories | |

These are neither the only categories nor are they exclusive ones. Some overlapping is present from one technique to another. Yet there are certain distinctive attributes which are characteristic of each.

Interview:

The personal interview is one of the oldest devices for measuring personality. It is obviously subjective. Yet, interviewers sometimes make reasonably good measuring instruments. Psychologists and psychiatrists, although, they secure all the information they can from objective tests, usually rely on the interview to round out their picture of a personality. Interviews are conducted for selection in educational and vocational fields and for guidance and counselling purposes, where it is used to assess the personality characteristics of an individual through the observation of his manner and speed, his expressions and ideas.

The interview may include the results of some projective material, such as TAT, free association test or sentence completion. It is this very multifariousness of the techniques used, which makes its results so uncertain, there is a large scope for the interviewer to jump to wrong conclusions. It depends almost on the skill

of the interviewer - and skill in interviewing is hard to teach. More serious than that is the difficulty of expressing the results of an interview in quantitative terms. In general the results can be communicated only in word descriptions, not in objective scores. This makes it difficult to compare people by means of the interview or even to tell whether the interviewer has, in fact, made any valid measurement.

In spite of its obvious drawbacks mentioned above, interview still remains to be the most frequently used technique because of its ease and acceptability.

Reliability and Validity of the interview Method:

On the whole, this technique is said to have poor reliability. Various reasons are responsible for this. The prejudices of the interviewers constitute one such reason. A notorious instance of this is quoted by Rice.¹ Two interviewers were investigating into the causes of destitution among people who had applied for relief. One of the two interviewers was a socialist and the other was a prohibitionist. The socialist attributed 39 per cent of his cases to the industrial conditions,

¹ Rice, S.A., 'Contagious Bias in the Interview,' Amer. J. Sociol., XXXV, 420-423, 1929.

and 22 per cent to alchoholism. The prohibitionist on the other hand attributed 62 per cent to alchoholism and only 7 per cent to the industrial conditions. Such subjectivity accounts for the unreliability of the results.

The evidence for the validity of the interview procedures is more scarce than that for its reliability. Vernon and Parry¹ have reviewed studies about the validity of the various selection procedures in the British Armed Forces, and have arrived at the conclusion that prediction based on the combination of tests and interviews were even worst than those based on the best tests alone. McClelland² arrived at the conclusion that the primary teachers' judgements about the personality qualities of their students had no predictive validity.

Kelley and Fiske³ investigated most thoroughly the predictive validity of the interviews, ability tests and personality tests in case of the selection of candidates for the clinical psychology course. Here again the evidence was unfavourable for the interview

¹ Vernon and Parry, as cited in Vernon, Op. Cit., p. 25.

² McClelland, W., Selection for Secondary Education, (London: University of London Press, 1942), pp. 152-153.

³ Kelley and Fiske, D.W., The Prediction of Performance in Clinical Psychology, (An Arbor University of Michigan Press, 1951), As Cited in P.E. Vernon, 1953, Op.Cit., p. 26.

method. However, Vernon's¹ findings in connection with the validity of the selection procedures of the Civil Service Selection Board present more favourable evidence of the validity of interviews.

Case Study Method:

This method is used generally by the clinical psychologists and the psychoanalysts while dealing with the cases of abnormal behaviour. When a child or a person encounters some problem which can be solved only by a trained psychologist, the latter is approached for help. Three kinds of data pertaining to past, present and future events are included in a case study:

- (1) information about the developmental history,
- (2) information about the present status of the individual and (3) his ideas or plans about the future or his orientation to it. All this information is gathered from a variety of sources. The information collected is then put together to get a composite view of the subject's personality. Thus, case study technique is not a unitary method of personality assessment. Rather it employs all other methods whenever suitable, to get a complete picture of the individual.

¹ Vernon, P.E., 'The Validation of Civil Service Selection Board Procedure', Occu.Psychol.XXXV, 75-95, 1950.

According to Allport, 'unskillfully used, it becomes a meaningless chronology, or a confusion of fact and fiction, of guesswork and misinterpretation. Properly used it is the most revealing method of all.'¹

Though this method has its limitations especially in the hands of the untrained workers, it has the potentiality to provide the complete and best possible picture of the individual's characteristics.

Rating Method:

Rating is a formal and articulate estimate of the strength of one or more qualities in a personality, arrived at on the basis of direct acquaintance with that personality. Teachers, army officers, foremen, personnel directors and social workers are among those who use ratings extensively for practical purposes. But it is also a method indispensable to psychologists who are forced in spite of their distrust of subjective judgements to rely time and again upon ratings as basic criteria in their investigations.

Rating yield quantitative comparisons of different people in respect to one variable at a time.

¹ Allport, G.W., Personality - A Psychological Interpretation, (New York: Henry Holt & Co., 1949), p. 390.

Of itself the rating method gives no consideration to the setting of the rated variable in the personal life. It is a tool for analysis and comparison, nothing more. But inspite of its limitations the method of rating has its uses, and many years of experience have shown how to make most of it. Some of the principles of effective and useful rating are as follows:

(1) Variables must be clearly defined: what variables the investigator will use, of course, depend upon his specific purpose. For purely practical purposes a schedule of 'characterial' variable may be chosen, e.g. biological dimensions as 'personal charm,' 'leadership,' 'citizenship,' 'value to business', and the like. But for the purposes of a strictly psychological analysis, a more orderly theory of the structure of personality is necessary, together with a guiding logic or common traits. In defining variables for the raters, it is possible (a) to give synonymous terms, (b) to describe in a general way the kind of behaviour to which the variable pertains, and (c) to give specific instances of its operation in overt conduct.

The rating scales are of different types. The important types according to Guilford¹ are numerical scales, graphic scales, standard scales, cumulated points scale, and forced-choice scales.

The chief advantages of these techniques are that they are less time-consuming, simple in use, and applicable in a wide variety of situations.

(2) However, there are two basic types of rating scales viz. numerical or scoring type and the graphical type. Of course, the scoring scale is more common. Here the subject is rated without direct comparison with other people in his immediate circle. It requires some frame of reference in the rater's mind, but generally a somewhat vague reference to the 'general population' suffices.

In the graphic rating scale, the judge makes as fine a discriminations as he chooses, placing a mark upon a straight line anywhere he wishes between the extreme low and extreme high. By this method the number of distinguishable intervals is theoretically infinite. In practice, since the graphic scale is more finely graduated than is warranted by its accuracy, it is always

¹ Guilford, J.P., Psychometric Methods., (New York: McGraw Hill Book Co., 1954), p. 263.

reconverted by the investigator into arithmetical or statistical units of some kind.

The rating scale is used only when a whole group of associated individuals are to be rated in relation to one another. The judge places their names in serial order in respect to their status for a given variable. In such a scale there are no true arithmetical units, since it is impossible to demonstrate that the intervals between successive individuals in the rank-order are equi-distant. If the distances were equal, there would then be a rectangular rather than 'normal' distribution of the variable, a most unusual situation. It is for these reasons that all ordinary statistical computations based upon rank-orders are precarious, though they are sometimes mistakenly attempted.

The usefulness of the rating or ranking method is limited to such investigations which from start to finish employ one single group of subjects, and require no units other than the serial positions of these subjects. Each subject can be compared with every other subject before his final position is determined. Such tangible comparisons are sometimes thought to be more accurate than the 'solitary' method of scoring. Extreme cases are

likely to be reliably ranked, but the more moderate individuals receive their ranks almost by guesswork.

(3) Judges require training: judges differ greatly in ability, but whatever their inherent skill, to do his best each judge requires: (a) instructions concerning the nature of the variables, (b) instructions concerning the intervals used together with a warning not to place his ratings within too narrow a range, but to make free use of the scale, (c) instruction to make each rating a fresh and independent judgement unprejudiced by preceding judgements. The rater also requires, (d) accurate acquaintance with the subjects, (e) broad enough experience with people at large to provide a suitable range of reference for his judgements, and (f) sufficient time, patience, and incentive to work carefully and honestly.

(4) Variables that are overt in expression are more reliably rated than variables that are covert. Agreement is especially good when a variable can be judged on the basis of past and present accomplishment known to all the judges. A safe generalization seems to be that variables pertaining to self-expression and

to social behaviour are among the most reliably rated, whereas variables pertaining to inner life and attitudes toward self are much less dependable. In addition to the overtness with which a trait is expressed, spielman and Burt found that actively emotional and socially acceptable qualities can be reliably rated.

(5) The subjective certainty of the judges is an indication of the reliability of their ratings. If judges are asked to indicate in conjunction with each rating the degree of certainty with which they have made the rating, it is found that judgements recorded with confidence are the most reliable.

(6) Some persons are more reliably rated than others. There are 'open' personalities about whom all judges agree exceptionally well, there are 'enigmatic' personalities about whom they agree hardly at all.

(7) Judges rate best those who are most like themselves. Commenting on their own finding that judges rate best those whom they resemble most, and rate poorly those whom they least resemble.

(8) In self-rating there is a tendency to overestimate those qualities considered desirable and to underestimate those considered undesirable.

(9) In general, ratings are complementary.

Unless prevented from so doing, a judge is likely to give over-generous ratings. He assign superior scores to too large a proportion of the subjects. The fallacy of generosity is found especially in the rating of friends. It extends in marked degree also to rating of members of one's own professional group, one's fraternal associates, and one's own sex.

(10) The 'halo effect' prevents the giving of strictly independent judgements on the separate variables.

According to Rugg¹ 'we rate our fellows in terms of a general mental attitude toward them' and, 'there is dominating this mental attitude toward the personality as a whole, a like mental attitude toward particular qualities.' The general impression thus influences any one specific judgement, and in turn some impressive specific observation may be the source of the general impression.

Though impossible to eradicate completely, there are many ways to diminish the halo effect, e.g. (a) by specific warning against it, (b) by employing distinctive

¹ Rugg, H.O., Jr. of Edu. Psychol., 1921, 12, 425-438
"Is the rating of Human Character Practicable?"

and well-defined variables, (c) by using alert and trained judges, (d) by avoiding characterial and censorial variables, (e) by so varying the presentation of the qualities to be rated that a fresh and independent consideration of each is demanded, (f) by avoiding haste and perfunctoriness in making the ratings and (g) by averaging together the ratings of several judges so that to some degree the prejudices of several judges will cancel one another.

(11) The reliability of ratings depends upon the extent of agreement among the judges. Since it is seldom possible to decide which judge is 'right' when judges disagree as to the score or rank of a given subject, the mean of all the judgements is ordinarily taken to represent the nearest approximation to the subject's true position. Yet if the judges have disagreed seriously with one another, this mean position may be entirely worthless. Therefore, before using the mean rating of a subject as his true rating, it is necessary to determine the extent of the judges' agreement, and as Guilford has shown, it is also beneficial to estimate even statistically the extent of likely errors make adjustment in ratings.

Personality Inventories

Typically personality inventories and questionnaires present the subject with a series of statements or questions concerning a variety of personal reaction, attitudes and emotions which may be experienced in life situation. The subject is asked to respond to these items so as to tell something about himself, sometimes by checking 'Yes' or 'No' or 'Uncertain', sometimes by ranking items according to what he feels to be their order with respect to feeling or value, or again by checking statements 'agree', 'disagree', or 'uncertain',

Review of Well Known Inventories

Some of the major inventories commonly used are reviewed below:

Woodworth Personal Data Sheet:

This inventory consists of 116 questions, to be answered 'Yes', or 'No'. The questions were framed on the basis of common psychoneurotic or preneurotic symptoms about such behaviour as is found in cases of abnormal fears, obsession and compulsions, nightmares and other sleep disturbances, excessive fatigue and other psychosomatic symptoms, feeling of unreality, motor disturbances and the like. The total number of neurotic answers was the score of an individual which was compared with the average scores

of the normals and the neurotics in the standardization sample.

After the end of the World War I, Woodworth Personal Data Sheet was published for use with civilian population. It was also revised and adapted a number of times.

Cornell Index:

This inventory was developed during the second World War. The items were constructed on the same basis i.e. common psychoneurotic symptoms. The following areas of disturbances have been covered in it:

- Defects in adjustment expressed as feelings of fear and inadequacy.
- Pathological mood reactions, especially depression.
- Nervousness and anxiety.
- Neurocirculatory psychosomatic symptoms.
- Pathological startle reactions.
- Other psychosomatic symptoms.
- Hypochondriasis and aethenia.
- Gastrointestinal psychosomatic symptoms.
- Excessive sensitivity and suspiciousness.
- Troublesome psychopathy.

The score is compared with the performance of normal and psychiatric subjects. An alternative scoring procedure

has also been devised in which the total score is interpreted and answers to particular individual items called 'Stop items' are given particular attention and importance in clinical diagnosis.

According to Anastasi¹, even the use of stop questions did not improve the validity of this instrument. It was of course designed to be only a rough screening device for personal and psychosomatic disturbances in the military selection. It is also used in civilian practice and norms for male adults are available.

The Bell Adjustment Inventory:

It consists of questions intended to evaluate the subject's status in respect to home (satisfaction or dissatisfaction with home life), health (extent of illness), social adjustment (extent of shyness, submissiveness, introversion), emotional adjustment (extent of depression, nervousness, ease of disturbance), and occupational adjustment (satisfaction with work, associates and conditions.)².

¹ Anastasi, Anne, Psychological Testing, 2nd ed. 1963, (New York: Macmillan Company). p. 531.

² Freeman, F. S., Theory and Practice of Psychological Testing, (New York: Holt, Rinehart & Winston), May 1963. (Third Edition).

There are two forms, one for students and one for adults. The items are of the usual kind, to be answered as Yes, No, or?

The inventory based on content validity raises a problem common to all devices of this kind. Do the questions and the scores for each category actually represent separate and distinct aspect of behaviour and adjustment? Are these aspects mutually exclusive? Some critics maintain they are not. They hold that the same personality variables influence adjustment in all situations, and therefore, that the more useful and significant inventories are those that probe the various psychological mechanisms such as hysteria, defense and escape techniques and psychosomatic manifestations. Other psychologists, while recognizing the instrument's inability to reveal the dynamics of behaviour, nevertheless believe that it is useful in placing the individual relative to a group in respect to the specified areas of behaviour, and as a basis for further psychological interviewing. While the first criticism is warranted, the Bell inventory has found wide and justified use for the latter purpose.

Bernreuter Personality Inventory:

This test consists of 125 items, describing both adjustment and interests. Each item is to be answered with 'Yes', 'No', or '?'. Four keys were prepared on the basis of results from four previous tests: Thurstone's Personality Schedule of Neurotic Tendencies (1930), Laird's Inventory of Extraversion-Introversion (1925), Allport's Ascendence-Submission Reaction Study (1928), and Bernreuter's Test of Self-Sufficiency. These four tests and the personality inventory were administered to adults selected to represent extreme group. Each item in the inventory was correlated with total scores on each of the four tests. The answer to each item were assigned points on the basis of these correlations; the higher the correlation, the greater the number of points allotted. The total score was the addition of such points on all the items of the scale. These scores correlated highly with the original tests. Bernreuter's score for neurotic tendencies correlated .94 with Thurstone's schedule. Laird's and Bernreuter's introversion scores correlated .79. Allport's measures of ascendancy and Bernreuter's dominance correlated .81, and the two measures of self-sufficiency .89.

The split-half reliability of the scores on the Bernreuter personality inventory was high, with median coefficient of .90. An interesting observation about the scores was that there were very high intercorrelation between the different scales. This led Flanagan to make factorial analysis of Bernreuter's scores of 305 eleventh-grade boys. Two factors were isolated which accounted for the inter-correlations of the four scores on the inventory. The first one was a large factor with high positive loading on neurotic tendencies, introversion and submission and high negative loading on the self-sufficiency items. This factor was named as lack of self-confidence. The other factor, a much smaller one, was called sociability. Two new scoring keys were prepared by Flanagan to measure these two factors in addition to the four old ones. As a matter of fact these two factors are not additional scales, because they have been derived from the old ones only. The intercorrelations between the old scales denoted that there was large overlapping in the old scales. By these two new scales, the intercorrelations are explained and overlapping is reduced, and therefore,

they could be taken as substitutes for the old scales.

This inventory became very popular in use because it measured the four traits through a single administration and in less time than the original four tests. The correlations between the original tests and the corresponding scales of this inventory were also very high to justify its use. And the scores were sufficiently stable, that is, the test-retest reliability was also high.¹

The investigator has used this inventory in Gujarati version to study the factors affecting personality measurement. Validity and reliability of this inventory in Gujarati version has been discussed in Chapter V.

The Minnesota Personality Scale:

This has separate norms for men and women and is intended to rate the following aspect of personality: morale, social adjustment, family relations, emotionality and economic conservatism. The inventory is devised for use in the last few years of high schools, with college students, and in some 'adult cases'. An aspect of this instrument infrequently found is the gradation of answers

¹ Farnsworth, P.R., Genetic Study of the Bernreuter Personality Inventory, J. Genetic Psychol. 11 : 3-13, 1938.

whereby the subject indicates the strength of his responses. Instead of the commonly used responses such as 'Yes', 'No' or '?', the subject in this instance has five choices, such as 'strongly agree, agree, undecided, disagree, strongly disagree,' or 'almost always, frequently, occasionally, rarely, almost never.' The score of each item is weighted from one to five corresponding to the degree of intensity represented by the choice of answer.

The particular selection of five personality aspects tested by the Minnesota inventory may appear to be a rather strange one. The author explains the selection as being '.... the result of work on problems of personality measurement in a clinical personnel program in the University of Minnesota. The personality aspects sampled with this instrument have been found valuable in identifying'... a substantial proportion of adjustment problems in a large scale student personnel programme' after a number of trends and attitudes had been experimentally investigated.

The Minnesota Multiphasic Personality Inventory (MMPI):

This is by far the most well-known of the personality inventories used in the clinical practice. It has aroused

such an amount of interest among its users and research workers that a tremendously large volume of work had gone into experimentation with it. The bibliographies listed in the third, the fourth and the fifth mental measurement year books only can give an idea about it. Moreover, it has been applied to measure more and more personality characteristics by independent workers.

MMPI has the reputation of being a test constructed along a very systematic procedure. The authors of the inventory, Hathway and Mckinley, collected about one thousand items on the basis of their own clinical experience, the case study records of psychiatric cases, literature on psychiatry and personality and adjustment inventories.

The inventory consists of 550 statements each of which is printed on a separate card. The cards are sorted by the subject into three groups - 'true', 'false', or 'cannot say', depending upon whether he regards the statement true of himself or not. The items have been classified under twenty six headings - for example, general health, gastrointestinal system, family and marital, religious attitudes, affect, delusions, phobias, masculinity - feminioity, interests. Originally

various items were selected and grouped to form separate scales for scoring in the nine categories. These nine categories are: hypochondriasis, depression, hysteria, psychopathic deviate, masculinity interests, paranoia, psychasthenia, schizophrenia and hypomania.

In addition to the above mentioned nine diagnostic scales, there are four special validating scales. These are the question score, the lie score, the validity score and correction score.

MMPI is a good instrument for general screening purposes. It also differentiates psychotics, neurotics and normals from one another. But the diagnostic validity of the separate clinical scales is questionable, especially, in the light of recent studies which employed factor-analytical procedures.¹

Like these, there are ~~so~~ many other inventories for measurement of personality: Allport Ascendence - Submission Reaction Study, Guilford Factorial Inventory, Gorden Personal Profile, the California Test of Personality etc.

General Evaluation of Personality Inventories:

There is tremendous amount of growth and use of the personality inventories. This is, inspite of the fact that

¹ Wheeler, W.M., et al. 'The Internal Structure of the MMPI.' Jn. of Consult. Psychol., XV, 134-141, 1951.

they are so vehemently criticised by extremists who would like to ban this instrument from the field of psychological testing. Most of the psychologists, however, occupy the intermediate position where they make use of the inventories with the full awareness of their limitations. It is, therefore, necessary to see what are the major points of criticism against the use of personality inventories in general.

(1) The behaviour of an individual is more changeable in the areas covered by the personality tests than those covered by ability and aptitude tests. In other words, the personality variables are not as stable as the ability variables. But this fact does not disqualify personality inventories in particular. Rather it is one of the unavoidable obstacles in the field of personality measurement, whatever be the method used. It particularly poses a problem for determining the reliability of the personality tests as such. When the behaviour itself is subject to change, the inconsistency of responses cannot be solely attributed either to the method of assessment or the behaviour itself. But the need for assessment is so great that one has to tolerate this fact on the ground that the deeper and more subtle

patterns of behaviour are more enduring and once formed, do not change so easily. This makes it possible to measure and predict behaviour with sufficient accuracy.

(2) Sometimes one does not know whether the behaviour itself has undergone change, but the responses are changed. Guilford¹ has reported studies in this connection. Eventhough there are some changes, they often act in the opposite directions to neutralise each other and the total effect due to such changes is negligible.

(3) An individual does not behave consistently in all the situations. For example, one who is extraverted and sociable in a classroom may not behave in the same manner at home and among relatives. If the items in the inventory cover some narrow field of behaviour, then of course this point of criticism stands. But as a general rule, in the item construction, the area of behaviour should be covered as widely and thoroughly as possible. If this is not done, it is a draw-back of that particular test and not of the personality inventories in general.

(4) Some critics say that the examinee does not know himself well enough to make a dependable self-report. Even

¹ Guilford, J.P., Personality, (New York: McGraw Hill Book Co., 1959), p. 193.

though it sounds absurd to some, there is still some element of truth in it. So far as the questions asked are simple such as, 'Do you get nightmares' or 'Do you keep a diary', it is very easy to report correctly. Majority of the items are of this kind, or rather they should be of such kind. But sometimes the subjects come across an item which does not merely ask to report a fact. It requires him to give his judgement or interpretation of a situation, which might be well beyond his ability. Firstly, such cases are rare. It is never so difficult. Any one can reasonably be expected to think for himself, make judgements and interpret facts. Secondly, during item analysis, the items which cannot be understood properly which are ambiguous, and which are beyond the grasp of the group on which the test is being standardized are most likely to be eliminated. Even though some such items remain, Guilford¹ says that whatever the subject reports is significant for him. Only thing for consideration is that the response should be properly keyed. Even if the subject misinterprets the facts, it can be taken as a significant indicator of his behaviour mechanisms.

¹ Guilford, J.P., 1959 Ibid., pp. 191-192.

(5) Different examinees interpret an item in different ways even though the examinees are generally supposed to have a common cultural and educational background with those included in the standardization sample, and even though the bad items are dropped through item analysis, there is bound to be some possibility that a few items can be misinterpreted or rather interpreted differently by different persons. Firstly one or two of such items would not matter if the rest of the items are carefully constructed and edited and have gone through rigorous item analysis procedures. There is no need to despise the value of the entire inventory if one or two bad items can be detected on this ground. Secondly Guilford¹ puts forth a different point of view. According to him if an item has gone through an item analysis process, it has some validity. It might be even due to the fact that item is ambiguous and is differently interpreted. It becomes a kind of projective test. 'If an item predicts or indicates trait positions of individuals, it does so inspite of, and perhaps in some cases because of misinterpretations.'

¹ Guilford, J.P., 1959 Ibid., pp. 193-194.

(6) Examinees are not always honest in answering the personality tests. The question arises: 'Do all examinees invariably falsify their answers?' The fact that there is always a scope for malingering in the self-report inventories, does not imply this. It can be done but it is not done invariably. For example, an applicant for a job would try to appear good by answering in the more desirable directions, but one who comes for solution of his difficulties to a counselling centre has more reasons to be honest in his responses. During the second World War, the prospective recruits tried to fake bad responses because they knew that emotionally unstable individuals were not sent for dangerous and taxing situations and under conditions of stress. So the malingering depends upon the purpose of taking the test. Edwards¹ conducted an investigation to determine the amount of relationship between the social desirability of a trait and the possibility of a trait being endorsed. The relationship found was very high ($r = .87$). Hanley² later on in an independent inquiry confirmed this

¹ Edwards, A.L., 'The Relationship Between the Judged Desirability of a Trait and the Probability that the Trait will be Endorsed.' J. Appl. Psychol. XXXVII, 90-93, 1953.

² Hanley, C., 'Social Desirability and Responses to Items from Three MMPI Scales.' D.Sc., and K.J. Appl. Psychol., XI, 324-328.

relationship. But there are a number of ways in which it can be interpreted. Firstly the examinees bias their answers in the direction of social desirability. Second, the socially desirable qualities are more common among people. Thirdly, what qualities individuals have, they consider those to be desirable. In the experiments¹ the possibility of influencing the test scores by giving instructions to do so was studied and it was found that it was quite possible. But this does not necessitate a total ban on the use of inventories. It is necessary to be more cautious. In fact, in the case of MMPI, there are some devices such as validation scores which act as checks against such practices. More recently, the forced choice technique has come to be explored as a very effective check against this.

(7) Another point of criticism is the response set or response bias in answering. A subject might be more prone to answer 'Yes' rather than 'No' or vice versa. This tendency certainly vitiates the actual score in either the positive or negative direction depending upon the nature of the items. This is not a criticism of

¹ Bernreuter, R.G., 'The Theory and Construction of Personality Inventory,' J. of Soc. Psychol., IV: 387-405, 1933.

personality inventories in particular, because this applies to tests of abilities, aptitudes and achievement as well. Generally the correct or keyed responses should be distributed evenly among the different alternative positions of the answers. The validation scores detect carelessness in responding or such other factors and act as check against these.

Though these are the various criticism against the personality inventories, they do not warrant ban on their use. There are explanations and ways out. There are no other tools that can be so easily handled by moderately trained workers in the field of psychological testing. The need for the assessment of personality qualities is so great that the few experts who can profitably use the projective or other methods, cannot be solely depended upon. Moreover, the value of other methods is also greatly debated and questioned. In the present circumstances, therefore, any attempt to improve upon technique which can be most widely used is welcome and needs due encouragement.

Summary

Personality inventories originated in the first world war as a quick screening device. Since then there has been a

tremendous growth in their number and the variety. Different inventories serve different purposes. Some are suitable in clinical use, some in counselling and some in vocational guidance or selection. The content of each inventory depends upon its purpose. Some inventories measure single factor, while other measures more than one. All of them are based on the principle of self rating.

Even though they are most widely used, there is criticism against them. If their role is understood properly as a technique which is easy, quick, reliable and economical, they serve the purpose very well. The point of criticism often leads to the development and improvement of the technique. In this case also, the criticism about faking behaviour of the subjects led to the development of the lie scales and the forced choice technique. There is yet scope for improvement.

Situational Tests

Among more recent developments in psychological testing are situational tests that either test the individual in action or confront him with situations related to his own life, in response to which he gives expression to his feelings for other persons. The

individual's behaviour is related or evaluated by his peers or by judges. Some of such methods to assess the individual in a specific situation are summarized below.

Sociometric Method:

This method credited to J.L. Moreno, may be defined as a technique for revealing and evaluating the social structure of a group through the measurement of the frequency of acceptance among individuals who constitute the group. It is an approach to the problem of studying interpersonal relationships. This technique permits the analysis of each person's position and status within the group with respect to a particular criterion. The method also reveals the organization of the group as well as identifying dominant individuals, cliques, cleavages, and patterns of social attraction and rejection. The reasons for the existing patterns of attraction and avoidance can then be determined if the personality traits of each individual are known and the values of the group as a whole established.

The sociometric test requires that each individual in a group would choose one or more other persons in that group for a specified purpose. In schoolroom, the pupils

may be asked to name their first and second preferences next to whom they wish to sit, or with whom they wish to attend the movies etc. The method was adopted for use in the armed forces in an effort to identify individuals for specific assignments requiring, for example, leadership and dependability. Thus each individual is viewed in his social relationship in the whole group.

A recent sociometric instrument is the Syracuse Scale of Social Relations, separate forms of which are available for elementary, junior high school and senior high school pupils. The results obtained with this scale are expected to indicate the extent to which each individual feels favourable to his classmates. The scores are intended to provide answers - at least tentative ones - to such questions as these: Does a particular pupil feel comfortable with his classmate? etc. Answers to such question should provide a basis upon which to proceed with any remedial measures that might be necessary, to be taken by the parent, teacher, guidance counsellor or psychologist.

Psychodrama:

This technique requires an individual to play spontaneously on assigned role in a specified situation. The central principle of the psychodrama is spontaneity which has been defined by Moreno as the ability of the subject to meet each new situation with adequacy as 'the most important vitalizer of living structure.' The aim of the psychodrama is to develop in the subject the capacity to play his life roles in a spontaneous and always creative manner that will enable him to meet adequately the demands of new and evolving situations, rather than by employing stereotyped patterns of response.

The psychological rationale of the psychodrama can be stated in the following lines: 'In therapy, the subject by acting, by participating in the reproduction of a life situation significant to him, experiences an emotional catharsis. In the process, while he gains insight into his own behaviour, he should learn how to meet a situation adequately through observations of himself and through interpretations and evaluations given by the therapist and members of the audience.'

Although Moreno's book on psychodrama was published in 1946, there is yet little sound empirical evidence to establish its value as a method of diagnosis and treatment suitable for wide use. The method has yet to achieve a satisfactory level of objectivity and systematic organization in respect to techniques of observation, rating of behaviour and interpretation of responses. Furthermore, the validity of the hypotheses regarding the value of the psychodrama as a technique to develop spontaneous, adequate and adjusted personalities has not been demonstrated. A serious obstacle to the experimental development and the use of the psychodramatic group technique is its heavy requirement of time, personnel, and equipment. Perhaps this accounts for the paucity of research data.

Office Strategic Services (OSS):

During the World War II a group of psychologists and psychiatrists were given the assignment of assessing the traits of men and women recruited for the OSS as it has come to be known.¹ The task was to devise test

¹ OSS Staff, Assessment of Men, (New York: Holt Rinehart and Winston, Inc., 1948).

procedures that would reveal the recruits' personalities and give reliable predictions of their future usefulness in this branch of military service. The assessment staff decided to use the 'wholistic' approach that is, to evaluate each personality as a whole: This meant that some members of the staff would provide an overall evaluation and description of each individual, based upon interview, each candidate would be tested, observed, and evaluated in respect to specific traits of personality, intellect and physique. Finally all information of each individual would be assembled, organized and interrelated to provide a complete description of each candidate, on the basis of their unified conception of each individual's personality traits, the staff estimated the probable level of future performance. For each recruit, an assignment was determined upon, using as criteria the statements of the qualifications required for each job as formulated by each branch of the OSS.

Since the task of the staff was to devise tests that would reveal personality traits for the purpose of predicting success in future assignments, it was necessary to appraise the forecasting value of the procedures being used. Even in ordinary civilian situations, where subjects

are under frequent or constant observation and where their effectiveness in performance can be judged in terms of relatively concrete outcomes, assigning rating presents serious difficulties. It was to be expected, therefore, that evaluations of the performance of persons accepted after OSS assessment would be even more difficult and less reliable for these men and women were not always under close observation in the field, it was always not possible to rate their work, because often the results were intangible and deferred, and for the most part, the primary judges on the job were inexperienced in making psychological evaluations.

Since the termination of war, little has been done along the lines of OSS situational tests, probably because of the difficulties and complexities inherent in the method.

Level of Aspiration Test:

This test is one of the methods used in situational test. The term 'aspiration' in psychology means expectation or judgement of future performance from previous knowledge of results on a particular task. It is likely that setting of level of aspiration differs from



individual to individual and depends upon several factors.¹

It determines the success or failure of persons according to their achievement. But Hoppe, Frank and others² showed that the level of performance in a task is not judged as 'success' or 'failure' in terms of the absolute degree of accomplishment, but in terms of the level of aspiration or goal one sets in that particular line of achievement.

Munn³ has defined the term 'aspiration' as striving to reach a certain level of performance.

Murphy and Newcomb⁴ define 'aspiration level' as not the highest level imaginable, but the level which the organism sets as its goal toward which it hopes it may possibly rise.

According to Hoppe⁵, an early German worker in this field 'level of aspiration' referred to the individual's expectations or goals in regard to his own future achievement in a given task.

-
- ¹ Murphy, G., and Newcomb, T.M., Experimental Social Psychology, New York and London: Harper and Brothers, p. 42
 - ² Sheriff, M., and Sheriff, C., An outline of Social Psychology, New York: Harper and Brothers, 1956, p. 51.
 - ³ Munn, N.L., Psychology, the Foundation of Human Adjustment, (London: George G. Harpapp & Co., Ltd., 1961), p. 717.
 - ⁴ Murphy, G., and Newcomb, T.M., Op.Cit., p. 212.
 - ⁵ Lapiere, R.T., and Fransworth, P.R., Social Psychology, (New York: McGraw Hill Book Co., 1949), p. 506.

J.D, Frank¹ defines 'Level of aspiration' as the level of future performance in a familiar task which an individual knowing his level of past performance in that task, especially undertakes to reach.

According to Lawshe², it may be said to describe the condition of an individual's state of motives or goals at any given time. It is the level of future performance. In varying up and down, with success and failure, it keeps the goal ahead of actual achievement and thereby protects the individual from disillusionment, discouragement and behaviour unaccepted by society.

According to McGehee³, 1940, the process of setting a level of aspiration appears to have much in common with the process of formulating a psychological judgement. In the former definition of Lewin and Frank they said that level of aspiration is an outcome from conflict from several motives.

Lewin's View About Level of Aspiration

A theoretic concept of level of aspiration was first presented by Escalona (1940) and elaborated by Festinger⁴

1 Maccoby, E.E., Newcomb, J.M. and Hartley, E.L., Readings in Social Psychology, (New York: Rinehart and Winston, Inc., 1958), p. 290.

2 Lashe, H., Psychology of Industrial Relations, (New York: McGraw Hill Book Co., Inc., 1953), p. 26.

3 Atkinson, J.W., An Introduction to Motivation. (Princeton: New Jersey New York), p. 239.

4 Atkinson, J.W., *Ibid*, p. 149.

(1942). It is referred as the 'resultant valence theory' of level of aspiration. The psychological situation of a person at the moment, he is asked to state his level of aspiration, can be characterized as a choice situation. The choice is determined by the valences which different levels of difficulty within the same activity have for the person. Each level of difficulty may be represented as a separate activity or goal region in the life space of the person. The choice of a particular goals region 'L' i.e. to say the level of aspiration will be determined by the resultant force towards 'L' (f_p, L) where 'L' is the region and (f_p, L) is resultant force.

The individual faces the possibility of succeeding or failing whatever level he chooses, so that positive valence of future success and negative valence of future failure become the basic determinants of his decisions. The attractiveness of success seems to increase with the level of difficulty and vice a versa.

In setting level of aspiration, the person is simultaneously confronted with a number of different levels of difficulty which may be represented as separate activity - regions in the life space.

From the above description of these various techniques of situational tests, it should be clear that the setting for testing requires elaborate arrangements and a team of trained psychologists. Secondly, in the situational tests, the real purpose of the test is hidden and if the subject knows it, he can beat the examiners. Thirdly, the evaluations from the situational tests are inferences derived on the basis of data obtained by observation during the limited period of testing. These ratings are open to the same criticism as are the usual rating procedures.

Projective Tests

These are similar to situational tests taken unawares. The general idea behind these projective methods is to confront the subject with an unstructured, ambiguous situation.... asks him to do something with it, The subject is given several degree of freedom to organize a plastic medium in his own way, and since little external aid is provided from conventional patterns he is all but obliged to give expression to the most readily available factors within himself. It is further characteristic of

projective methods that the subject does not know what kind of inferences the experimenter intends to make.¹

The projective techniques are sort of indirect methods for studying the inner life of an individual. The basic theory underlying all these devices is that each person unconsciously 'projects' his private feelings and attitudes in his dealings with the everyday situations of the external world and his actions thus have a symbolic as well as literal references. Plastic materials (are employed) that permit a wide variety of symbolic structuralisations.²

The Essential Features of Projective Techniques:

Primary Criteria.- Perhaps the most distinctive feature of projective techniques is their sensitivity to unconscious or latent aspects of personality.

The second consideration of major importance is the multiplicity of response permitted by the subject by most projective techniques. It permits the subject to select the particular responses that he wishes to make from a theoretically unlimited number of response alternatives.

¹ White, R.W., Interpretation of Imaginative Productions, Inc., J. Mcy. Hunt (Edu.), Personality and the behaviour disorders Vol. I (New York: Ronald 1944) pp. 215.

² Goodenough, Florence, L., Mental Testing (New York; Holt, Rinchart & Winston, 1949) p. 562.

A further distinctive feature is multidimensionality of these devices.

A fourth feature of these tests is the lack of subject's awareness of the purpose of the test.

A fifth distinguishing feature of the projective technique is the profusion and richness of the response data they elicit.

Secondary Criteria.- One of the most frequently proposed qualities for distinguishing projective techniques from other tests is the ambiguity of the stimulus presented to the subject.

A further distinctive feature of the project tests is their appropriateness for holistic analysis.

A third characteristic that applies to most projective techniques has to do with their tendency to evoke fantasy responses from the subject. A high proportion of these techniques ask the subject to take an 'as if' set, to respond without reality restriction, to imagine or to invent.

Closely related to the fantasy aspects of these responses is the fact that subject's responses have no right or wrong status. The individual is to respond in whatever manner seems most natural and appropriate, with the

assurance that there is no criterion of correctness against which his responses can be arranged and scored. Some well-known examples of projective technique are described below:

The Rorschach Inkblots: The best known and most widely discussed projective technique is undoubtedly the Rorschach Developed by the Swiss psychiatrist, Herman Rorschach, this technique was first discussed in 1921.

The Rorschach utilizes ten cards. Each card is handed to the examinee with the instructions that give him the greatest possible freedom. Besides keeping a verbatim record of the subjects responses to each card, the examiner notes the true responses, positions in which the cards are held, spontaneous remarks during the test session etc.

The scoring of Rorschach ~~test~~ is highly complex and individualized. The most common scoring categories employed with the Rorschach include location determinents and content. A major complicating factor in the interpretation

1 Rorschach, H. (Transl. by P. Lemken and B. Kronebug) Psychodiagnostic: a diagnostic test based on perception. Berna: Humber, 1942 (1st German ad; 1921, U.S. distributor, Gruse and Stratton).

of Rorschach score is the total number of response known as response productivity or R. Because of large individual differences in R, the practice of considering the absolute number of responses in various categories is obviously misleading.

The most important advantage of this technique is the extensiveness with which it has been employed. Another advantage is the existence of a relatively well-structured system of scoring with a set of general interpretative principles attached to the results of scoring principles.

Thematic Apperception Test (TAT):

The TAT is found useful in any comprehensive study of personality, and in the interpretation of behaviour disorders, psychosomatic illness, neurosis and psychosis.

The TAT as devised by Murray¹ consists of a set of 20 cards, the nature of which is varied somewhat depending upon whether the subject is male or female, child or adult. In each case 19 of the cards present pictures of varying content and degree of ambiguity and the remaining one card is blank. The cards are presented to the subject individually with the request that he would create a story

¹

Murray, H.A., Thematic Apperception Test Manual, 1943.

about the picture that describes what is going on in the pictured scene, what the people are thinking and feeling, what led up to this scene, and what will be the outcome.

In interpreting TAT stories, the examiner first determines who is the 'hero', the character of either sex, with whom the subject has presumably identified himself. The content of the stories is then analysed principally in reference to Murray's list of 'needs' and 'press.'

A fair amount of normative information has been published regarding the most frequent characteristics of each card, including the way each card is perceived, the themes developed, the roles ascribed to the characters, emotional tones expressed, speed of responses, length of stories and the like. Although these normative data provide a general framework for interpreting individual responses, most clinicians rely heavily on 'subjective norms' built up through their own experience with the test. A number of quantitative scoring themes and rating scales have been developed that yield good scorer reliability.

Besides Rorschach and TAT, there are also other projective tests e.g. Word Association Technique, Michigan Picture Test, Blacky Picture, Rosenzweig Picture-Frustration Test, Scondi test, etc.

Critical Evaluation

Rapport and applicability:

Most projective techniques represent an effective means for 'breaking the ice' during the initial contacts between subject and examiner. It tends to divert the subjects' attention away from himself and thus reduces embarrassment and defensiveness.

Faking:

In general, projective techniques are less acceptable to faking than are self-report inventories. Moreover, the subject soon becomes absorbed in the task and hence is less likely to resort to the customary disguises and restraints of the interpersonal communication.

Standardization:

It is obvious that most projective techniques are inadequately standardized with respect to both administration and scoring. Even when employing identical instructions some examiners may be more encouraging or reassuring, others more threatening, owing to their general manner

and apperance. Such differences may affect response productivity, defensiveness, stereotype, imaginativeness and other basic performance characteristics.

Equally serious is the lack of objectivity in scoring. Even when objective scoring systems have been developed, the final steps in evaluation and integration of the raw data depend upon the skill and clinical experience of the examiner. In the first place, it reduces the number of examiners who are properly qualified to employ the technique and thus limits the range of its effective application. It also means that the results obtained by different examiners may not be comparable, a fact that complicates research with the instrument.

Norms:

Another conspicuous deficiency common to most projective instruments pertains to normative data. Such data may be inadequate or based upon vaguely described population. In the absence of adequate objective norms, the clinician falls back upon his 'general clinical experience' to interpret projective test performance. But such a frame of reference is subject to all the distortions of memory that are themselves reflections of theoretical bias, preconceptions and other idiosyncrasies of the clinician. Moreover, any one clinician's contacts may have been limited largely to

persons who are typical in education, socio-economic level, sex ratio, age distribution and other relevant characteristics. Interpretation of projective test performances often involves sub-group norms, of either a subjective or an objective nature. Thus, the clinician may have a general subjective picture of what constitutes a 'typical' schizophrenic or psychoneurotic performance on a particular test. Thus, the sub-group norms may lead to faulty interpretations unless the sub-groups were to be equated in other respects.

Reliability:

In view of the relatively unstandardized scoring procedures and the inadequacies of normative data, scorer reliability becomes an important consideration in projective testing. For projective techniques, a proper measure of scorer reliability should include not only the more objective preliminary scoring, but also the final integrative and interpretive stages.

Retest reliability also presents special problems. With long intervals, genuine personality changes may occur, which the test should detect. It is also relevant to note that many scores derived from projective techniques are based upon very inadequate response samples.

Ratios and percentages computed with such unreliable measures are even more unstable than the individual measures themselves.

Validity:

Most empirical validation studies of projective tests have been concerned with concurrent validity. A few studies have been investigated predictive validity against such criteria as success in specialised types of training or response to psychotherapy.

The large majority of published validation studies on projective techniques are inconclusive because of procedural deficiencies in either experimental controls, statistical analysis or both. Similarly the examiner may have obtained cues about the subject's characteristics from conversation with the subject in the course of test administration or from case-history material and other non-test sources. The customary control for the latter type of contamination in validation studies is to utilize 'blind analysis' in which the test record is interpreted by scorer who has had no contact with the subject and who has no information about him other than contained in the test protocol.

Inadequacies of experimental design may also have the effect of underestimating the validity of a diagnostic

instrument. It is widely recognized, for example, that traditional psychiatric categories such as schizophrenia, manic depressive psychosis, and paranoia represent crude and unrealistic classification of personality disorders actually manifested by patients. Hence, if such diagnosis categories are used as the sole criterion for checking the validity of a personality test, negative results are inconclusive.

Objective Tests

The term 'objective test' has not yet gained an unequivocal meaning as a category of personality tests. Many times objective tests are conceived as a very broad category of tests as against the subjective tests, and include the personality inventories, situational tests, etc., which are relatively objective approaches to personality assessment. But this term is being used now-a-days in a more restricted sense to denote only those experimental procedures which stand the tests of empirical validation. The application of experimental procedures in the field of personality assessment has been very limited. It might be due to the fact that this field is not amenable to treatment by experimental methods, or may be that those who worked in this field lacked this bias in their work.

Whatever be the reason, those working in this direction have created a ray of hope that this approach is likely to yield better results than other methods, all of which are relatively subjective.

Those who have been working with such techniques are mainly Eysenck, Benker, Harrington and Sears. These investigators tried to correlate the underlying psychological attributes of personality with more tangible and measurable characteristics, such as autonomic activity, dark vision, effects of aspiration, suggestibility, the effect of success or failure on activity and so on.

Many of these studies have been inconclusive. This approach has the chief merit that it is subject to greater control. At the same time, these techniques are very elaborate, expensive and time consuming.

Similar recent approach to measure personality through objective tests is adopted by a group of psychologists like Whitney who have undertaken experimental investigations to study personality through perception. In this approach a number of tests of perceptual processes are administered to subjects and the results are correlated with different personality traits.

Summary

Assessment of personality is done through different techniques, singly or in combination. Each one of them has specific merits and drawbacks. Nobody claims that any particular technique is infallible and should be preferred to all the rest.

In fact, each one has its place under particular circumstances. This suitability would depend upon the purpose of measurement.

It is likely that there may be found variation in attempts to measure some traits through different forms of personality also. The hypothesis deals with this aspect of personality. Attempt is made in the succeeding chapter to review briefly some of the works undertaken in this direction to study variation in measurement of personality by different forms of the tools.
