

CHAPTER IV

A BRIEF REVIEW OF RELEVANT WORKS

Interest in personality is so wide spread among the behavioural sciences that it would be pretentious to regard it as the province of psychologist alone. The economist, the political scientist and the educationist share the interest of psychologists in the study of personality. Different authors use different styles of forms of tests to assess personality. Some of the major investigations in area of personality assessment have been summarised below:

Klett and Lovaas¹ (1956) have investigated the social desirability scale values of the statements in personal preference schedule. Kletts judges consisted of

¹ Edwards, Allen, L., 'The Social Desirability variable in (Edited by) personality assessment and Research; The drydan Press, New York, 1957, p. 12.

high school males and females from three different socio-economic classes. He found that the standards of social desirability of high school males and females are much the same with respect to the statements investigated. This was also true when the two socio-economic groups were compared with another. Pooling the judgements from all his group, Klett obtained successive interval scale values for the statements. The product-moment correlation between Klett's scale values based upon the judgements of 206 high school students and the values I originally obtained from judgements of 152 college students was .93.

Lovaas (1956) translated the statements in the personal preference schedule from English into Norwegian and obtained judgements from 50 male and 36 female students at the gymnasium in Oslo, Norway, in the fall of 1954. The average age of his subjects was 17 years. Scale values for the statements were found by the method of successive intervals. The product moment correlation between the scale values based upon the judgements of Norwegian group and those based upon American college group of 152 judges was .78.

A set of 140 personality statements was judged for S.D. by a group of 152 judges (Edwards 1953 a). Scale

values for the statements were obtained by the method of successive intervals. The statements were then printed in inventory form and the inventory was administered to a group of 140 students. One of the statements in the inventory was "I like to be loyal to my friends". This statement had an extremely high social desirability scale value on the successive interval continuum and it was endorsed by only 6% of the students.

For the complete set of statements it was found that as the social desirability scale value of the statement increased, so also did the percent endorsing the statement. It was evident that probability of endorsement was linearly related to S.D. scale value. The product-moment correlation between the two variables was .87.

Calvin Wright (1957) has repeated this study with a minor variation. He gave the same 140 statements to a group of 127 college students. Approximately half of the students in the group were males and the other half females. He asked the students to rate the degree to which each statement characterized them on a nine-point scale. For each of the statements he found the mean rating assigned in self-description. These means were then correlated with S.D. (Social Desirability)

scale values of the statements. The correlation between the two variables was .88.

Kenny (1956) has scaled Zimmer's traits for S.D. using the method of equal appearing intervals. Judges consisted of 28 female and 39 male students at the university of British Columbia. The traits ranked in order from most to least socially desirable, in term of their scale values, were: "respectful, energetic, orderly, refined, ambitious, trusting, spontaneous, precise, determined, obedient, economical, ardent, persistent, deliberate, cautious, conventional, leisurely, daring, sentimental, poetic, emotional, wary, dominant, meek and lusty." Having determined socially desirability scale values for these traits, Kenny then stated them in first person form: 'I am respectful' and so on. A new and independent group of 65 college students responded to the traits in the inventory form and for this sample the probability of endorsement was found for each trait. The rank order correlation between probability of endorsement and social desirability scale value was .82.

Fujita (1956) scaled personality statements from the Personal Preference Schedule for social desirability

using as judges group of 50 Japanese - American college students. It was his high belief that the culture and background of this group differed in many respects from the college group of 152 judges he originally used in scaling the same statements for social desirability.

Navaran and Stauffacher (1954) asked a group of 25 nursing students to provide self-descriptions by ranking 15 personality traits from most to least descriptive. The same group then ranked the fifteen traits according to social desirability. Mean ranks were obtained for each of the traits under each of the condition of instructions. The rank order correlation between the mean ranks assigned to the traits in self-description and in social desirability was .90.

All these studies cited above have supported the statement made earlier that probability of endorsement of an item in a personality inventory is positively and highly correlated with the social desirability scale value of the item.

The attempt to control for social desirability in personality inventories by means of forced-choice technique assumes the unidimensionality of the construct.¹

¹ Wiggins, Nancy, 'Individual Viewpoint of Social Desirability, Psychological Bulletin Vol. 66 No. 2. August 1966.

Utilising a methodology which allows multidimensionality to emerge both within individuals and within stimuli, the present study demonstrated that at least six individual viewpoints exist with respect to social desirability judgements of selected MMPI items. The perceptual spaces of six "idealized individuals" differed from one another and from the perceptual space of the 'average' individuals. It was concluded that controlling social desirability by pairing of items on the basis of 'average' social desirability scale value was an inadequate control at the individual level.

That the social desirability is prominent influence in personality tests is witnessed by the substantial correlations obtained between item endorsement frequency and the items independently scaled social desirability value for such diverse media as the Edward Personal Preference Schedule (EPPS) in the true-false format (Edwards 1953), the interpersonal check list (Edwards 1957 a), the Q sort (1955) and selected MMPI Scale. The S.D. is confirmed in meaning to the set of operations necessary to assign a numerical social desirability value to an item. In distinction the term 'social desirability set' is defined as the subjects tendency to respond to an item in the socially desirable direction.

In the literature of the past few years a number of individual difference measures of the response styles of social desirability and acquiescence have appeared, employing explicitly or implicitly the concept of controversiality, communality and related social desirability. Hanley (1961) has recently suggested a classification scheme for social desirability scales based on whether response frequencies played a role in item selection and whether the social desirability of the items was determined explicitly or implicitly. Another basis for classifying social desirability scales which has been stressed by Wiggins is the method of contrasted groups, in which responses of a group of subjects who because of special instructions or special circumstances are considered to be a group of high social desirability respondents are compared with the responses of a control group.

The method of contrasted groups was employed by Cofer, Chance and Judson (1949) and Wiggins (1959) in the development of two role playing scale of social desirability. The responses of a control group were contrasted with those of subjects instructed to answer MMPI in terms of social desirability. In construction of

L scale of MMPI, a group of two clinicians constructed items in terms of social desirability and guessed at the frequencies that would occur in a control group. The 'K' scale of the MMPI was constructed by comparing the responses of one group of presumably non-faking patients with those of another group thought to include a large number of social desirability respondents. Unfortunately, the interpretation of the K scale is further complicated by the fact that an additional set of items was added to the scale that had been shown not to discriminate between role playing and control college groups.

In another study of the relationship between probability of endorsement and S.D. scale value for a set of 2824 personality statements, a group of 47 male and 48 female judges rated 2824 personality statements for S.D. using a nine point rating scale.¹ Another group of 110 male and 111 female subjects described themselves in terms of the same set of 2824 statements by answering each one 'true' or 'false'. The correlation between probability of 'true' response and social desirability scale value for the combined sex groups was .892. The distribution of the social desirability scale value of the 2824 statements was

¹ Edwards, A.L., 'Relationship between Probability of endorsement and Social Desirability scale value for a set of 2824 Personality Statement.'" Journal of Applied Psychology. Vol. 50. No. 3 June 1966.

distinctly bimodal. These results are in accord with another large scale study in which 1647 personality statements were investigated. In view of the large personality statements involved in these two studies, it is suggested that a correlation of .90 between probability of endorsement and social desirability scale values and a bimodal distribution of the scale values of the personality statements may be characteristic of the population.

One of the most active research areas in psychology during the last decade has been the study of test taking response sets or styles.¹ In particular, attention has been devoted to investigating their influence on personality inventory scores. This work has been confined almost exclusively to only three types of response tendency: (i) the social desirability set, characterized by consistent endorsement of desirable traits and the denial of undesirable ones, (ii) the deviation of a pattern of scores from a typical pattern produced by a given population of responses, and (iii) the acquiescence set which consists of tendencies to choose the "true", "agree" or "like"

¹ Edwards, A.L., Response, Style as a Personality variable - by what Criterion? Psy. Bulletin Vol. 59, No. 4 July 1961.

option rather than respective negative alternatives.

The degree to which individual subjects tend to give socially desirable responses in self-description on each of B MMPI scales was determined by computing, for each subject, a point-biserial correlation between his true/false responses and the social desirability scale values of the items.¹ The results indicated that on the average, from 10% to 20% of the variance in individual self-description was associated with SD scale values. In comparison with these results, individual differences in the tendency to give S.D. responses, as measured by these point biserial correlations and Edwards SD scale, account for much more of the variance in raw scores on many of the MMPI scales. The differences between these methods of appraising SD aspects of test performance and implication for constructing SD free scales were considered.

Arthur Couch and Kenneth Keniston's study attempt to analyze agreeing response set as a personality variable - using both objective test measurements and a programme of clinical assessment. The first part of this report outlines the development of a reliable and valid measure of agreeing

¹ Boe, E.E., and Kogan, W.S., Social Desirability in individual performance on 13 MMPI Scales. British J. of Psy. 1966. 57 (1-2), 161-170.

response set as it manifests itself in questionnaire items. The pattern of results suggested a set of hypotheses concerning the personality differences between individuals who tend to agree (yes-sayers) and those who disagree (no-sayers). The second part of this report presents a set of hypotheses about the personality dynamics behind the agreement tendency and attempts to validate these formulations by a clinical study of subjects with marked tendency to 'agree' and "disagree". Results of a structured interview and a sentence completion test given to these "Yes-sayers" and "No-sayers" enabled the investigatory to verify and conceptualize more clearly the inner dynamics behind this 'agreeing' personality syndrome.

The measure of response bias showed considerable internal reliability, consistency over time and generally over tests. The agreeing tendency was highly related to 'true saying' 'yes-saying' on other standardized tests. When a factor analysis of individual items was performed, a major dimension of stimulus - acceptance VS stimulus - rejection emerged as control to the agreeing - disagreeing response set. Personality test results also indicated pervasive difference in ego functioning, particularly as

regards the high VS low psychological inertia of secondary processes.

Development of the Over-all Agreement Score (OAS):

As a part of the assessment procedure in the selection of subjects for a clinical research project, a large battery of objective personality tests was administered to 61 paid volunteer students from Harvard college. The subjects gave responses to 681 items, using a scale of the Likert type. The response categories were given value from 1 to 7, (1) Strongly disagree (2) Disagree (3) Slightly disagree (4) No answer (5) Slightly agree (6) Agree (7) Strongly agree.

An over all agreement score (OAS) was computed for each subject by taking the mean of their responses to the 360 items. The range of scores was from 3.1 to 4.5, with a mean value of 3.9 and a standard deviation of 0.3. The first sign of the importance of this measure was indicated by the high (+ .85) spearman - Brown split - half (even-odd) reliability of the entire 360 items scale. The OAS thus provided a reliable measure of agreeing response set. The results also demonstrated that the OAS was most highly related to a response preference for the categories of 'Agree' (in a positive way). These correlations were affected by the relative contribution to the OAS of each

response category - as reflected in the Mean Percentage Frequency. This "part-whole" effect may explain the lower correlations for 'strongly agree' and 'strongly disagree', but since other response categories have about the same mean frequency, the uniquely high correlation of the 'Agree' and 'Disagree' categories point to a real relationship of response preference and the agreement tendency. The preference for definite "agreement" responses support the interpretation of the OAS measure. It should also be noted that the frequency of using 'disagreement' categories is an equally prominent indicator of the agreement tendency - only the direction of scoring is reversed. The results obtained in developing in Overall Agreement Score, namely the high reliability, stability over time and generality of agreeing response set support the view that this set is a manifestation of a personality syndrome. The score of OAS was then correlated with a series of objective tests domains. This step-wise research for the correlates of the agreement response set gradually built up a consistent and meaningful pattern which characterized the personality of individuals who tended to agree on questionnaires regardless of item content.

In the past half century, various aspects of test administration and scoring procedures have been studied, since their standardization has been regarded as the cornerstone of valid and reliable testing. Effects of variation in test instructions and phrasing of questions on test results are currently receiving much attention under such topics as social desirability and response sets.

If, however, a testee is in fact influenced by general social considerations (social desirability) and by personal tendency to agree or disagree (response sets) when he reads and reacts to test questions, then he should also be thus influenced when he receives and reacts to test instructions. Therefore, if a test is introduced differently to two groups of subjects as an instrument measuring traits A and B respectively, and if trait A is socially more desirable than trait B, it should then be predicted that subjects will be differentially motivated, resulting in differential achievement, under the two instructional conditions. An attempt to study this type of hypothesis has been made on intelligence test.

The Kuhlmann - Anderson Test was administered to a total of 557 subjects of fourth, seventh, tenth and twelfth grades under three different instructional sets.¹ In each grade,

¹ Kaory Yamamoto and Dizney, F. Henry, 'Effects of three sets of test instructions on scores on an intelligence scale' - Edu. and Psy. measurement Vol. XXV No. 1 1965. pp. 87-94.

subjects were randomly assigned to three groups and the same test was introduced to these groups as a test of intelligence, a test of achievement and a routine test respectively. Because of the differential social emphases upon intelligence and achievement, it was predicted that: (1) Performance of the first group (intelligence). Would be the best, followed by the second group (achievement), and the third (routine) in social order, (2) That preceding differences would be greater as subjects became older and (3) That sex differences would interact with grade level.

When the results were converted to deviation IQs and analysed by an analysis of variance, it was found that (1) the mean IQ obtained by the intelligence group (116.3) was significantly higher than those obtained by the achievement group (112.1) and the routine group (109.1), while the latter two group means were not significantly different, (2) girls obtained a mean IQ (114.4) significantly higher than that (111.7) obtained by boys, and (3) no other effects, including the (instructional set-grade) interaction and the (instructional set sex) interaction were statistically significant.

In another study, Rosen (1956) gave subjects a shortened form of the MMPI under two conditions: under the standard instructions to provide self-description and under instructions to give what he has called personally desirable responses.¹ Taking the records of subjects obtained under the desirable instructions, Rosen counted the number answering 'True' to each of the 409 items. Obtaining the correlations between these counts separately for males and females, Rosen in both instances found product-moment correlations of .87. These correlations are identical with the correlation of Edwards, A.L., obtained between probability of endorsement and social desirability scale value of an item in his study. In Rosen's study, a discrepancy score was obtained for each subject. The score was based upon the no. of items answered in one way under standard instructions and in a different way under instructions to give desirable responses. The discrepancy score was considered a measure of the extent to which an individual's responses to items were influenced by the special instructions to give desirable responses.

¹ Edwards, A.L., 'Social Desirability Variable in Personality Assessment and Research' New York. The dryden Press, 1957, p. 12.

In the study,¹ "Test Wiseness" on self-report personality scales was explored, using measures of accuracy in estimating the frequency of endorsement of personality items, estimating their social desirability, and identifying and "keying" items that measured the same factor, as well as indexes of ability to change scores on standard personality scales when they were administered with fake-good and fake-bad instructions. These variables generally did not correlate with each other and they had only moderate and scattered correlation with personality scales administered with standard instructions. The test-wisness measures were generally uncorrelated with ability and cognitive style tests and defensiveness scales, but they did correlate consistently with social desirability response style scale.

The scientific focus of another study² was to investigate the differential effects of employing first person and

¹ Stricker, J. Lawrence., 'Test Wiseness on Personality Scales' Ja. of Applied Psychology monograph. Vol. 53, No. 3 Part II June 1969.

² George, Stricker and Darrall D. Dawson, 'The effect of the first completion responses' and Ja. of projective techniques and personality assessment Vol. 30 No. 2 April 1966 pp. 169-171.

third person instructions and stems, varied separately to produce four test forms, on the responses of male and female psychiatric in-patients as scored by an objective scoring system. The sample consisted of 72 male and 72 female patients of mixed diagnosis drawn from comparable male and female wards. Two alternative forms of the Rotter IsB were employed. Form S1 consisted of the 22 items from the ISB which contained a first person pronoun in the stem, with the 18 items without a pronoun omitted. Two different instructions were also employed. Instruction II was the standard self reference instruction employed with the ISB and the subject was to express his own 'Real feelings'. Instruction I asked the subject to express "The real feelings of the average person of his sex and age." A three dimensional analysis of variance was utilized to analyse the data. Separate analyses were performed for dependency, anxiety and hostility scores. No significant effects were obtained in analysing hostility scores. The only significant main effect with dependency and anxiety scores was one of sex, with males producing more anxiety responses and females more dependency responses. Aside

from the secondary substantive finding that the ISB was more effective in eliciting dependency responses from females and anxiety responses from males, the major methodological focus showed no consistent differences in efficacy among the various forms employed. The implication of this finding, if it can be appropriately extended, is that the profusion of sentence completion forms utilising a multiplicity of combinations of instructions and stems serve to reduce our ability to utilize the research literature with little likelihood that a dramatic break through in efficacy will be achieved.

Another study¹ reports on the correlations between the 14 scales in booklet IB of the EPI and the same scales when the items are stated in first person form. In a large scale testing project, 171 females and 115 males were administered Booklets IA, II, III, IV and IB of the EPI in that order, at one testing session. Then they completed the Allport, Vernon, and Lindzey study of values. At a second testing session, two days later, the same subjects

¹ Edwards, A. L., 'Correlations between Scores on Personality Scales when items are stated in the first person - third person form' *Edu. and Psychological measurement*. Vol. 29, No. 3, Autumn, 1969, pp. 561-563.

were administered the California Psychological Inventory and then Booklet Ic. Booklet Ic contained the same items in the same order as the items in the booklet IB of the EPI, except that all items were stated in the first person form instead of the third person form. After completing booklet Ic, the subjects was administered a number of other personality inventories.

Correlations between the scales in booklets IA and IB were calculated with those in booklet IC. The correlation between the scales in the two comparable forms, IA and IB, were also found out. It was noted that the values of r_{AC} are quite comparable to the values of r_{AB} . Correlations of the scales with the social desirability (S.D.) scale of the Minnesota Multiphasic Personality Inventory, the Kuder-Richardson Formula and means and SD of the scales also tend to be much the same for corresponding scales in all three test booklets. Because Booklet IA of the EPI is a comparable form of booklet IB, it seems reasonable to believe that similar results would have been obtained if the items in booklet IA had been put in first person form rather than those in Booklet IB. It is possible that translating the items in Booklet III, IV, and V of the EPI into first person form might result in scores that differ considerably from those obtained with the standard EPI booklets.

A limitation of this study is that the same subjects responded to the items in each of the three tests booklets. Thus there is no way knowing to what degree the scores obtained with booklet Ic were influenced by memory and other carry-over effects resulting from the corresponding values for booklet IA and IB.

The purpose of yet another study¹ was to compare the personality questionnaire stems on the 16 PF Test, when they were presented to the college students in the second person as also in the first person statements and to find out those areas of personality where a change from "You" to "I" forms of items brings about a predisposing effect on the responses of the subjects. Further, it was also to present evidence of subject's test taking attitudes, specially of "Lying", in taking the personality questionnaire and to show its relation with scores on other factors.

The results of the above study indicated that a change in the item-form from second person to first person pronoun does have a predisposing effect on the response attitudes of the subjects in general. The difference was significant in the scores of six out of sixteen factors assessed. These are the factors of Ego strength, Super-ego

¹ Kapoor, S.D., 'Comparision of the personality questionnaire items presented in the first person and the second person An abstract of the PhD. thesis awarded under the supervision of Prof. S. Jalota, Banaras Hindu University, 1965.

strength, Self-sentiment formation, timidity, Ergic tension and Parmia. Besides, the correlational studies with the lie-score indicate that the said factors have in general higher correlations with the lie scores than other factors.

Thus it would be seen that a number of studies have investigated the efficacy of different methods of forms of Tests for personality measurement. Results have been of varying nature. One more attempt on similar lines have been made by the present author to study some specific issues in methodology of personality measurement.
