

12/11/71 P.R. only

Pam 159.923.072 TEM

19

Mastercopy



SPECIAL REPORT

PERS 151

THE ASSESSMENT OF PERSONALITY
USING PSYCHOLOGICAL TESTS

An Overview of the Theoretical Background and
Current Status of Personality Measurement

NATIONAL INSTITUTE FOR PERSONNEL RESEARCH
COUNCIL FOR SCIENTIFIC AND INDUSTRIAL RESEARCH

001.3072068 CSIR NIPR PERS 151

CSIR SPECIAL REPORT NO. PERS 151

UDC 159.923.072

Johannesburg, South Africa.

August 1971.

HSRC Library and Information Service

HSRC
Private Bag X41
PRETORIA
0001

Tel.: (012) 202-2903
Fax: (012) 202-2933



RGN
Privaatsak X41
PRETORIA
0001

Tel.: (012) 202-2903
Faks: (012) 202-2933

RGN-Biblioteek en Inligtingsdiens

PB 95809

THE ASSESSMENT OF PERSONALITY USING PSYCHOLOGICAL TESTS

An Overview of the Theoretical Background and
Current Status of Personality Measurement

RECEIVED
S.P.R. P.O. BOX 10319
JOHANNESBURG
1971
PSYCH
JOHANNESBURG



HSRC Library and Information
Service

RGN-Biblioteek en Inligtingsdiens

DATE DUE - VERVALDATUM

--	--

N

ARCH

COUNCIL FOR SCIENTIFIC AND INDUSTRIAL RESEARCH

Johannesburg

August, 1971



* PB 95089 *

RGN BIBLIOTEK

1995 1.03.

H SRC LIBRARY

STANDKODE	AANWINSNOMMER
COI 3072068	PB 005006
CSIR NDR	
PERS ISI	

ACKNOWLEDGEMENTS

This report represents the background research to the following project: Personality Evaluation Questionnaire. (7035/4387).

This project is directed by Mr. D.J.M. Vorster, Director of the National Institute for Personnel Research.

The research is being carried out as part of the activities of the Psychometrics Division under the guidance of Mr. D.W. Steyn.

SUMMARY

This report presents a comprehensive review of the theoretical background and current status of personality measurement. The particular emphasis of the review is on the questionnaire method of personality assessment and on the construction of questionnaires which will accord with the highest psychometric criteria of adequacy. The report begins by considering such background questions as the definition of personality and the psychological processes underlying measurement. The various general approaches to test construction are compared and evaluated, and special note taken of the contribution of Guilford, Eysenck, Cattell and Comrey to personality assessment. Such methodological questions as the role played by response style variables are considered, as well as the important topic of the validity of personality assessment. The report concludes that adequate personality measurement can be achieved using questionnaires, provided the construction of the questionnaire is characterized by methodological rigor and sensitivity to theoretical issues. A comprehensive bibliography on the questionnaire approach to personality measurement forms part of the report.

TABLE OF CONTENTS

	Page No.
1.0 Introduction	1
2.0 The Question of Definition	2
3.0 The Process of Measurement	4
4.0 Methods of Assessment	6
5.0 Test Construction	8
5.1 General Approaches to Test Construction	8
a. Global-Rational Approach	9
b. Separated-Rational Approach	10
c. Empirical Approach	11
5.2 Some Practical Guidelines for Test Construction	12
6.0 The Contribution of Guilford, Eysenck, Cattell and Comrey	14
6.1 Guilford's Contribution	15
6.2 Eysenck's Contribution	15
6.3 Cattell's Contribution	17
6.4 Comrey's Contribution	21
6.5 The Relationship Between the Personality Dimensions of Guilford, Eysenck, Cattell and Comrey	23
7.0 Methodological Concerns	26
7.1 Introduction	26
7.2 The Vanishing Variance Component	26
8.0 The Validity of Personality Assessment	32
9.0 Conclusion	34
10.0 Bibliography	35

1.0 INTRODUCTION

Systematic and scientific research into the area of personality developed out of initial attempts to measure individual differences in personality. Personality testing has never really lost its position at the forefront of personality research and remains an important topic today. There are differing opinions however, as to the future prospects for personality testing. Certain psychologists hold pessimistic views and point to the frequent failure of current personality tests to demonstrate adequate validity, and their failure to demonstrate the theoretical relevance of many of the traits they measure. There has even been some disillusionment expressed about the very possibility of the valid measurement of personality.

Such a pessimistic viewpoint of the prospects for personality testing is however, an overreaction to some of the difficulties of assessing personality. Jackson (1970) believes that very real progress has been made in research into personality and that this progress can be translated into valid and meaningful measures of personality. Some encouraging prospects for the assessment of personality are the advent of the computer and the presence of a vigorously growing empirical and theoretical background to personality measurement. This is not to deny that the measurement of personality is a complex undertaking. Even after a vast number of measures have been made, one may be little further along the road to the understanding of an individual's personality. Perhaps "understanding" is a mirage-like goal which can never be achieved by, and should not be the sole aim of personality measures.

One fact which will become clearer as this review continues is that many of the current tests of personality have not made use of this background of theory and research to which Jackson points. Far too many devices appear to have been selected more from the point of view of availability and expediency than of actual appropriateness to the problem being considered. A vast number of personality tests have been

devised to measure just about every conceivable personality "factor". Unfortunately there is a lack of clear evidence that the various personality "factors" carrying the same label do in fact coincide. Many of the concepts in the personality field are so broad and heterogeneous that the same concept does not mean the same thing when applied to different people. One of the results of these problems is that many of the trait measures lack adequate convergent and discriminant validity. The central problem in personality assessment is not methodological ineptitude but this lack of consensus in conceptualization. There are far too many definitions of personality. This introduces the discussion of definitions in personality assessment.

2.0 THE QUESTION OF DEFINITION

The question of definition in the assessment of personality is crucial since it is without doubt a fact that the type of research done and the measures used depend on the definitions of personality used. Certain workers such as Marlowe and Gergen (1969) claim that the term "personality" has outlived its usefulness in any exact denotative sense. They suggest that the term should simply refer to an area of investigation and no more. This approach to the definition of personality is an unrealistic one for it is merely a vain attempt to side-step the issue, - and this issue cannot be avoided if coherent personality measurement is to be achieved. Though they do not give a systematic definition of personality, it is clear from their article that Marlowe and Gergen have worked within the framework of a definition. They see personality as a dynamic and dependent relationship affecting social interaction.

On the one hand, personality connotes the human qualities of an individual, and yet on the other hand also implies that which is unique about a person. A definition of personality which takes both these ideas into account is one from the Oxford English Dictionary: "That quality or assemblage of qualities which makes a person what he is, as distinct

from other persons". A definition must take both the idea of "organization" and "distinctiveness" of personal attributes into account. Because of the difficulty in defining personality, psychologists in practice have generally confined their study to certain specific personality attributes, rather than considering the total field at one time. Thus it has become the practice to make reference to particular personality traits of an individual in defining personality. This has made assessment easier by encouraging the use of multi-dimensional and multi-factor instruments which are more reliable and valid than the earlier global devices. Since many workers now use the same plethora of "factors" in different contexts however, it has reduced the possibility of agreement upon the precise definition of concepts for personality assessment.

As might be expected there is not too much agreement about which factors fall into the personality domain. Personality has almost been defined by default, - what is left over after covering human ability, interests and attitudes. Holtzman (1964) points out that there is a similar lack of agreement about what is meant by personality assessment. What a given psychologist means by personality assessment seems to be determined mainly by the particular techniques he chooses to employ and the assumptions implicit in them. It is surely a more satisfactory procedure to make some effort to clarify what is involved in personality assessment before starting to construct measures of personality.

From the point of view of constructing personality tests it is necessary to adopt some type of working definition of personality. Thus it might be advisable to consider only those personality traits which can be used to describe the majority of individuals in populations on which the test will be used. In addition it might be necessary to consider only those personality differences which are observable by a number of observers, i.e. which result in high interrater agreement.

Before leaving the topic of definition in personality assessment, the importance of theory in personality test construction should be noted. The value of theory in scale

construction is the requirement it places upon explicit definitions of what is being measured. Unfortunately advances on a theoretical level seem to lag far behind the production of new scales. A test may continue to be used in the absence of any convincing evidence in its favour. There is a need to coordinate measures with concepts and theory in order to reduce the number of trivial measures which are constructed.

3.0 THE PROCESS OF MEASUREMENT

An area in which more theoretical and empirical knowledge is vital is in the process of measurement. There is no substitute in the creative task of defining dimensions and constructing personality tests for a sensitivity to the diverse ways in which psychological tendencies can be revealed in behaviour. The fundamental question is, how many things must we really know about an individual to measure his personality.

It has long been a criticism of personality tests that they do not consider a sufficient number of variables in measuring individual differences. A few traits are measured in isolation, not realizing that an individual's personality is part of a total system. Dahlstrom (1970) points to the need for a person-orientated, rather than a variable-orientated approach. Research and measurement should not only consider discrete variables, but also the fully integrated and functioning personality system interacting with the environment. It is particularly necessary that test constructors realize that the test behaviour of subjects is itself meaningful and worthy of investigation.

The items of a tests are most certainly not the only stimuli for test responses. The context of testing will determine responses as well as purely dispositional variables. It might be valuable to include a post-test interview to ascertain what variables are operating and in what ways, such as: the subject's perception of the study; his reactions to the procedures; and why he behaved as he did. Sechrest (1968) makes an

important point in emphasizing the need to analyse all the responses a subject makes in the test situation. It might turn out that what is important is not what the subject responds to the test stimuli, but how he makes these responses.

The test situation is not an isolated one for the subjects; he brings with him all the prejudices, biases and cultural variables which characterize him in everyday life. An individual's responses to tests may reflect overall cultural values rather than his unique opinions. Eisenman and Hare (1970) demonstrated this in a study of three different cultural groups of subjects. They administered five personality tests to 121 subjects from a Quaker secondary school, 32 subjects from a liberal arts university, and 70 subjects from a nursing school. They examined the subject's responses whenever any one group agreed 80% or more. They found high commonality of responses to the personality tests, from which they were able to infer the values of the groups. Holtzman (1964) points to the danger that we may be building a theory and technique of assessment which is culture-bound with a western, industrialized society. It must not be assumed that a test based upon university students and hospitalized schizophrenics can be meaningful to widely differing subjects.

Though there are a large number of variables which must be taken into account in personality measurement it should not be concluded that adequate assessment is not possible. There is a number of techniques and procedures which have made it a distinct possibility to do justice to the complexity of personality variables. Research workers and test constructors in the field of personality do seem to be taking more account of the vast number of variables which should be considered. The advent of the computer has made it relatively easy to adopt complex, multivariate approaches to assessment. An increasing number of studies, such as that of Kusyszyn and Jackson (1968) are using factorial methods such as multitrait-multimethod factor analysis to clarify relationships between the personality variables.

A study such as that of Turner and Fiske (1968) reveals that it is possible to investigate the processes underlying measurement. They studied the way subjects go about answering questionnaires and what types of items elicited higher frequencies of so-called "appropriate" response processes. They revealed that more acceptable items could be defined in terms of certain psychometric criteria such as item-test correlations and stability.

It should be clear from the research reviewed that a great deal more must be done to learn more about the process of measurement. Certainly it is vitally important that this type of research be undertaken at the same time as work in the development of further measures of personality.

4.0

METHODS OF ASSESSMENT

One thing that can certainly be said about the field of personality assessment is that there is no lack of variety in the measuring devices. There are measures of almost every conceivable characteristic using a number of different techniques. The techniques available for the assessment of personality can be divided roughly into three types: observational techniques; objective test techniques; and questionnaire and self-assessment techniques. The concern of this report is primarily with the questionnaire techniques.

Observational techniques are probably the oldest available and are still in use - particularly in survey-type research. These techniques include situational sampling and the contrived situation experimental method. One of the main reasons why these methods are not widely used is that they are essentially individual methods, not easily applied to group measurement.

Perhaps the most significant recent contribution to objective test techniques has been the publication by Cattell and Warburton (1967) of their compendium of objective tests of personality and motivation, providing more than 600 performance variables. The meaning Cattell attaches to "objective" is the goal of every test constructor: for Cattell "objective" means

perfect inter-scorer agreement, and a test situation and mode of response in which the subject is not able to fake the response to fit a subjective self concept. Unfortunately a careful look at Cattell's objective tests suggests that many of the tests are not as subtle as the authors might have hoped. Objective tests do have many potential advantages but have not really come into their own as yet. Perhaps because of the high degree of ingenuity required to construct objective tests of personality there are relatively few available. Those tests which are available have not been supported by much research.

It can be said without any fear of contradiction that questionnaire and self-assessment techniques are by far the most widely used instruments in personality measurement. The main reason for this is that they are easier to construct than objective tests and quickly administered to any number of subjects. Sarason and Smith (1971) point out that new scales have been developed to measure areas as diverse as: empathy, social caution, suspiciousness, moral attitudes, impulse control, defense mechanisms, novelty seeking, fear of death, and personality conflict. The vast number of questionnaires available should not be taken as indicating the necessary excellence of this method of personality assessment. As things stand at present, however, the questionnaire approach remains the most useful one for group personality assessment. Research must continue in an effort to devise the most acceptable questionnaires.

It is encouraging to note the growing number of personality inventories which measure a number of dimensions, rather than just a few characteristics. In addition, many of these inventories such as the Comrey Personality Scales (Comrey, 1970) are based upon a substantial body of theory and appear to be carefully constructed. Of particular significance is a paper by Jackson (1970) which sets out a sequential system for personality scale development. Following Jackson's outline it should be possible to produce a psychometrically elegant device. Certainly there are some challenging possibilities for the construction of new measures of personality.

5.0 TEST CONSTRUCTION

5.1 General Approaches to Test Construction

After a consideration of the personality measures which are in existence and following an outline given by Fiske and Pearson (1970), it can be said that there are three general approaches to test construction: Global-Rational; Separated-Rational; and Empirical. The Rational approaches begin from a priori-"rational" notions about the nature of the concept being measured. Tests using any particular measuring technique are constructed to cover the preconceived ideas about the concept. In contrast, the Empirical approach begins by deriving measuring instruments by operational techniques such as factor analysis, and then goes on to conceptualize or, more typically to label its variables by induction. The Global-Rational and Empirical approaches assume that their measures cover the total domain of the concept being measured. The Separated-Rational approach, however, tends to take the position that tests only cover part of the construct being measured and so quite frequently breaks up the construct into a number of parts or "facets".

Taking into account the tremendous complexity of the personality domain it is perhaps more realistic to take the position of the Separated-Rational approach, that measures only cover part of a total concept. On the other hand it is frequently more useful to have a measure which covers the total domain of a concept. Since the Empirical approach makes no a-priori assumptions about the nature of the concept considered, it often results in tests of higher psychometric standing than the Rational approaches. Empirical tests are relatively more "objective" than Rational devices and tend to have taken into account the fact that observations vary with the stimulus, the environment and the subject's perceptions. The advantage of the Rational approach is that it is less likely to result in tests that are statistical artifacts than the Empirical approach, and that it produces measures of theoretically known concepts, thus permitting construct validation operations.

It is not possible to state categorically that one approach is better than another, for the various approaches have advantages and disadvantages. It is clear that individual approaches are best suited for different requirements and so the purpose of a test ideally should determine the approach used. In the light of these last points it is not surprising to find that many test constructors cannot be categorised as belonging to a particular test construction approach. The inventories constructed by Eysenck (1959) would seem to fall into the Empirical approach, but those constructed by Cattell (Cattell and Eber, 1949) and Comrey (1970) fall into both the Global-Rational and the Empirical approaches. Cattell and Comrey begin from rational definitions of their constructs, but then go on to refine and develop instruments using techniques typical of the Empirical approach. In considering the important workers in the field of personality assessment in the next section of this paper, it will be possible to see to what extent the approaches are combined in practice.

Though the three approaches are combined to some extent in practice, it is valuable for the purposes of this paper to consider briefly some of the work done within each approach. For a fuller survey the reader is referred to Fiske and Pearson (1970).

(a) Global-Rational Approach

The majority of personality assessment is carried out within the framework of the Global-Rational approach. This is probably because test constructors want a particular test and so do have preconceived ideas before they begin. The work of Cattell, Eysenck, Guilford and Comrey - reviewed later, can be regarded as falling at least partly in this approach.

From the point of view of demonstrated validity there are some newer scales which show much initial promise, such as the Internal-External Control Inventory (I-E) (Rotter, 1966) and the Sensation Seeking Scale (SSS) (Zuckerman, et al, 1964). Fiske and Pearson (1970) conclude that in general, however, the older self-report scales within this approach fail to meet the

requirements of convergent and discriminant validation. For example, results suggest that the Taylor Manifest Anxiety Scale (Taylor, 1953) measures general emotionality and defensiveness as much as it measures "anxiety".

The largest single problem facing investigators in this approach is the definition of the constructs to be measured. Empirical studies have revealed the multi-dimensional nature of most of the constructs, and so there is a need for greater conceptual specificity.

The reliability of the I-E Scale is quite acceptable and results obtained on the standardization sample appear to concur with theory. (Lefcourt, 1966; Rotter, 1966). It appears however as if the Internal control subjects form a more homogeneous group than the External control subjects. This suggests that it would be valuable to further subdivide the External Control factor (Hersch and Schlieber, 1967).

Similarly a factor analysis of the SSS Scale (Zuckerman and Link, 1968) produced four clear factors which could each lead to independent scales.

To gain clarity on the nature of the concepts of authoritarianism and dogmatism, Kehrlinger and Rokeach (1966) factor analysed the Fascism and Dogmatism scales. They found that it was possible to differentiate between the scales and that the concept of authoritarianism could be empirically subdivided into a number of different aspects.

These are just a few examples of studies which have shown the multi-dimensional nature of many of the Global-Rational concepts, and the need for greater conceptual specificity. The Separated-Rational approach attempts to fulfil this need by dividing up the concept measured on a rational basis.

(b) Separated-Rational Approach

In this approach concepts are broken down into their rational components or "facets", before measurement is attempted.

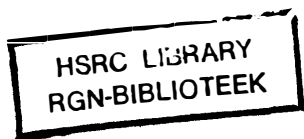
One example of the Separated-Rational approach in action is found in a paper by Kasl, Sampson and French (1964). Independence and dependence were conceptualised as separate constructs rather than polar opposites, and each dimension specified in terms of motivations. The authors proposed four dimensions: approach need dependence and independence; and avoidance need dependence and independence. They devised four separate measures of these needs which they reported to be relatively independent of one another.

One of the difficulties facing this approach is that most of the theoretical work in personality research has been conceptualised at the global level. Thus there are few studies providing evidence that the rational separation of aspects of constructs corresponds with reality, or even that such separation is theoretically advantageous. Until theory itself becomes more specific, it will just have to be assumed on logical grounds alone that definitions are improved by the separate measurement of parts of constructs. This lack of theoretical support is likely to continue to deter workers from using this approach on its own. This approach may be a useful second step however, in the construction of scales begun with the Global-Rational or Empirical approaches.

Though the Separated-Rational approach is still in an early stage of development it appears to show some promise. The approach has the advantage of being less likely to result in measures of vague and diffuse constructs of little demonstrable meaningfulness, than global personality scales. A dimension which might very well benefit from the application of this approach is the large field-dependence-independence dimension of Witkin, et al (1962).

(c) Empirical Approach

In essence the Empirical approach can be outlined as follows: A number of items are drawn up which have no specific relationships to personality factors or to one another. The items are formed into a questionnaire and applied to subjects.



098000

The results are factor analysed and studied carefully to reveal specific personality factors.

This approach has the advantage of relative objectivity and high methodological sophistication. It does have certain disadvantages, however. Since no theory guides the selection of items, really good items are constructed only by chance and may form a relatively small proportion of the total item pool. There is a problem in interpreting the factors which are found, and no guarantee that subsequent researchers will find the same factors.

Possibly the most influential empirically constructed personality scale is the Minnesota Multiphasic Personality Inventory (MMPI) which was drawn up on the basis of its ability to diagnose the pathology from which patients were suffering. As might be expected from an instrument devised using Empirical test construction procedures, there has been much disagreement on the meaning of the areas covered by the MMPI (Adcock, 1967). It appears as if the MMPI can be used as a screening instrument to distinguish normals from abnormal, but cannot distinguish adequately between types of abnormal. Ellis (1967) concludes that the MMPI may be used for group testing, but not for individual diagnosis.

Empirical research has revealed that the majority of measures constructed within the Global-Rational Approach are multidimensional. Adequate Empirical measures would not fall into the danger of being multidimensional, though it will be necessary to ensure that they are psychologically meaningful. Therefore it is likely that the most adequate personality tests will be those constructed using both Rational and Empirical approaches. A Rational approach would be used to establish the construct to be measured, and the Empirical approach to empirically confirm that the test measures that construct.

5.2 Some Practical Guidelines for Test Construction

In this section it is proposed to consider briefly some practical guidelines which have been suggested to aid the constructor of personality questionnaires. Following some sort

of systematic plan, or at least having a clear aim is essential in the development of psychometrically sound devices.

There is a growing realization of the need for research into the efficacy of different ways of constructing questionnaires. Neill and Jackson (1970) undertook a study into the relative effectiveness of 15 different item selection strategies, beginning from a pool of items reflecting Emotional Sensitivity. They give a wealth of useful material on the construction of a personality scale. They found that all purposeful strategies produced better psychometric devices than those using a random strategy, though there was little difference between the purposeful strategies. Neill and Jackson concluded that the most important single consideration in scale construction is the development of a carefully constructed, substantively defined item pool.

Hase and Goldberg (1967) studied the comparative validity of four different test construction strategies using California Psychological Inventory items. When the items were applied to a sample of 200 female students they found that the four primary systematic strategies were significantly more valid than stylistic and random methods. The results were similar to those of Neill and Jackson in that there was little difference found between the primary systematic strategies.

Butt and Fiske (1968) carried out a study which was meant to overcome some of the limitations of the Hase and Goldberg investigation. Hase and Goldberg only considered criterion-related validity, whereas Butt and Fiske compared the different strategies from a wide range view of their psychometric quality. They reached the important conclusion that different strategies of test construction are best suited for different requirements. They classified scales in terms of two dimensions: facet vs. trait and rational vs. factorial. They found: the facet strategy superior to the trait for basic research; the rational-facet approach most promising for theoretical work; and the factorial-facet approach yielded the most relationships with outside variables.

Jackson's (1970) viewpoint on the best strategy to use for test construction is worth noting here. He feels that it is not really meaningful to ask which is the best strategy to employ since the various methods are not mutually exclusive. He argues that with the aid of modern computer technology it is entirely feasible to employ a number of strategies simultaneously or sequentially, to capitalize upon the advantages of each. Certainly consideration must be given to which methods will fit the aims of an instrument best, but there should be no hesitation to make use of any number of approaches to achieve these aims.

Jackson (1970) has provided the most useful outline to guide test construction that has appeared for a long time. He provides a sequential system for scale development that covers all aspects from initial definitions to test validation. He points to four essential principles of personality test construction (pg. 63).

- (a) The overriding importance of psychological theory.
- (b) The necessity for suppressing response-style variance.
- (c) The importance of scale homogeneity, as well as generalizability.
- (d) The importance of fostering convergent and discriminant validity at the very beginning of a programme of test construction.

Further guidelines for test construction are fully documented in Jackson's paper, so there is little point in pursuing them here. The would-be test constructor is well advised to consider Jackson's ideas.

THE CONTRIBUTION OF GUILFORD, EYSENCK, CATTELL AND COMREY

In this section it is proposed to consider the work which Guilford, Eysenck, Cattell and Comrey have done in personality measurement. No review of the area of personality assessment can overlook these workers since they have together revolutionized the field - particularly in their construction of wide-range personality inventories.

6.1 Guilford's Contribution

Guilford has made an important contribution to the methodology as well as to the theory of personality measurement. All the personality scales he has constructed have been devised within the framework of a rigorous and scientific methodology. His work has paved the way for the latest advances in test theory.

Guilford began his personality studies with an investigation into the nature of the dimension Extraversion-Introversion. He came to the conclusion that this dimension was in fact a multidimensional one, made up of the following five factors: Social Introversion; Emotionality; Masculinity; Thinking Introversion; and Rhathymia - i.e. a "happy-go-lucky" attitude. Guilford used the first three factors to form the basis for the development of his Nebraska Personality Inventory (Guilford, 1934). In its original form this inventory proved to be rather ambiguous in meaning and unreliable.

Guilford's next test was the Guilford-Martin Questionnaire (Guilford and Martin, 1943a), followed by the Guilford-Martin Personnel Inventory (Guilford and Martin, 1943b). Both these instruments tended to be rather transparent and so were more useful as experimental than selection devices.

Guilford and Zimmerman (1949) adapted the 10 most important and least interdependent factors from the previous scales to form the Guilford-Zimmerman Temperament Survey. This was a valuable contribution to factor analytical studies, but is not currently itself a particularly significant instrument.

Guilford has constructed a 13-factor structure of temperament model which is fully outlined in his book, "Personality" (1959). In conclusion it can be said that Guilford's primary contribution is likely to be seen in his factor-analytical methods for the study of personality, rather than in the questionnaires he constructed (Michael et al, 1963).

6.2 Eysenck's Contribution

Eysenck's contribution to personality assessment is

also primarily in the factor analytical methods he has demonstrated, though he has constructed two useful instruments - the Maudsley Personality Inventory (MPI), and the Eysenck Personality Inventory (EPI). In addition Eysenck has made a significant contribution to the wider field of personality in his personality theory which links biological, social and personality phenomena together (Eysenck, 1957; Eysenck and Eysenck, 1969).

Eysenck emphasises the value of multivariate analysis procedures in the study of personality, stating that these provide the only way to construct an empirically valid personality theory. Eysenck claims that the essential differences in personality between individuals can be covered by only a few factors. In particular he points to differences in neuroticism and extraversion as covering most of what is meant by personality.

Eysenck's two questionnaires, the MPI and EPI (Eysenck, 1959; Eysenck and Eysenck, 1963) are rather similar to one another. Both measure extraversion and neuroticism, and are supported by an enormous theoretical research background. The reliabilities of the two scales are high (ranging between 0.75 and 0.90), and the two sub-scales of extraversion and neuroticism are reported to be independent of one another. The EPI is a later development of the MPI. Two parallel forms for retesting are available, and nine items from the lie-scale of the MMPI are included to measure test-taking attitudes.

Though a great deal of work has gone into these scales they are of somewhat limited usefulness. The validity of the scales is not easy to determine because of the many psychiatric correlates Eysenck uses in his test construction theory. Unless one is prepared to accept Eysenck's personality theory uncritically, it is difficult to regard these inventories, covering only two factors, as wide-range personality batteries. The scales are more useful when used in conjunction with other tests, or when used simply to give a short and speedy index of the two factors.

Eysenck is using increasingly sophisticated statistical techniques in his personality studies. His book on "Personality

Structure and Measurement" (Eysenck and Eysenck, 1969) is an excellent guide to the application of factor analysis to personality measurement. Unfortunately, it is disturbing that Eysenck's basic theoretical ideas have hardly changed over the last two decades. Psychology is as yet too young a science for any particular theory to expect to remain unchanged for a long period. Klein, et al (1967) point out that there are many conflicting findings in Eysenck's work that he too easily ignores, - particularly in his physiological studies.

6.3 Cattell's Contribution

Cattell has probably made the most significant contribution of any one person to personality measurement using factor analytical procedures. He has published the greatest number of articles in personality assessment and his well-known questionnaire, the 16 Personality Factor Questionnaire is already translated into ten languages. Just about all workers in this field have cause to refer to Cattell's work sooner or later.

In his first article in this area, Cattell (1943a) discusses the conceptual integration of clinical and statistical methods to define traits. He distinguishes between "unique" and "general" traits, and between "surface" and "source" traits. Source traits are indicators of basic behavioural tendencies closely tied to constitutional factors; whereas surface traits reflect less stable tendencies which may be the result of motivational and instrument effects upon personality.

In a second article, Cattell (1943, b) experimentally applies the methods set out in principle in the earlier article. He uses the term "personality sphere" to denote the total personality domain, - a personality domain which he derives in the last resort from the dictionary. The first factor analysis is based upon his verbally defined variables and is subsequently used as the framework for the construction of further tests. A representative list of 171 personality traits was drawn up from the comprehensive list of personality traits of Allport and Albert (1936). Cattell describes how the list was reduced to

151 traits and used to rate 208 adult subjects. Tetrachoric correlations were determined and 67 trait-clusters found which covered these 151 traits. The trait-clusters were reduced in number to 35 and then regarded as the basic variables for the factor analysis.

Twelve factors were extracted by the centroid method of factor analysis (Cattell, 1945a). These factors are fully described by Cattell (1945,b). A number of other investigations followed from this first one: Cattell's confirmation and classification of these "primary personality factors" (1947); a study of the primary factors in the objective test field (1948); and a comparison of the personality factor structure of men and women (1947-8).

Three further factors were extracted and added to the previous 12 factors, together with an intelligence factor to form the basis for the 16 P.F. questionnaire (Cattell, 1950, 1956). There are three forms of the 16 PF available. Forms A and B comprise 187 items each thus providing between 10 and 13 items per factor. The short form C consists of 105 items. The list of personality traits measured by the 16 PF test is presented in Table 1.

Table 1
List of Personality Traits Measured by the 16 PF Test

Trait Designation by Letter	Title of Trait
A	Affectothymia v. sizothymia (warm, easy going v. critical, reserved).
B	General intelligence v. mental defect
C	Ego strength v. dissatisfied emotionality
E	Dominance or ascendance v. submission
F	Surgency v. desurgency ('enthusiasm' v. melancholy')
G	Superego strength v. lack of internal standards
H	Parmia v. threctia ('adventurous' v. 'timid')
I	Premisia v. harria (protected emotional sensitivity v. tough maturity)

Trait Designation by Letter	Title of Trait
L	Protension v. alaxia ('suspecting' v. 'accepting')
M	Autia (autistic temperament) v. practical concernedness
N	Sophistication v. rough simplicity (or 'shrewdness' v. 'naivete')
O	Guilt-proneness v. confident ('insecure' v. 'confident')
Q ₁	Radicalism v. conservatism
Q ₂	Self-sufficiency v. lack of resolution
Q ₃	Strong self-sentiment v. weak self-sentiment
Q ₄	High ergic tension v. low ergic tension (tense v. relaxed)

It can be noticed that Cattell makes use of highly esoteric labels for his traits in an effort to reduce the number of non-scientific associations attached to them. This does have the disadvantage however of making understanding difficult and meaning obscure.

In a further factor analysis Cattell (1956) extracted four second-order factors from the questionnaire: Anxiety; extraversion v. introversion; cyclothyme v. schizothyme constitution; and success v. frustration. (It is interesting to note that the first two factors are the same as those that Eysenck reports). Although Cattell claims that each of the 16 factors concurs with a source trait in a behaviour situation, this has not been demonstrated in research. A report by Becker (1961) casts some doubt upon the independence of the 16 factors. Becker found only eight independent factors.

The questionnaire has been validated for clinical, developmental and industrial use. (Cattell, Blewett and Belhof, 1955; Cattell, Day and Meeland, 1952; and Wrights, 1955). More information is needed on its construct validity. The 16 PF has been compared with other questionnaires such as the MMPI (Gocka and Marks, 1961; La Forge, 1962); the Vocational

Preference Inventory (Holland, 1960); and the Guilford-Martin Personality Inventory (Becker, 1961). Harsh (1953) points out that there is too small a relationship between the 16 PF and these other instruments. He particularly criticises the 16 PF for ambiguity of its items and the fact that its factors overlap.

The 16 PF is an important instrument for it has set the pace for large-scale personality tests. Wittenborn (1953) points to the value of Cattell's selection of items on the basis of their factor-analytic grouping, for this attempts to ensure homogeneous behavioural areas that are likely to have good validity estimates. There is however, a number of inadequacies in the 16 PF which should be rectified in any tests which are constructed along the same lines. Wittenborn cannot agree with Cattell that his questionnaire covers the whole personality sphere because its construction is dependent upon the behaviour of the norm sample which can never be completely representative.

In 1965 Lorr suggested that the norms of the 16 PF were inadequate because such characteristics of the sample as occupation, social status and background were not taken into account in their calculation. In the latest handbook on the 16 PF however, Cattell, Eber and Tatsuoaka (1970) have attempted to provide full information on the test. This new handbook is an impressive piece of work providing norms and test profiles for the use of the 16 PF in clinical, educational, industrial and research psychology. Cattell has made every effort to make the 16 PF the most comprehensive and useful instrument available. Further research using the questionnaire must tell how successful he has been in his aim.

Cattell is currently more interested in the use of objective tests for the measurement of personality, because these are less easily faked by the subject. Cattell now measures some 20 factors by objective tests as part of the Objective-Analytic Personality Test Battery (Cattell, Baggeley, et al, 1955). About 30 independent multivariate studies were carried out to investigate the personality structure as assessed by

objective tests. This structure is reported in a monograph by Hundleby, et al (1964). An important contribution is the compendium of objective tests (Cattell and Warburton, 1967) yielding over 600 performance variables. According to Klein, et al (1967) this work on objective tests may turn out to be Cattell's greatest contribution to the measurement of personality. Cattell is frequently criticised out of ignorance by critics who simply are not aware of the level of sophistication of his research. It is likely that many reviewers simply do not understand the meaning of Cattell's factors or procedures because of the vast number of variables he considers.

6.4

Comrey's Contribution

Comrey's most significant research has only been carried out relatively recently, but has already made a very significant contribution to personality assessment. He is particularly concerned about clearing up some of the problems facing personality test constructors. He points out that there is far too little agreement about which personality variables should be studied. There is not very much agreement even when the same variables are studied.

The factor structure of existing scales is often very complex, certainly not corresponding to clearly defined "pure" factor scales. Not only do existing scales not measure pure factors, but they correlate with other scales from which they should have been very different according to their definitions (Comrey and Soufi, 1960). Comrey (1961) suggests that one of the most important sources of disagreement between different personality studies is the use of items as the basic variables in factor-analytic personality studies. Items are usually unreliable and so give rise to unstable correlation matrices and tests. He suggests rather that the basic unit variable should be the "Factored Homogeneous Item Dimension" - (FHID). A FHID is a small cluster of about four or more items that all load highly in a factor analysis to define a "pure" factor. Total scores over such dimensions will give more reliable measures and hence a more stable picture of the factor structure of personality in the domain of inventories.

The procedure which Comrey uses is more or less the following: firstly the personality dimensions must be clearly defined on the basis of theory, previous research or intuition. Secondly items are drawn up to give a homogeneous, internally consistent cluster of items for each dimension. These items are applied to subjects and the results factor analysed. A clear factor should emerge for each dimension if the dimension was adequately defined and is independent, and if the items were drawn up well. The items with the highest loadings on the factor are chosen to represent the dimension and form a FHID. Comrey (1961) makes a very important point which must not be forgotten in evaluating his later work. He says that a factor produced by analysing collections of FHIDs does not necessarily represent a phenomenon of deep underlying significance for the understanding of personality dynamics. All that a FHID represents is a well-defined variable which may be useful in the description of personality. In other words, Comrey's approach does not free the investigator from the need to search for meaningful and relevant personality dimensions and the FHIDs to represent them.

Comrey began his research by a large number of factor-analytic studies of the MMPI in 1957 and 1958. Of particular importance to this paper however, is his research which led to the development of his own personality scales. Comrey (1961) chose over 40 personality dimensions from his factor analytical studies of the MMPI, from the factors of other researchers such as Guilford and Cattell, and from other dimensions that he thought might throw some light on the expected factors. He wrote four to six items per dimension and applied them to 300 subjects. The results were factor analysed in four separate studies (Comrey, 1961, 1962; Comrey and Soufi, 1960, 1961), to produce FHIDs.

In an attempt to gain results that were reliable Comrey (1961) took the following precautions: He improved item reliability by changing the item-format from two-choice to nine-choice responses; he was careful not to include two dimensions in the analysis which were merely alternate forms of

one another; and he extracted a large number of factors. The FHIDs that he obtained were combined with factors from Guilford and Cattell and the analyses carried out again. This process was repeated four times, (Comrey, 1964, 1965) until four stable FHIDs were found: Shyness, Neuroticism, Compulsion and Hostility. In a later study Comrey and Jamison (1966) added two other FHIDs, namely Dependence and Empathy. Table 2 shows the eight scales that are currently to be found in the published form (Comrey, 1970) of the Comrey Personality Scales.

Table 2

List of Personality Dimensions Measured by the Comrey Personality Scales

Dimension Designation by Letter	Title of Dimension
T	Trust v. defensiveness
O	Orderliness v. lack of compulsion
C	Social conformity v. rebelliousness
A	Activity v. lack of energy
S	Emotional stability v. neuroticism
E	Extraversion v. introversion
M	Masculinity v. femininity
P	Empathy v. egocentrism

Comrey suggests that researchers should not be encouraged to construct new scales so much as to seek for new dimensions. Comrey's own work is of high standing but there is still much to be done. There is as yet little theoretical foundation to the FHIDs of his scales and clearly a need for more items for each dimension. The relationship between the different dimensions on various samples should be investigated, and attempts made to find external criteria for the meaning of the dimensions.

6.5

The Relationship Between the Personality Dimensions of Guilford, Eysenck, Cattell and Comrey

Having reviewed the contribution of the "big four" names in the factor analytical approach to personality assessment

the question arises, how do the personality dimensions of these investigators relate to one another? Before considering some specific studies a few general observations are in order.

When a number of personality inventories are factor analysed together it is often found that the factors extracted do not correspond with the factors or dimensions that the individual tests were supposed to measure. Borgatta (1968) points to the disconcerting fact that the major concepts arising out of "second order" factor analyses, or analyses of scores (rather than items) of wide-range personality inventories, tend to appear rather similar to the concepts measured by older and simpler tests. The concepts which emerge are frequently such dimensions as extraversion-introversion, ascendance-submission, emotional stability, and social and emotional adjustment. Studies frequently reveal an apparent lack of coincidence in tests of similar constructs. This may reflect the tendency there has been to move away from the older social adjustment inventories with only two or three scores. The development of tests adding more subscores may not correspond to the empirical experience of sampling of items but may indicate the pursuit of particular avenues of cluster construction.

Sells, Demaree and Will (1970) carried out a factorial study of items from Guilford's and Cattell's questionnaires, selected so as to represent 15 Guilford and 17 Cattell factors. They applied the items to a sample of over 2,000 male airforce recruits. They found that analysis at the item level was highly destructive to the factors previously assembled. They concluded that Guilford and Cattell's so called "source" traits have an overlapping item composition.

Eysenck and Eysenck (1969) report a combined analysis of personality test items taken from Guilford, Cattell and Eysenck inventories. Their main interest was in second-order factors, or what they call "super" factors. They found results which tend to confirm what has been stated above. They found that the primary factors of Guilford and Cattell were not replicable across sex, nor were they replicable from one

investigator to another. In the combined analysis they were able to extract four second-order factors. The two most powerful second-order factors corresponded with Extraversion and Neuroticism. The other two factors were unstable and appeared to be made up of Guilford and of Cattell items respectively.

Mitchell (1963) undertook a comparison of the first and second-order dimensions of the 16 PF and California Psychological Inventory (CPI). He also found little congruence between the first-order domains of the two instruments. The second-order domains of the two instruments were more co-extensive than the first, with both instruments playing a major role in defining five factors: General adjustment v. neuroticism, introversion v. extraversion, intellectual resourcefulness, emotional sensitivity v. tough-mindedness, and super-ego strength. Mitchell concluded that it appeared as if the second-order domains of personality tests may define a similar picture.

A study which included the Comrey Personality Scales appears to substantiate the conclusion of Mitchell. Comrey and Duffy (1963) studied the relationship between the EPI, 16 PF and the Comrey Personality Inventory. They found that the EPI Extraversion (E) and Neuroticism (N) factors appeared to match with the Comrey E and N factors. In addition the Cattell second-order anxiety factor matched the Comrey Neuroticism factor. The Cattell primary factors however, overlapped but did not match with the Comrey factors.

This research has shown that personality measurement is still by no means a straightforward matter. It appears as if the questionnaires covering a few, carefully factor analysed dimensions are more likely to measure replicable characteristics than the more rationally constructed tests covering a large number of dimensions.

7.0 METHODOLOGICAL CONCERNS

7.1 Introduction

Having considered some of the questions that are

being asked in personality assessment and some of the major tests that are available it would be valuable at this point to turn to look at the methodology being used.

Though very real progress is being made it cannot be concluded that all the research being done and the tests being constructed are the best possible. Even the most cursory glance at the literature reveals that by far the majority of studies are still restricted to university student samples. This seriously limits the generality of the results found. There are also still too many studies using the single-variable correlation approach. This method may be able to relate an almost infinite number of dimensions, but gives little meaningful insight into interrelationships. Multivariate analysis is a step in the right direction, though it should not be assumed that increasing the number of personality traits measured in one sitting will necessarily increase the number of meaningful relationships discovered. Marlowe and Gergen (1969) point to the need to take account of non-linear relationships in order to cover simultaneously interacting personality processes.

7.2 The Vanishing Variance Component

Possibly one of the most important questions which has been asked about the methodology of personality assessment is concerning the relationship of personality variance and method variance. In essence the question relates to the validity of measurement, - whether test scores reflect personality or method differences.

Campbell and Fiske (1959) have suggested that for most personality constructs, it is possible to design experiments employing a multi-trait multi-method matrix for teasing out the relative contributions of content and method. Given two or more traits and two or more methods for assessing the traits, the resulting scores can be intercorrelated and the results arranged in a rectangular matrix for inspection. More or less the same information can be received by the judicious use of factor analysis, comparing similar measures and similar traits. It is important, however, that a multimethod factor analysis rather

than the traditional principal components factor analysis be used. For example, in a study by Kusyszyn and Jackson (1968) multimethod factor analysis revealed important method variance which was not revealed in a standard principal components factor analysis. Boruch and Wolins (1970) provide a somewhat more recent outline of a multi-method factor analytical procedure for uncovering method variance.

It can be seen that there are a number of procedures available for uncovering method variance, but what accounts for this variance and how important is it? Method variance is the result of characteristics of the personality tests used that elicit responses to them which do not follow from personality differences. Fiske and Pearson (1970) suggest that method variance might be better understood if termed, "methods at variance" with one another. They emphasize the need for using a number of different measures of a construct to establish the independent existence of the construct.

It is not easy to answer the question relating to the importance of method variance. As shall be seen there are widely differing opinions on this score. Perhaps it can be said in general though, that researchers are tending to emphasize the positive aspects of gaining demonstrated validity more than the negative attempt to reduce particular aspects of method variance. Certain workers such as Borgatta (1968) imply that the careful utilization of factor analytical techniques of test construction more or less precludes excessive method variance anyway.

The most important sources of method variance in the questionnaire method are the result of response styles. The two response styles usually considered are social desirability set and acquiescent set. Social desirability set refers to the tendency by subjects to respond to items in terms of their perceived social desirability. Thus a subject may try to look well-adjusted, extraverted, happy, self-confident, and so on. Acquiescent set refers to the tendency by subjects to agree with items ("yea-saying") irrespective of item content. Though social desirability and acquiescent sets are perhaps the most

important response styles, they are of course not the only ones. Subjects may adopt tendencies for extreme responding or random responding to items which do not reflect the content of the items.

Possibly the most important worker in the field of social desirability response sets is Edwards. He has devised a scale to measure social desirability (1965, c) and has written two important books (1957, 1970) which give information on the topic. Edwards' (1957) method of judging the social desirability of items is to give each item to a group of judges who rate it in terms of its social desirability on a Likert-type scale. A mean social desirability scale value (S.D.S.V.) is calculated by the judges for each item. A whole sample of items having similar S.C.S.V.s can be put together in this way. Edwards has done a lot of research on social desirability as found in the MMPI. He has found quite a substantial amount of evidence for the influence of social desirability in the MMPI. (Edwards, 1957, 1964,c, 1965,b, 1965,c).

There is by no means complete agreement on the nature of Acquiescent set. (Block, 1965; Messick, 1966; Rorer, 1965). According to Christie and Lindauer (1963) the origin of acquiescent set is to be found in vague and ambiguous items. Particular items have a particular meaning for an individual so that an item which is understandable for one person is obscure to another.

One of the important workers in the field of acquiescent set and other response styles is Messick (1966). He takes the point of view that response styles should not just be written off as artifacts of questionnaire measurement, but should be seen as important personality variables in their own right. Damarin and Messick (1965) have written a very interesting paper on response styles as personality variables. They argue that response styles may depend on intelligence or upon attitudes toward the self, and that such traits may play important roles in the psychic economy. They report evidence suggesting that correlations among measures of any one response style are

determined by more than one basic personality dimension. Thus social desirability encompasses both a social desirability and a lie factor, while acquiescence may reflect "yea-saying" as well as an intellectual confusion (or ability) factor.

As was pointed out earlier there is a growing feeling that response style does not play as large a part as was originally thought by workers such as Edwards (1957). A certain amount of research supports this position.

Dicken (1963) investigated the difference made to the validity of the California Psychological Inventory (CPI) on a student sample by the suppression of such variables as social desirability. He found that such suppression made very little difference and concluded that response set variables do not account for enough criterion-irrelevant predictor variance to warrant their use in correcting the scores of subjects. He suggests that it is not surprising that the majority of normal subjects endorse what is considered to be the acceptable response by the majority of normal subjects. He appears to take a position similar to that of Damarin and Messick in arguing that social desirability variables need not, and perhaps should not be removed from personality scales. The important task is the positive one of developing predictors that account for more criterion variance than is usually the case.

In an article entitled, "The Great Response-Style Myth", Rorer (1965) concludes that the literature contains no unequivocal evidence of the importance of acquiescence in personality scales. He questions the contention that present psychological inventories may be improved by correcting for stylistic variables. He and Goldberg (1965,b) carried out an experimental study of the role of acquiescence in the MMPI using reversed items and the so-called "Acquiescence" scales of the MMPI. They found that only trivial proportions of the total variance were found to be attributable to response variance.

Jackson and Lay (1968) also conclude that personality items are far less affected by stylistic variables than has been

thought traditionally. They found that item content dimensions were orthogonal to acquiescence and desirability factors. They do advise, however, that test construction techniques to maximize content saturation and minimize response biases should continue to be used.

There is a number of methods that have been suggested for the measurement and control of response styles. Firstly some of the measures and secondly some of the methods of controlling response style will be considered.

Edwards (1957) has devised a Social Desirability scale which is used frequently in research but has been criticised (Fiske and Pearson, 1970). It is suggested that Edwards' scale measures both good adjustment and dissimulation, and appears to have a limited generality across different methods of measurement. There is also the Marlowe-Crowne Social Desirability Scale (Crowne and Marlowe, 1960) which was constructed to ensure a minimal correlation with clinical criteria. (This correlation with clinical criteria being an important criticism of Edwards' scale) The Marlowe-Crowne scale is considered to be a measure of the approval motive and is linked to social learning theory. Although the Marlowe-Crowne scale is regarded as a better one than the Edwards scale, it has also been criticised - particularly on the grounds of its not very adequate psychometric characteristics (Goldfried, 1964).

There are a few specific scales of acquiescence that have been devised, such as: Bass' (1956) "Social Acquiescence Scale", Couch and Keniston's (1960) Overall Agreement Scale; and Christie, Havel and Sudenberg's (1958) "discrepancy matrix". In a large number of studies, however, acquiescent set is measured simply by adding up the number of "true" responses. There are so many and varied methods of measurement that Fiske and Pearson (1970) describe the area as chaotic. Measures of social desirability and acquiescence typically are uncorrelated (Cruse, 1966; Feder, 1967) and usually emerge as separate factors in factor analysis.

The main method that has been used for controlling social desirability is the matching of items in terms of social desirability by the adoption of a forced-choice response format. This method of matching items requires agreement on what constitutes the socially desirable. Scott (1963), however, strongly argues (and demonstrates with data) that there are wide differences in what constitutes the desirable. N. Wiggins (1966) supports Scott's position, and found at least six viewpoints in social desirability judgements of MMPI items. She concluded that the forced-choice format is inadequate as a control for social desirability at the individual level.

Acquiescence has generally been controlled by eliminating vague and ambiguous items and by reversing the keying of some items in a scale. Item reversal has been severely criticised on methodological grounds by researchers such as Liberty (1965). It has been suggested that an item that is reversed is not psychologically equivalent to its non-reversed form. The research on response styles has led to the attempt to measure a "repression v. sensitization" dimension (Byrne, 1964; Couch and Keniston, 1960). "Sensitizers" are "yea-sayers" but also tend to show less evidence of being affected by social desirability sets than "repressors" or "nay-sayers".

Because of the lack of agreement about the importance of response styles in questionnaire measurement it would seem to be advisable to attempt to minimize their effects, but not to expend too much energy in their control and measurement. It would be advisable to include a scale of social desirability in a personality inventory and to avoid test samples (e.g. psychology students) that are likely to be "test-wise". Though it is true that research such as that of Dicken (1963) has found social desirability to be less important than was thought previously, it can certainly not be ignored in personality measurement. Dicken's research was based on student samples that may have had little cause to fake results. Research based on job-applicant samples might have produced very different results. The effect of acquiescence may not be

so marked in a carefully devised personality scale, but the effect of social desirability cannot be discounted, at least at the present stage of research. Marlowe and Gergen (1969) also recommend the routine use of an intelligence measure, for they consider this to be a major source of method variance in the area of personality. Borgatta (1968) points out that it may turn out to be more important to work on the validity of tests from a positive viewpoint, than to concentrate on the elimination of response sets from a negative viewpoint. This sets the scene for the next section.

8.0

THE VALIDITY OF PERSONALITY ASSESSMENT

The validity of a personality test is universally acknowledged as being very important, but is not always given pride of place in actual test construction. Campbell and Fiske (1959) brought out an important point in an article on "convergent and discriminant" validity: a test should not only relate to conceptually similar measures, but should not correlate highly with theoretically unrelated constructs.

Yet a consideration of the performance of personality tests in practice (Sarason and Smith, 1971) provides rather disconcerting findings. Frequently a bewildering array of inconsistent and unexpected findings suggest that measures are not measuring what they were originally assumed to be. Some of the inconsistent patterns of results obtained with personality measures may be outgrowths of the questionable implicit assumption that subjects who occupy a similar position on a given personality dimension, are (at least as a group) also alike on other variables which may influence the dependent variable behaviour. Personality and behaviour in test situations are too complex to be covered by such an assumption. Validity studies must concomitantly take other personality variables into account, and cannot be delayed until after the test is constructed. Jackson (1970) makes a very important point in referring to the need to consider validity in the test development phase. He says (pg. 85):

"... test validation ... implies a program of test construction wherein successive attempts are made to approach a specified variety of optimal properties, rather than leaving the outcome entirely to capricious nature."

Jackson's paper is an important one for, although there has been much written on validity of completed tests (e.g. Anastasi, 1954), there has been very little on the consideration of validity while constructing a test.

Jackson (1970) chooses to consider the validity of personality tests in terms of two components: a structural and an external one. It is the structural component that should be considered in particular while constructing a test.

The structural component of validity requires that a test conform to the hypothesized structural model of content-homogeneity. The test should be homogenous and "content saturated", i.e. not attempt to measure a number of different things or be affected by response styles. A test that is content heterogeneous will have a low discriminant validity - probably the major source of error in personality measures. Such a test will probably also have a low reliability since it measures a number of characteristics.

It is of course not enough that a test be structurally valid. If it is to be at all useful it must be valid in terms of correlations with external criteria. It should be noted that the presence of one of the two components of validity does not necessarily imply the presence of the other. A test may be content-homogeneous but have few practical applications; while another test may have numbers of relevant external correlations but be content-heterogeneous. Ideally the two components do go together, but in practice this may not occur unless careful test construction has specifically aimed at both components of validity. Validation of a test against external criteria should not be carried to an extreme, lest a test become a good measure of the criterion but a poor one of the trait concerned. This implies the need for a continued search for good criteria in the validation of personality tests.

A point that was made earlier was that validity studies must take other personality variables than the one being measured into account. If other variables influencing criterion behaviours are not taken into account validity studies have little chance of success, unless the personality variable being considered is an extremely prevalent one. It follows from this point that multitrait rather than unitrait measures should be developed to cover these constellations of personality traits. As Sarason and Smith (1971) put it, ignoring individual differences because of the problems involved in measuring them seems tantamount to throwing the baby out with the bath water.

It is worth remembering that the validation of personality measures is a complex undertaking. Even the structural validation of tests has to take individual differences into account, for as Scott (1963) showed, such characteristics as social desirability are qualities of persons and not merely of items. Fiske (1968) undertook a study of the interaction between persons and items in personality measurement. In considering the response process he found that item indices vary from person to person. Ghiselli (1963) also found that persons have differences in the errors that they show. He concluded that it is not sufficient only to consider test variables. It is also necessary to consider the use of moderator variables to predict individual differences in errors. Ghiselli proceeds to outline some methods of including such moderator variables.

9.0

CONCLUSION

It seems in order to conclude this survey of personality measurement on an optimistic note. Holtzman (1964) and Jackson (1970) point to a number of problems and yet also to encouraging aspects of personality assessment. In particular, the advent of more sophisticated methodology is a promise of better things. Fiske and Pearson (1970) note that progress is being made - even if the goal of technical adequacy

is still a long way off. The multiple determination of test responses is being increasingly recognized.

Personality measurement must be seen as part of the whole body of psychology. New methods of measurement may be found in some of the related fields of psychology. Two valuable trends in the measurement of personality are noted by Sarason and Smith (1971): the use of unobtrusive and naturalistic measures, and the assessment of behaviour in ongoing social situations. New measures of personality may become more and more important if the outcry against traditional personality testing mentioned by Holtzman (1964) continues. Unobtrusive measures may be particularly useful where pencil and paper tests are feared.

Because of the expense surrounding these newer measures of personality and the difficulty of measuring groups of subjects with them, the questionnaire method must remain the most important one. It is still going to be necessary for personality theorists to construct personality questionnaires, for the existing instruments are by no means the best possible available. This last point applies particularly to the questionnaires available for measuring personality in South Africa. There simply is no multifactor personality questionnaire constructed in South Africa particularly for local conditions.

Perhaps the most important conclusion that can be drawn on the basis of this review of research is that much can be done to ensure valid and reliable personality measurement using questionnaires. This requires methodological rigour and sensitivity to theoretical issues from the earliest stages of constructing a questionnaire. It is too late to attempt to change an already constructed instrument into one that will be both methodologically sound and practically useful.

10.

BIBLIOGRAPHY

ADCOCK, C.J. (1965) Rev. 143 In Buros, O.K., "The Sixth Mental Measurements Yearbook". New Jersey: The Gryphon Press.

- ALLPORT, G.W. and ODBERT, H.S. (1936) Trait names: A Psycholexical Study. Princeton: Psychological Review Co.
- ANASTASI, Anne. (1954) Psychological Testing. New York: Macmillan Co.
- BARRCWS, G.A. and ZUCKERMAN, M. (1960) Construct validity of three Masculinity-Femininity Tests. Journal of Consulting Psychology, 24, 441-5.
- BASS, B.M. (1956) Development and Validation of a scale for measuring Social Acquiescence. Journal of Abnormal and Social Psychology, 53, 296-99.
- BECKER, W.C. (1961) A comparison of the factor structure and other properties of the 16 PF and the Guilford-Martin Personality Inventories. Educational and Psychological Measurement, 21, 393-404.
- BECKWITH, J., IVERSON, M.A. and REUDER, Mary E. (1965). Test anxiety, task relevance of group experience and change in level of aspiration. Journal of Personality and Social Psychology, 1, 579-88.
- BENTON, A.L. (1945) The MMPI in clinical practice. Journal of Nervous and Mental Diseases, 100, 416-20.
- BERKOWITZ, N.H., and WALKIN, G.H. (1964) A forced-choice form of the F-scale - free of Acquiescent Response set. Sociometry, 27, 54-65.
- BLOCK, J. (1965) The Challenge of Response Sets: Unconfounding meaning, acquiescence and social desirability in the MMPI. New York: Appleton Century Crofts.
- BOE, E.E. (1964a) Acquiescence set in the Edwards Social Desirability Scale: an evaluation of the Adams and Kirby paper. Journal of Consulting Psychology, 28, 472-473.
- BOE, E.E. (1964b) Social desirability considerations in separate acquiescence set and content scores. Psychological Reports, 15, 699-702.

- BOE, E.E., GOCKA, E.F. and KOGAN, W.S. (1967) Factor analysis of individual social desirability scale values: Second-order analysis. Multivariate Behavioral Research, 2, 239-40.
- BOE, E.E. and KOGAN, W.S. (1963) Social desirability response set in the individual. Journal of Consulting Psychology, 27, 369.
- _____ (1964a) An analysis of various methods of deriving the social desirability score. Psychological Reports, 14, 23-29.
- _____ (1964b) Effect of social desirability instructions on several MMPI measures of social desirability. Journal of Consulting Psychology, 28, 245-51.
- _____ (1964c) Social Desirability in Individual Performances on Thirteen MMPI Scales. Address presented at the American Psychological Association Meeting in Los Angeles.
- BORGATTA, E.E. (1968) Traits and Persons, in Borgatta, E.F. and Lambert, W.W. (eds), Handbook of Personality Theory and Research. Chicago: Rand McNally & Co.
- BORUCH, R.F. and WOLINS, L. (1970) A procedure for estimation of trait, method and error variance attributable to a measure. Educational and Psychological Measurement, 30, 547-74.
- BRAUN, J.R. (1965) Effects of specific instructions to fake on Gordon Personal Profile scores. Psychological Reports, 17, 847-50.
- _____ (1966) Effects of "typical neurotic" and "after therapy" sets on Personal Orientation Inventory scores. Psychological Reports, 19, 1282.
- _____ and GOMEZ, B.J. (1966) Effects of faking instructions on the Eysenck Personality Inventory, Psychological Reports, 19, 398-90.
- BURTGN, R.V. (1963) Generality of honesty reconsidered. Psychological Review, 70, 481-99.

- BUTCHER, J.N. and TELLEGEN, A. (1966) Objections to MMPI items. Journal of Consulting Psychology, 30, 527-34.
- BUTT, D.S. and FISKE, D.W. (1968) Comparison of strategies in developing scales for dominance. Psychological Bulletin, 70, 505-19.
- BYRNE, D. (1961) The Repression-Sensitization Scale: rationale, reliability and validity. Journal of Personality, 29, 334-49.
- - and BOUNDS, C. (1964) The reversal of F-scale items. Psychological Reports, 14, 216.
- CAMPBELL, D.T. and FISKE, D.W. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin. 56, 81-105.
- CAMPBELL, D.T. and SORENSON, W.W. (1963) Response set on interest inventory triads. Educational and Psychological Measurement, 23, 145-52.
- CARVER, R.P. (1968) Brief report: On the danger involved in the use of tests which measure factors. Multivariate Behavioral Research, 3, 509-12.
- CATTELL, R.B. (1943a) The description of personality: I. The foundations of trait measurement. Psychological Review, 50, 559-94.
- - (1943b) The description of personality: II. Basic traits resolved into clusters. Journal of Abnormal and Social Psychology, 38, 476-506.
- (1945a) The description of personality: Principles and findings in a factor analysis. American Journal of Psychology, 58, 69-90.
- (1945b) Interpretation of the twelve primary personality factors. Character and Personality, 13 55-91.
- (1947) Confirmation and classification of the primary personality factors. Psychometrika, 12, 197-220.

- _____ (1947-8) The primary personality factors in women compared with those in men. British Journal of Psychological Statistics, Section 1, 114-130.
- _____ (1948) Primary personality factors in the realm of objective tests. Journal of Personality, 16, 459-87.
- _____ (1950) The main personality factors in questionnaire self estimate material. Journal of Social Psychology, 31, 3-38.
- _____ (1956a) Validation and intensification of the Sixteen Personality Factor Questionnaire. Journal of Clinical Psychology, 12, 205-14.
- _____ (1956b) A shortened "basic English" version (Form C) of the 16 PF Questionnaire. Journal of Social Psychology, 44, 257-78.
- _____ (1956c) Second-order personality factors in the questionnaire realm. Journal of Consulting Psychology, 20, 411-18.
- _____ (1957) Personality and Motivation Structure and Measurement. Yonkers-on-Hudson: World Book Co.
- _____ (1964) Objective Personality tests: A reply to Dr. Eysenck. Occupational Psychology, 38, 69-86.
- _____ BAGGELEY, A.R., CHECOR, L., COGAN, E.A., FLENT, D., GRUEN, W., HUSIK, C., MEELAND, T., SAUNDERS, D.R., and SCHIFF, H. (1955) Objective-Analytic Personality Test Batteries. Urbana: IPAT.
- CATTELL, R.B., BLEWETT, D.B. and BELOFF, J.R. (1955) The inheritance of personality. A multiple-variance analysis determination of approximate nature-nurture ratios for primary personality factors in Q-data. American Journal of Human Genetics, 7, 122-146.
- CATTELL, R.B., COAN, R.W., and BELOFF, A.H. (1953) Junior-Senior High School Personality Questionnaire. Bobbs-Merrill Co.

- CATTELL, R.B. and COAN, R.W. (1958) Personality dimensions in the questionnaire responses of six and seven year olds. British Journal of Educational Psychology, 28, 232-42.
- (1963) Early School Personality Questionnaire. Urbana: IPAT.
- CATTELL, R.B., DAY, M., and MEELAND, T. (1956) Occupational profiles for the 16 PF Questionnaire. Occupational Psychology, 30, 10-19.
- CATTELL, R.B. and EBER, H.W. (1949) The Sixteen Personality Factor Questionnaire. Urbana: IPAT.
- CATTELL, R.B., EBER, H.W. and TATSUOKA, M.M. (1970) Handbook for the Sixteen Personality Factor Questionnaire. Illinois, IPAT.
- CATTELL, R.B. and GRUEN, W. (1954) Primary personality factors in the questionnaire medium for children 11 to 14 years old. Educational and Psychological Measurement, 14, 50-76.
- CATTELL, R.B., KNAPP, R.R. and SCHIER, I.H. (1961) Second-order personality factor structure in the objective test realm. Journal of Consulting Psychology, 20, 345-54.
- CATTELL, R.B. and LAYTON, W.L. (1963) Children's Personality Questionnaire, Urbana, IPAT.
- CATTELL, R.B. and WARBURTON, T.W. (1964) Principles in Personality Measurement and a Compendium of Objective Tests. Urbana: University of Illinois Press.
- (1967) Objective Personality and Motivation Tests: A Theoretical Introduction and Practical Compendium. Urbana: University of Illinois Press.
- CHRISTIE, R., HAVEL, J. and SUDENBERG, B. (1958) Is the F-scale reversible? Journal of Abnormal and Social Psychology, 56, 43-59.

CHRISTIE, R. and LINDAUER, F. (1963) Personality Structure. Annual Review of Psychology, 14, 201-30.

COMREY, A.L. (1957a) A factor analysis of items on the MMPI Hypochondriasis scale. Educational and Psychological Measurement, 17, 568-77.

————— (1957b) A factor analysis of items on the MMPI Depression scale. Educational and Psychological Measurement, 17, 578-85.

————— (1957c) A factor analysis of items on the MMPI Hysteria Scale. Educational and Psychological Measurement, 17, 586-92.

————— (1958a) A factor analysis of items on the MMPI Psychopathic scale. Educational and Psychological Measurement, 18, 91-8.

————— (1958b) A factor analysis of items on the MMPI Paranoia scale. Educational and Psychological Measurement, 18, 99-107.

————— (1958c) A factor analysis of items on the MMPI Psychasthenia scale. Educational and Psychological Measurement, 18, 293-300.

————— (1958d) A factor analysis of items on the MMPI Hypomania scale. Educational and Psychological Measurement, 18, 313-23.

————— (1958e) A factor analysis of items on the F-scale of the MMPI. Educational and Psychological Measurement, 18, 621-32.

————— (1958f) A factor analysis of items on the K-scale of the MMPI. Educational and Psychological Measurement, 18, 633-39.

————— (1961) Factored Homogeneous Item Dimensions in personality research. Educational and Psychological Measurement, 21, 417-31.

————— (1962) A study of thirty-five personality dimensions. Educational and Psychological Measurement, 22, 543-52.

- (1964) Personality factors of Compulsion, Dependence, Hostility and Neuroticism. Educational and Psychological Measurement, 24, 75-84.
- (1965) Scales for measuring Compulsion, Hostility, Neuroticism and Shyness. Psychological Reports, 16, 697-700.
- (1970) The Comrey Personality Scales. San Diego: Educational and Industrial Testing Service.
- and DUFFY, K.E. (1963) Cattell and Eysenck factor scores related to Comrey personality factors. Multivariate Behavioral Research, 3, 379-392.
- COMREY, A.L. and JAMISON, K. (1966) Verification of six personality factors. Educational and Psychological Measurement, 26, 945-53.
- COMREY, A.L. and MARGRAFF, W. (1958) A factor analysis of items on the MMPI Schizophrenia scale. Educational and Psychological Measurement, 18, 301-11.
- COMREY, A.L. MESCHIERI, L., MISITI, R. and NENCINI, R. (1965) A comparison of personality factor structure in American and Italian subjects. Journal of Personality and Social Psychology, 1, 257-61.
- COMREY, A.L. and SOUFI, A. (1960) Further investigation of some factors found in MMPI items. Educational and Psychological Measurement, 20, 777-86.
- (1961) Attempted verification of certain personality factors. Educational and Psychological Measurement, 21, 113-27.
- COUCH, A. and KENISTON, K. (1960) Yea-sayers and Nay-sayers: agreeing response set as a personality variable. Journal of Abnormal and Social Psychology, 60, 151-74.
- CRITES, J.O., BECHTOLDT, H.P., GOODSTEIN, L.D., and HEILBRUN, A.B. (1961) A factor analysis of the California Psychological Inventory. Journal of Applied Psychology, 45, 408-14.

- CRONBACH, L.J. (1946) Response sets and test validity. Educational and Psychological Measurement, 6, 475-94.
- _____ (1950) Further evidence on response sets and test design. Educational and Psychological Measurement, 10, 3-31.
- _____ (1959) Rev. no. 37, In Buros, O.K.: The Fifth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- CROWNE, D.P. and MARLOWE, D. (1960) A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 24, 349-54.
- CROWNE, D.P. and MARLOWE, D. (1964) The Approval Motive: Studies in Evaluative Dependence. New York: Wiley.
- CRUSE, D.B. (1965) Social desirability scale values of personal concepts. Journal of Applied Psychology, 49, 342-44.
- _____ (1966) Some relations between minimal content, acquiescent-dissentient, and social desirability scales. Journal of Personality and Social Psychology, 3, 112-19.
- DAHLSTRÖM, W.G. and WELSH, G.S. (1960) An MMPI Handbook. University of Minnesota.
- DAMARIN, F. and MESSICK, S. (1965) Response styles as Personality Variables: A Theoretical Integration of Multivariate Research. New Jersey: Educational Testing Service.
- DEMPSEY, P. (1965) The concept of item difficulty in personality measurement: a note on Hanley's formulation. Educational and Psychological Measurement, 25, 143-7.
- DICKEN, C. (1963) Good impression, social desirability and acquiescence as suppressor variables. Educational and Psychological Measurement, 23, 699-720.

- DIERS, C.J. (1964) Social desirability and acquiescence in response to personality items. Journal of Consulting Psychology, 28, 71-7.
- DUFFY, K.E., JAMISON, K. and COMREY, A.L. (1969) Assessment of a proposed expansion of the Comrey personality factor system. Multivariate Behavioural Research, 4, 295-307.
- EDWARDS, A.L. (1957) The Social Desirability Variable in Personality Assessment and Research. New York: Dryden Press.
- EDWARDS, A.L. (1963) A factor analysis of experimental social desirability and response set scales. Journal of Applied Psychology, 47, 308-16.
- (1964a) The assessment of human motives by means of personality scales. In Nebraska Symposium on Motivation, 134-62. Nebraska: University of Nebraska Press.
- (1964b) Prediction of mean scores on MMPI scales. Journal of Consulting Psychology, 28, 183-5.
- (1964c) Social desirability and performance on the MMPI. Psychometrika, 29, 295-308.
- (1964d) Correlation of a Unidimensional Depression Scale for the MMPI with the Social Desirability scale. (Unpublished Data).
- (1965a) The social desirability variable: a broad statement. In Berg, I.A. (ed) Response Set in Personality Assessment. pgs. 32-47. Chicago: Aldine.
- (1965b) The social desirability variable: A review of the evidence, Ibid, pgs. 48-70.
- (1965c) Measurement of individual differences in ratings of social desirability and in the tendency to give socially desirable responses. Journal of Experimental Research in Personality, 1, 91-8.

- _____ (1966) Relationship between the probability of endorsement and social desirability scale values for a set of 2824 personality statements. Journal of Applied Psychology, 50, 238-9.
- _____ (1970) The Measurement of Personality Traits by Scales and Inventories. New York: Holt, Rinehart and Winston.
- EDWARDS, A.L. and DIERS, C.J. (1963) Neutral items as a measure of acquiescence. Educational and Psychological Measurement, 23, 687-98.
- EDWARDS, A.L., DIERS, C.J., and WALKER, B.N. (1962) Response sets and factor loadings on sixty-one personality scales. Journal of Applied Psychology, 46, 220-5.
- EDWARDS, A.L., GOCKA, E.F., and HOLLOWAY, H.C. (1964) The development of an MMPI acquiescence scale. Journal of Clinical Psychology, 20, 148-50.
- EDWARDS, A.L. and PANKS, P. (1964) Note on the normative data of tests measuring flexibility in cognitive processes. Psychological Reports, 14, 741-2.
- EDWARDS, A.L. and WALSH, J.A. (1963a) The relationship between the intensity of the social desirability keying of the scale and the correlations of the scale with Edwards' Social Desirability Scale and the first factor loading of the scale. Journal of Clinical Psychology, 19, 200-3.
- _____ (1963b) The relationship between various psychometric properties of personality items. Educational and Psychological Measurement, 23, 227-38.
- _____ (1964a) Response sets in standard and experimental personality scales. American Educational Research Journal, 1, 52-61.
- _____ (1964b) A factor analysis of 'I?' scores. Journal of Abnormal and Social Psychology, 69, 559-63.
- EDWARDS, A.L., WALSH, J.A. and DIERS, C.J. (1963) The relationship between social desirability and internal consistency of personality scales. Journal of Applied Psychology, 47, 255-9.

- EISENMAN, R. and HARE, R.T. (1970) Commonality of response to personality tests. Perceptual and Motor Skills, 30, 767-773.
- ELLIS, A. (1965) Rev. no. 86. In Buros, O.K.: The Sixth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- ENDLER, N.S. (1961) Conformity analysed and related to personality. Journal of Social Psychology, 53, 271-83.
- ENDLER, N.S. and HUNT, J. McV. (1968) S-R Inventories of Hostility and comparisons of the proportions of variance from persons, responses and situations for hostility and anxiousness. Journal of Personality and Social Psychology, 9, 309-15.
- EYSENCK, H.J. (1949a) Rev. no. 43. In Buros, O.K. The Third Mental Measurements Yearbook. New Brunswick: Rutgers University Press.
- EYSENCK, H.J. (1949b) Rev. no. 55. Ibid.
- (1957) The Dynamics of Anxiety and Hysteria. London: Routledge and Kegan Paul.
- (1959) The Maudsley Personality Inventory. London: University of London Press.
- and EYSENCK, S.B.G. (1963) The Eysenck Personality Inventory. London: University of London Press.
- EYSENCK, H.J. and EYSENCK, S.B.G. (1969) Personality Structure and Measurement. London: Routledge and Kegan Paul.
- EYSENCK, S.B.G. and EYSENCK, H.J. (1963a) Acquiescence response set in personality questionnaires. Life Science, 2, 144-7.
- (1963b) An experimental investigation of "desirability" response set in a personality questionnaire. Ibid., 343-55.

- _____ (1964) "Acquiescence" response set in personality inventory items. Psychological Reports, 14, 513-4.
- FEDER, C.Z. (1967) Relationship of repression-sensitization to adjustment status, social desirability, and acquiescence response set. Journal of Consulting Psychology, 31, 401-6.
- FISKE, D.W. (1966) On the coordination of personality constructs and their measurement. Human Development, 9, 74-83.
- _____ (1968) Items and persons. Formal duals and psychological differences. Multivariate Behavioral Research, 3, 393-401.
- _____ and PEARSON, P.H. (1970) Theory and techniques of personality measurement. Annual Review of Psychology, 21, 49-86.
- FLANAGAN, J.C. (1938) Rev. no. 922. In Buros, O.K.: The Nineteen Thirty-Eight Mental Measurements Yearbook. New Brunswick: Rutgers University Press.
- FOSTER, R.J. (1961) Acquiescent response set as a measure of acquiescence. Journal of Abnormal and Social Psychology, 63, 155-60.
- _____ and GRIGG, A.E. (1963) Acquiescent response set as a measure of acquiescence: Further evidence. Journal of Abnormal and Social Psychology, 67, 304-6.
- GHISELLI, E.E. (1963) Moderating effects and differential reliability and validity. Journal of Applied Psychology, 47, 81-6.
- GOCKA, E.F., and MARKS, J.B. (1961) Second-order factors in the 16 PF and the MMPI Inventory. Journal of Clinical Psychology, 17, 32-5.
- GOLDBERG, L.R. (1963) A model of item ambiguity in personality assessment. Educational and Psychological Measurement, 23, 467-92.

- GOLDFRIED, M.R. (1964) A cross-validation of the Marlowe Crowne Social Desirability Scale Items. Journal of Social Psychology, 64, 137-45.
- GUILFORD, J.P. (1934) Nebraska Personality Inventory. Lincoln: Sheridan Supply.
- _____ (1940) An Inventory of factors STDCR: Manual of Directions and Norms. Lincoln: Sheridan Supply.
- _____ (1959) Personality. McGraw-Hill Book Co.
- _____ and MARTIN, H.G. (1943a) The Guilford-Martin Inventory of factors GAMIN: Manual of Directions and Norms. Beverley Hills: Sheridan Supply.
- _____ (1943b) The Guilford-Martin Personnel Inventory. Beverley Hills: Sheridan Supply.
- GUILFORD, J.P. and ZIMMERMEN, W.S. (1949) The Guilford-Zimmerman Temperament Survey: Manual of Instructions and Interpretations. Beverley Hills: Sheridan Supply.
- GUTTMAN, L. and SCHLESINGER, I.M. (1967) Systematic construction of distractors for ability and achievement test items. Educational and Psychological Measurement, 27, 569-80.
- HAND, J. (1964) Measurement of response sets. Psychological Reports, 14, 907-13.
- _____ and BRAZZEL, C.O. (1965) Contamination in measures of acquiescence and social desirability. Psychological Reports, 16, 759-60.
- HANLEY, C. (1962) The "difficulty of a personality inventory item. Educational and Psychological Measurement, 22, 577-84.
- HARSH, C.H. (1953) Rev. no. 87. In Buros, O.K.: The Fourth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- HASE, H.D. and GOLDBERG, L.R. (1967) Comparative validity of different strategies of constructing personality inventory scales. Psychological Bulletin, 67, 231-48.

- HEILBRUN, A.B. Jr. (1964) Social learning theory, social desirability and the MMPI. Psychological Bulletin, 61, 377-87.
- HERSCH, P.D. and SCHLIEBER, L.E. (1967) Reliability and validity of internal-external control as a personality dimension. Journal of Consulting Psychology, 31, 609-13.
- HOLLAND, J.J. (1960) The relationships of the Vocational Preference Inventory to the 16 PF Questionnaire. Journal of Applied Psychology, 44, 291-6.
- HOLTZMAN, W.H. (1964) Recurring dilemmas in personality assessment. Journal of Projective Techniques, 28, 144-50.
- (1965) Personality Structure. Annual Review of Psychology, 16, 119-56.
- HUNDLEBY, J.D., PAWLIK, K., and CATTELL, R.B. (1964) The First Twenty-One General Personality Dimensions in Objective Behavioral Measurements. Chicago: University of Chicago Press.
- (1965) Personality Factors in Objective Test Devices: A Critical Integration of a Quarter Century's Research. San Diego: Knapp.
- JACKSON, D.N. (1970) A sequential system for personality scale development. In Spielberger, C.D. (ed.): "Current Topics in Clinical and Community Psychology". Vol. 2. New York: Academic Press.
- and LAY, C.H. (1968) Homogeneous dimensions of personality scale content. Multivariate Behavioral Research, 3, 321-37.
- JACKSON, D.N. and MESSICK, S. (1962) Response styles on the MMPI: Comparison of clinical and normal samples. Journal of Abnormal and Social Psychology, 65, 285-99.
- JENKINS, J.J. and SCHUELL, H. (1965) Acquiescent response set in Aphasics. Journal of Abnormal and Social Psychology, 70, 111-13.

- JERSON, A.R. (1965) Rev. no. 138. In Buros, O.K.: The Sixth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- KAPLAN, M.F. and ERON, L.D. (1965) Test sophistication and faking the T.A.T. situation. Journal of Projective Techniques and Personality Assessment, 29, 498-503.
- KASL, S.V., SAMPSON, E.E. and FRENCH, R.P. (1964) The development of a projective measure of the need for independence: a theoretical statement and some preliminary evidence. Journal of Personality, 32, 566-86.
- KATKIN, E.S. (1964) The Marlowe-Crowne Social Desirability Scale: Independence of psychopathology? Psychological Reports, 15, 703-6.
- (1966) Sex differences and the relationship between the Marlowe-Crowne Social Desirability Scale and MMPI indices of psychopathology. Journal of Consulting Psychology, 30, 564.
- KEHRLINGER, F. and ROKEACH, M. (1966) The factorial matrix of the F and D scales. Journal of Personality and Social Psychology, 4, 391-99.
- KLEIN, G.S., BARR, H.L. and WOLITZSKY, D.L. (1967) Personality. Annual Review of Psychology, 18, 467-560.
- KUSYSZYN, I. and JACKSON, D.N. (1968) A multimethod factor analytical appraisal of endorsement and judgement methods in personality assessment. Educational and Psychological Measurement, 28, 1047-61.
- LA FORGE, R. (1962) A correlational study of two personality tests. Journal of Consulting Psychology, 26, 402-11.
- LEFCOURT, H.M. (1966) Internal vs. external control of reinforcement: A review. Psychological Bulletin, 65, 206-20.
- LEVONIAN, E. (1961) A statistical analysis of the 16 PF Questionnaire. Educational and Psychological Measurement, 21, 589-96.

- COMREY, A.L., ZEVY, W., and PROCTER, D. (1959) A statistical evaluation of the Edwards Personal Preference Schedule. Journal of Applied Psychology, 43, 355-9.
- LIBERTY, P.J. Jr. (1965) Methodological considerations in the assessment of acquiescence in the MA and SD scale. Journal of Consulting Psychology, 29, 37-42.
- LIEBERMAN, L.R. (1966) On actuarial methods as appropriate strategy for the validation of diagnostic tests. Psychological Review, 73, 262-4.
- LINGOES, J.C. (1965a) Rev. no. 138. In Buros, O.K.: The Sixth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- (1965b) Rev. no. 143. Ibid.
- LORR, M. (1965) Rev. no. 171. Ibid.
- MADDI, S.R., and ANDREWS, S.L. (1966) The need for variety in fantasy and self-description. Journal of Personality, 34, 610-25.
- MADDI, S.R. and BERNE, N. (1964) Novelty of productions and desire for novelty as active and passive forms of the need for variety. Journal of Personality, 32, 270-7.
- MADDI, S.R., PROPST, B.S. and FELDINGER, J. (1965) Three expressions of the need for variety. Journal of Personality, 33, 82-90.
- MARLOWE, D. and GERGEN, K.G. (1969) Personality and social interaction. In Lindzey, G. and Aronson, E. (eds.): The Handbook of Social Psychology. Vol. 3. Addison-Wesley Publishing Co.
- MAURICE, L. (1965) Rev. no. 174. In Buros, O.K.: The Sixth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- McGEE, R.K. (1962a) Response style as a personality variable: By what criterion? Psychological Bulletin, 59, 284-95.

- (1962b) The relationship between response style and personality variables: I. The measurement of response acquiescence. Journal of Abnormal and Social Psychology, 64, 229-33.
- (1962c) The relationship between response style and personality variables: II. The prediction of independent conforming behavior. Ibid, 65, 347-51.
- MEEHL, P.E. (1945) The dynamics of "structured" personality tests. Journal of Clinical Psychology, 1, 296-303.
- MESSICK, S. (1961) Personality structure. Annual Review of Psychology, 12, 93-128.
- (1966) The Psychology of Acquiescence: An Interpretation of Research Evidence. Princeton, New Jersey: Educational Testing Service.
- MICHAEL, W.B., COMREY, A.L. and FRUCHTER, B. (1963) J.P. Guilford: Psychologist and teacher. Psychological Bulletin, 60, 1-34.
- MITCHELL, J.V. Jr. (1963) A comparison of the first and second order dimensions of the 16 PF and CPI inventories. Journal of Social Psychology, 61, 151-166.
- and PIERCE-JONES, G. (1966) A factor analysis of Gough's CPI. Journal of Consulting Psychology, 24, 453-6.
- MOSHER, D.L. (1966) The development and multitrait-multi-method matrix analysis of three measures of three aspects of guilt. Journal of Consulting Psychology, 30, 25-9.
- NEILL, J.A. and JACKSON, D.N. (1970) An evaluation of item-selection strategies in personality scale construction. Educational and Psychological Measurement, 30, 647-61.
- NORMAN, W.T. (1963) Personality Measurement, faking, and detection: An assessment method for use in personnel selection. Journal of Applied Psychology, 47, 225-41.

- PAWLIK, K. and CATTELL, R.B. (1964) Third-order factors in objective personality tests. British Journal of Psychology, 55, 1-18.
- PETERSON, D.R. Scope and generality of verbally defined personality factors. Psychological Review, 72, 48-59.
- RORER, L.G. (1963) The Function of Item-Content in MMPI Responses. Doctoral Thesis, Minneapolis: University of Minnesota.
- (1955) The great response-style myth. Psychological Bulletin, 63, 129-156.
- and GOLDBERG, L.R. (1965a) Acquiescence on the MMPI? Educational and Psychological Measurement, 25, 801-17.
- RORER, L.G. and GOLDBERG, L.R. (1965b) Acquiescence and the vanishing variance component. Journal of Applied Psychology, 49, 422-30.
- ROTTER, J.B. (1966) Generalized expectancies for internal vs. external control of reinforcement. Psychological Monographs, 88 (1), 28 p.p.
- SARASON, I.G. and SMITH, R.E. (1971) Personality. Annual Review of Psychology, 22, 393-446.
- SASSENATH, J.M. (1964) A factor analysis of rating scale items on the Test Anxiety Questionnaire. Journal of Consulting Psychology, 28, 271-77.
- SAUNDERS, D.R. (1959) Rev. no. 65. In Buros, O.K.: The Fifth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- SCHNEIDER, J.M. (1968) Skill vs. chance activity preference and locus of control. Journal of Consulting and Clinical Psychology, 32, 333-37.
- SCOTT, W.A. (1963) Social desirability and individual conceptions of the desirable. Journal of Abnormal and Social Psychology, 67, 574-85.

- SECHREST, L. (1968) Testing, measuring and assessing people. In Borgatta, E.F. and Lambert, W.W. (eds.): Handbook of Personality Theory and Research. Chicago: Rand McNally and Co.
- SEGALL, M.H. (1963) Acquiescence and "identification with the aggression" among acculturating Africans. Journal of Social Psychology, 61, 247-62.
- SELLS, S.B., DEMAREE, R.G. and WILL, D.P. (1970) Dimensions of Personality: I. Conjoint factor structure of Guilford and Cattell trait markers. Multivariate Behavioral Research, 5, 391-422.
- SIMON, W.B. (1967) A simple method for obtaining social desirability values. Educational and Psychological Measurement, 27, 135-7.
- SOLOMON, C. and KLEIN, E. (1963) The relationship between agreeing response set and social desirability. Journal of Abnormal and Social Psychology, 66, 176-9.
- STEPHENSON, W. (1953) Rev. no. 49. In Buros, O.K. The Fourth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- (1965) Rev. no. 138. In Buros, O.K.: The Sixth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- STOLLAK, G.E. (1965a) EPPS performance under social desirability instructions. Journal of Personality and Social Psychology, 2, 430-42.
- (1965b) EPPS performance under social desirability instructions: College females. Psychological Reports, 16, 119-22.
- STONE, L.A. (1965a) Relationships between responses to the Marlowe-Crowne social desirability scale and MMPI scales. Psychological Reports, 17, 179-82.
- (1965b) Reliability estimate for favourability (social desirability) ratings of MMPI items. Psychological Reports, 16, 720.

- STRICKER, L.J. (1963) Acquiescence and social desirability response style, item characteristics, and conformity. Psychological Reports, 12, 319-41.
- TAYLOR, J.A. (1953) A personality scale of manifest anxiety. Journal of Abnormal and Social Psychology, 48, 285-90.
- THORNDIKE, R.L. (1959) Rev. no. 37. In Buros, O.K.: The Fifth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- TURNER, C.B. and FISKE, D.W. (1968) Item quality and appropriateness of response processes. Educational and Psychological Measurement, 28, 297-315.
- VERNON, P.E. (1964) Personality Assessment: A Critical Survey. Butler and Fannes, Ltd.
- WAHLER, H.J. (1965) Item popularity and social desirability in the MMPI. Journal of Applied Psychology, 49, 439-45.
- WEBB, E.J., CAMPBELL, D.T., SCHWARTZ, R.D. and SECHREST, L. (1966) Unobtrusive measures: Nonreactive Research in the Social Sciences. Chicago: Rand McNally.
- WIENER, D.N. (1948) Subtle and obvious keys for the MMPI. Journal of Consulting Psychology, 12, 164-70.
- WIGGINS, J.S. (1962) Strategic method and stylistic variance in the MMPI. Psychological Bulletin, 59, 224-42.
- (1964) Convergences among stylistic response measures from objective personality tests. Educational and Psychological Measurement, 24, 557-62.
- (1966) Social desirability estimation and "faking good" well. Educational and Psychological Measurement, 26, 329-41.
- (1968) Personality Structure. Annual Review of Psychology, 9, 293-350.
- and LEVELL, V.R. (1965) Communality and favourability as sources of method variance. Educational and Psychological Measurement, 25, 399-412.

- WIGGINS, N. (1966) Individual viewpoints of social desirability. Psychological Bulletin, 66, 68-77.
- WITKIN, H.A., DYK, R.B., FATTERSON, H.F., GOODENOUGH, D.R. and KARP, S.A. (1962) Psychological Differentiation. New York: J. Wiley and Sons, Inc.
- WITTENBORN, J.R. (1953) Rev. no. 87. In Buros, O.K.: The Fourth Mental Measurements Yearbook. New Jersey: The Gryphon Press.
- WRIGHTS, S. (1955) Some Psychological and Physiological Correlates of Certain Academic Underachievers. Unpublished Doctoral Dissertation: University of Chicago.
- ZUCKERMAN, M., KOLIN, E.A., PRICE, L., and ZOOB, I. (1964) Development of a Sensation-Seeking Scale. Journal of Consulting Psychology, 28, 477-82.
- ZUCKERMAN, M. and LINK, K. (1968) Construct validity for the Sensation Seeking Scale. Journal of Consulting and Clinical Psychology, 32, 420-6.



406-9077

