

THE FALLACY OF PERSONAL VALIDATION: A CLASSROOM DEMONSTRATION OF GULLIBILITY*

BY BERTRAM R. FORER

Veterans Administration Mental Hygiene Clinic, Los Angeles

THIS paper is concerned with some of the methodological errors which can affect estimations of the validity of personality interpretations and measuring instruments. Of prime significance is the nature of the interpretations themselves. Personality evaluations can be, and often are, couched in such general terms that they are meaningless in terms of denotability in behavior. Or they may have "universal validity" and apply to everyone. Bobertag (2) refers to the universally valid personality trait as *Universalscharakteristik*.

Possession of two eyes is a characteristic of all vertebrates, hence is of no value as a differentiating factor among vertebrates. The opposing thumb does not distinguish one human being from another. At the psychological level the acceptance of some cultural taboos appears to be universal among human beings who live within social groups. Virtually every psychological trait can be observed in some degree in everyone. For the purpose of characterizing a particular individual, stipulation of those traits which he demonstrates is a meaningless procedure. It is not in the presence or absence of a trait that individuals differ. The uniqueness of the individual, as Allport (1) amply documents, lies in the relative importance of the various personality forces in determining his behavior and in the relative magnitude of these traits in comparison with other persons. Thus the individual is a unique configuration of characteristics each of which can be found in everyone, but in varying degrees. A universally valid statement, then, is one which applies equally well to the majority or the totality of the population. The universally valid statement is true for the individual, but it lacks the quantitative specifica-

tion and the proper focus which are necessary for differential diagnosis. In a sense a universally valid statement is a description of a cultural group rather than a personal psychological datum.

A universally valid personality description is of the type most likely to be accepted by a client as a truth about himself, a truth which he considers unique in him. Many, if not most, individuals are able to recognize the characteristics in themselves—when it is not to their disadvantage—while oblivious to their presence in others. An example is the tendency for students to perceive their own problems in textbooks of abnormal psychology. In such cases the individual lacks the quantitative frame of reference necessary for a critical comparison of the printed description and his own self-evaluation.

At times confirmation by a client or by some other person familiar with his history is used as a criterion in the validation of diagnostic inferences and procedures (4). Test results may suggest certain problems and characteristic modes of behavior which therapists or the client, himself, can confirm or deny. Testing the correctness of inferences about a client by requesting his evaluation of them may be called "personal validation." When the inferences are universally valid, as they often are, the confirmation is useless. The positive results obtained by personal validation can easily lull a test analyst or a therapist into a false sense of security which bolsters his conviction in the essential rightness of his philosophy of personality or his diagnostic prowess. Such false validation increases his comfort in using what may have been a dubious instrument. A great danger arises when the confirmation of a prediction is extended uncritically to the instrument or conceptual system or person making the prediction. Such uncritical extensions occur too frequently in the clinical field.

Confirmation of a prediction does not necessarily prove the validity of the propositions from which the prediction was inferred. An identical prediction may be made from a group of propositions which contradict the original ones (3, p. 140). Taylor (12) has shown empirically that judges of case histories may arrive at identical predictions for different reasons. Confirmation of a variety of predictions which will differentiate among a number of clients is necessary if validation is to be accepted with any degree of confidence.

The crystal-gazer is likely to be aware of some of these points and other pseudo-diagnosticians, though they may be unaware of the fallacies inherent in their procedures, make effective use of "universal validity" and "personal validation" in deceiving their clients. Allport (1, p. 476) states that "one way in which character analysts secure a reputation for success is through the employment of ambiguous terms that may apply to any mortal person." A naïve person who receives superficial diagnostic information, especially when the social situation is prestige-laden, tends to accept such information.¹ He is

¹D. G. Paterson, in a personal letter to the writer, describes and includes a universally valid personality sketch which he uses in luncheon club lectures. It is reproduced here with his permission.

"Above average in intelligence or mental alertness. Also above average in accuracy—rather painstaking at times. Deserves a reputation for neatness—dislikes turning out sloppy work. Has initiative; that is, ability to make suggestions and to get new ideas, open-mindedness.

"You have a tendency to worry at times but not to excess. You do get depressed at times but you couldn't be called moody because you are generally cheerful and rather optimistic. You have a good disposition although earlier in life you have had a struggle with yourself to control your impulses and temper.

"You are strongly socially inclined, you like to meet people, especially to mix with those you know well. You appreciate art, painting and music, but you will never be a success as an artist or as a creator or composer of music. You like sports and athletic events but devote more of your attention to reading about them in the sporting page than in actual participation.

"You are ambitious, and deserve credit for wanting to be well thought of by your family, business associates and friends. These ambitions come out most strongly in your tendency to indulge in day-dreams, in building air-castles,

impressed by the obvious truths and may be oblivious to the discrepancies. But he does more than this. *He also validates the instrument and the diagnostician.* Crider's students (4) found surprisingly accurate the analyses they received from a pseudo-diagnostician. Crider, himself, seems to have been beguiled by the results and decries a priori rejection of the claims of these persons. While the use of matching procedures has revealed fairly high validity for inferences derived from projective tests by trained clinicians (6, 7, 8, 9, 10), it has not supported the claims of persons employing non-standardized graphological techniques (11).

Recently the writer was accosted by a night-club graphologist who wished to "read" his handwriting. The writer declined and offered to administer a Rorschach to the graphologist. An amiable discussion ensued, during which the graphologist ventured proof of the scientific basis of his work in that his clients affirmed the correctness of his interpretations. The writer suggested that a psychologist could make a blindfold reading and attain the same degree of verification.

EXPERIMENT

The following experiment was performed in the writer's class in introductory psychology to demonstrate the ease with which clients may be misled by a general personality description into unwarranted approval of a diagnostic tool. The writer had discussed his Diagnostic Interest Blank (5)² (hereafter referred to as DIB) in connection with the role of personal motivational factors in perceptual selectivity. Class members requested

but this does not mean that you fail to get into the game of life actively.

"You ought to continue to be successful so long as you stay in a social vocation. I mean if you keep at work bringing you in contact with people. Just what work you pick out isn't as important as the fact that it must be work bringing you in touch with people. On the negative side you would never have made a success at strictly theoretical work or in pure research work such as in physics or neurology."

²The DIB consists of a list of hobbies, reading materials, personal characteristics, job duties, and secret hopes and ambitions of one's ideal person. The test is interpreted qualitatively and personality dynamics are inferred along lines similar to projective tests.

* Published with permission of the Chief Medical Director, Dept. of Medicine and Surgery, Veterans Administration, who assumes no responsibility for the opinions expressed or conclusions drawn by the author.

that they be given the test and a personality evaluation. The writer acquiesced. At the next meeting the 39 students were given DIB's to fill out, and were told that they would be given a brief personality vignette as soon as the writer had time to examine their test papers. One week later each student was given a typed personality sketch with his name written on it. The writer encouraged the expressed desire of the class for secrecy regarding the content of the sketches. Fortunately, this was the day on which a quiz was scheduled; hence it was possible to ensure their sitting two seats apart without arousing suspicion. From the experimenter's point of view it was essential that no student see the sketch received by any other student because *all sketches were identical*.³ The students were unsuspecting.

The personality sketch contains some material which overlaps with that of Paterson, but consists of 13 statements rather than a narrative description. A further difference lies in the fact that this sketch was designed for more nearly universal validity than Paterson's appears to have been. The sketch consists of the following items.

1. You have a great need for other people to like and admire you.
2. You have a tendency to be critical of yourself.
3. You have a great deal of unused capacity which you have not turned to your advantage.
4. While you have some personality weaknesses, you are generally able to compensate for them.
5. Your sexual adjustment has presented problems for you.
6. Disciplined and self-controlled outside, you tend to be worrisome and insecure inside.
7. At times you have serious doubts as to whether you have made the right decision or done the right thing.
8. You prefer a certain amount of change and variety and become dissatisfied when hemmed in by restrictions and limitations.
9. You pride yourself as an independent thinker and do not accept others' statements without satisfactory proof.
10. You have found it unwise to be too frank in revealing yourself to others.

³ These statements came largely from a newsstand astrology book. The writer was not aware of Paterson's sketch at the time this problem was formulated and carried out.

11. At times you are extroverted, affable, sociable, while at other times you are introverted, wary, reserved.

12. Some of your aspirations tend to be pretty unrealistic.

13. Security is one of your major goals in life.

Before the sketches were passed to the students, instructions were given first to read the sketches and then to turn the papers over and make the following ratings:

A. Rate on a scale of zero (poor) to five (perfect) how effective the DIB is in revealing personality.

B. Rate on a scale of zero to five the degree to which the personality description reveals basic characteristics of your personality.

C. Then turn the paper again and check each statement as true or false about yourself or use a question mark if you cannot tell.

In answer to their requests students were informed that the writer had another copy of their sketch and would give it to them after the data were collected. After the papers had been returned to the writer students were asked to raise their hands if they felt the test had done a good job. Virtually all hands went up and the students noticed this. Then the first sketch item was read and students were asked to indicate by hands whether they had found anything similar on their sketches. As all hands rose, the class burst into laughter. It was pointed out to them that the experiment had been performed as an object lesson to demonstrate the tendency to be overly impressed by vague statements and to endow the diagnostician with an unwarrantedly high degree of insight. Similarities between the demonstration and the activities of charlatans were pointed out. That the experience had meaning for them was indicated by the fact that at least one-third of the class asked for copies of the sketch so that they might try the trick on their friends.

RESULTS

The data show clearly that the group had been gulled. Ratings of adequacy of the DIB included only one rating below 4. Thus the instrument received a high degree of personal validation. In the evaluation of the sketch as a whole there were five ratings below 4 (Table 1). While a few students

were more critical of the sketch than of the DIB, most of them were ready to admit that basic personality traits had been revealed.

The number of specific items accepted as true varied among the group from 8 to 13 except for one individual who accepted only 3 (Table 2). This same individual rated the test at 4 and the sketch at 2. Mean acceptance was 10.2 items.

No significant relationships were found between any of the ratings and sex, age, occupational background, or grades on the subsequent quiz.

considered sufficient evidence for acceptance of the sketch as perfect. For others, high, but imperfect, validity was indicated by the acceptance of 12 of the 13 items. It may be said, then, that among this group of students individuals varied in the degree to which they weighted the truth and falsity of the descriptive items in arriving at an overall evaluation.

Ratings of the DIB as a diagnostic instrument (rating A) and number of items accepted as true show no significant relationship (the probability value of the chi-

TABLE 1
DISTRIBUTIONS OF RATINGS

RATING	0	1	2	3	4	5	N
A (DIB)	0	0	0	1	25	13	39
B (Sketch)	0	0	1	4	18	16	39

TABLE 2
DISTRIBUTION OF "TRUE" RESPONSES

Number True	5	6	7	8	9	10	11	12	13	N
Frequency	1	0	0	5	5	10	9	7	2	39

In addition to the high ratings of the DIB which indicate a degree of gullibility or fallacious judgment, further evidence can be seen in the degree to which ratings were made on other than evidential grounds or contrary to the evidence. If the individual accepts all of the items as applying to himself, he is somewhat justified in accepting the instrument; if he rejects all of the items in the sketch, he is justified in rejecting the DIB.

The chi-square test indicates a degree of association, significant at the 1-per-cent level, between ratings of the sketch (rating B) and the number of items checked as true. However, the operation of other factors in judgment from part to whole is clearly indicated. For some individuals the presence of 8 true statements among the 13 was con-

square is .4). On the one hand, estimation of the adequacy of the personality sketch was partially dependent upon the amount of confirmatory evidence. On the other hand, the degree of approval of the test was independent of the degree to which test results agreed with self-evaluations. That is, validation of the test instrument was an all-or-none affair depending on a certain minimum amount of evidence. The amount of confirmatory evidence set up as a standard varied among the students.

All of the students accepted the DIB as a good or perfect instrument for personality measurement. Most of them can be accused of a logical error in accepting the test on such scanty evidence. Those who accepted the test with a rating of 5 while accepting fewer than all of the 13 statements have

demonstrated a disregard for the evidence of their own criticisms. The same can be said for those who rated the test higher than the personality sketch. It is interesting that the student most critical of the personality sketch, as indicated in an overall rating of 2 and acceptance of only 5 items, at the same time rated the DIB at 4.

The degrees of group acceptance for the 13 items are indicated in Table 3. None of the items attained complete universal validity, though more than half of them were close to complete group acceptance.

RECALL OF RATINGS

Since many of the class had indicated their embarrassment at having been "taken

lowered their ratings from 5 to 4. On the other hand, rating B (of the sketch) tended to be lowered. Seven ratings of 5 were lowered to 4 and one rating of 5 was lowered to 3. None was raised. The two distributions of ratings on the sketch are shown in Table 4. The *t*-test for differences between related means indicates significance at the 1-per-cent level. Thus, there is confirmation of a significant lowering in the level of acceptance of the sketch among those who had been most credulous.

CONCLUSIONS

1. Claims of validity for their methods and results by pseudo-diagnosticians can be duplicated or surpassed in the laboratory

TABLE 3

GROUP ACCEPTANCE OF SKETCH ITEMS

RESPONSE	ITEM NUMBER												
	1	2	3	4	5	6	7	8	9	10	11	12	13
True	28	38	23	31	18	35	38	37	34	35	34	12	28
False	4	0	1	0	9	3	0	1	3	2	1	9	7
Uncertain	7	1	15	8	12	1	1	1	2	2	4	18	4

in," the writer suspected that the dynamics of the memory process would operate in the direction of healing the results of this assault to self-esteem. The class had been informed of the distributions of ratings. Three weeks later the students were told that the writer had erased the names from their rating sheets as he had promised. Unfortunately he would have liked to compare their ratings with their grades on the quiz. Perhaps they would be willing to jot down from memory the ratings they had made of the DIB and the sketch. The rating scales were written on the blackboard. The students were understandably skeptical at first, but ultimately cooperative. Only 32 of the students were present who had taken the DIB and received the sketch.

Results were more or less as expected. In the case of rating A (of the DIB) no general trends were noted: two students raised their ratings from 4 to 5 and three others

TABLE 4
RATING B, ORIGINAL AND RECALL

RATING	2	3	4	5	N
Original	1	3	12	16	32
Recall	1	4	19	8	32

without the use of a diagnostic instrument. Blindfold personality estimates can be shown to be valid when the method of personal validation (confirmation by the client) is used for descriptive items of approximate universal validity.

2. Validation of a test instrument or of a personality sketch by means of personal validation is a fallacious procedure which presupposes objectivity of self-evaluation and an understanding of other persons on the part of the client.

3. Using the method of personal validation,

a fictitious personality sketch can easily deceive persons into approving a diagnostic device even when there is incomplete acceptance of the sketch itself. A minimum degree of correspondence between the sketch and self-evaluation appears to engender an attitude of acceptance of the total sketch and this attitude of acceptance is carried uncritically to the test instrument.

4. The personal validation procedure is likely to yield more fallacious results in the case of overall evaluations of a personality sketch than when specific statements are evaluated individually.

5. When self-esteem is threatened, memory functions operate in such a manner as to avert the threat and enhance self-esteem. Such memory changes are defensive distortions of recall rather than simple forgetting.

6. Clinical psychologists and others who make inferences about personality characteristics may be led into ascribing an excessively high degree of significance to these inferences. There is pressing need for clinicians to submit their own procedures, presuppositions, and, perhaps, projections to experimental scrutiny.

REFERENCES

1. ALLPORT, G. W. *Personality, a psychological interpretation*. New York: Holt, 1937.
2. BOBERTAG, O. Bemerkungen zum Verifikationsproblem. *Z. f. ang. Psychol.*, 1934, 46, 246-249.
3. COHEN, M. R., & NAGEL, E. *Introduction to logic and the scientific method*. New York: Harcourt, Brace, 1934.
4. CRIDER, B. A study of a character analyst. *J. soc. Psychol.*, 1944, 20, 315-318.
5. FORER, B. R. A diagnostic interest blank. (In press.)
6. HARRISON, R. Studies in the use and validity of the Thematic Apperception Test with mentally disordered patients. II. A quantitative validity study. III. Validation by the method of "blind analysis." *Char. & Pers.*, 1940, 9, 122-133; 134-138.
7. HARRISON, R. The Thematic Apperception and Rorschach methods of personality analysis in clinical practice. *J. Psychol.*, 1943, 15, 49-74.
8. HARRISON, R., & ROTTER, J. B. A note on the reliability of the Thematic Apperception Test. *This Journal*, 1945, 40, 97-99.
9. HERTZ, M. R., & RUBINSTEIN, B. B. A comparison of three "blind" Rorschach analyses. *Amer. J. Orthopsychiat.*, 1939, 9, 295-314.
10. MURRAY, H. A., & STEIN, M. Note on the selection of combat officers. *Psychosom. Med.*, 1943, 4, 386-391.
11. PASCAL, G. R., & SUTTELL, B. Testing the claims of a graphologist. *J. Personality*, 1947, 16, 192-197.
12. TAYLOR, D. W. An analysis of predictions of delinquency based on case studies. *This Journal*, 1947, 42, 45-56.