

SHOULD THE 16 PF BE USED IN PERSONNEL SELECTION ?

P. E. BULL

University of Waikato

The validity of the 16 PF is considered in terms of the relationship between the factors and behaviour ratings, the stability of the item-factor relationships, and the general problem of factorial validity. The conclusion is that the 16 PF is still a research hypothesis and is not suitable for use in selection decisions until specific predictive models have been established.

One of the attractions of the Sixteen Personality Factor Questionnaire (16 PF) is the skill with which Cattell has presented his information over the years since the first documents were published. Unfortunately this presentation is a mixture of clarity and complexity that is difficult to assimilate. The questionnaire purports to measure the most salient of the personality traits found by Cattell in his factorization of the personality sphere (Cattell and Butcher, 1968). With the recent publication of new British standards (Saville, 1972; Adcock, 1973) it seems likely that this test may remain one of the main instruments of personality assessment. In New Zealand it is probably the predominant instrument of personality assessment in use by management consultants and many government bodies to select personnel for such diverse occupations as intelligence staff, salesmen, and accounting managers (Hesketh, 1973). A question that needs to be asked is whether this questionnaire should be in use at all in applied fields, on the basis of the accumulated evidence. In his review of Saville's booklet Adcock notes that the British normative studies do not validate the questionnaire. What would be needed to establish factorial validity is a refactorization of the questionnaire to assess the integrity of the item groupings. But for applied psychologists, the kind of validity that needs to be established rests on test-criterion relationships. It is clear that there is not sufficient information to establish norms in New Zealand and that there is no substantial evidence on the validity of the constructs that underlie the 16 PF. In spite of this situation the book review by Adcock seems to accept that the test will be used. "It would appear, therefore, that New Zealand users of the test would be well advised to use the British norms until more N.Z. data are available." (Adcock, 1973, p. 44). The problem is that there is a lack of evidence that the factors in the 16 PF will maintain stable relationships with the kinds of criterion variables typically available to consultants. I had the disturbing experience of having a corrected multiple correlation of .89 shrink on validation to .26 (Bull, 1973). A close analysis of the data showed that the means

were stable over the sample and the predicted group, and that there was no attenuation due to the rejected applicants, which strongly suggests that the correlations were not stable between the test and the criterion. In view of this problem it could be useful to consider some of the theoretical issues that relate to the operational validity of the test.

Essentially there are two kinds of evidence offered by Cattell, Eber, and Tatsuoka (1970) to demonstrate the validity of the 16 PF; the correlation between questionnaire data derived from the test and life rating data, and secondly, a validity that is said to be confirmed by co-ordinated factor analyses.

Schaie's (1962) research is referred to in the 16 PF handbook as part of the evidence for the trait structure. The study examined the relationship of life and questionnaire data but any relationship between the 16 PF and the study must at best be regarded as showing an inferred, transitive relationship, as the High School Questionnaire was utilized and not the 16 PF. In the study, rating data were gathered from the supervisors of 43 institutionalized delinquent girls and correlated with the questionnaire. The eleven correlations between the test scores and ratings were low with only three significant at the .05 level. It is difficult to see how this data can be used as evidence to support a relationship between life rating data and the 16 PF.

Central to the conception of the 16 PF and its validity is the author's commitment to factor analysis. Brody (1972) argues that the difficulty of factor replication or the instability of factors is the central problem that undermines the position of the trait theorists. This argument is illustrated in a study reported by N. V. Adcock (1972) which examined 14 primary personality factors with items selected mainly from the 16 PF. The number of factors extracted was decided by the Kaiser-Guttman criterion. Unfortunately for the original hypothesis 26 factors qualified. She decided that 10 were 'non trivial' and that the rest were "multidetermined". Six of the twelve factors subsequently discussed were not marked by hypothesized item-factor relations. It could be that this finding is not to be entirely unexpected if we think about the nature of factor analysis. The vector of eigenvalues expresses the linear components of the variance of the data correlation matrix. When we make significant changes to the composition of this data matrix we can expect the factor loadings to change and the deviation of the factors between the original set and the replication will increase down the chain, as each is extracted, from the point at which one of the factors in the replication study assumes a different position from the same factor in the original study. To expect that the factors would replicate when the data matrix has been changed means that we consider that the items deleted played no part in locating the original factors, and that the items we have added will also not contribute to the location of the factors that are to be replicated. When we add to this type of problem the effects of different precisions relating to par-

ticular computer programs and the rounding off problems when a large number of factors are extracted, then failure to replicate when the raw data matrix is markedly different might be expected. From a psychologist's point of view it is disturbing that investigators who have done a great deal of work in this area were unable to exercise predictive control over their variables.

In assessing work on the 16 PF at the I.A.A.P. conference Adcock, Adcock, and Walkey (1972) reported

The evidence we have been considering seems to provide a damning indictment of the personality scheme which Cattell and many others have devoted so much time to developing, but before we decide that the product of half a life-time is crumbling about our ears, let us examine the position more fully. Certainly all the recent evidence seems to indicate that the items themselves do *not* define the factors to which they are alleged to relate, but one point which appears to have been overlooked is that the factors that *do* emerge are in many cases strikingly similar to the 16 PF factors as described (p. 2).

The problem now is that the scores on the original factors that are being used in selection are derived from items that may or may not load on the factors which receive their contribution.

There are two central problems emerging from studies on the 16 PF: first, researchers are recovering different numbers of factors, and the items in the questionnaire are loading in a manner which was not predicted.

The literature shows that control is not easily attained in factor analysis studies (Glass and Taylor, 1966), and issues relating to the number of factors are not unequivocally resolved in the types of factor analyses used in the Adcock studies. This is a crucial problem in evaluating the 16 PF as the validity of the test is established on the fidelity with which it represents the underlying structure. A few minutes spent using a standard computer program, altering the number of factors should convince the sceptic that this is not a minor problem. Francis (1972) constructed a model that had two underlying factors distributed over ten variables. When five factors were extracted, all with eigenvalues greater than one, there was no resemblance to the true loadings. The compelling conclusion of his study is that factor analysis should only be used in psychological research when the number of factors is known. This point of view creates problems because Cattell uses factor analysis to uncover relationships. Cattell (1959) contends that psychologists are apeing the physical sciences in the inappropriate use of univariate statistics when multivariate methods are more appropriate to the state and nature of psychological knowledge. He believes that it is better to reduce the large number of possible variables prior to hypothesis formulation and manipulative experimentation, and one does not disagree with this (Ashby, 1960): but he proceeds to this reduction by the repeated application of factor analysis, and there is a subtle transition in his writings from a state of cautious investigation to one

of having shown a stable and unambiguous structure that occupies a central space in questionnaire and personality rating variables. It is the use of factor analysis to discover structure that places Cattell in the position of proceeding without due regard for the state and certainty of mathematical theory, and this would contribute to the difficulties of researchers in trying to achieve similar item-factor relationships.

In the adventures of "Tom Swift and His Electric Factor Analysis Machine", Armstrong (1967) examined the hazards of derivation of theory by means of factor analysis. He argues that a minimal requirement for the use of factor analysis is a stipulation of prior assumptions as to the number of factors and the nature of their relationships. The dangers of using repeated factor analysis to sustain a theory has been shown by the studies of Humphreys, Ilgen, McGrath, and Montanelli (1969), and Francis (1972). The former paper showed that ostensibly meaningful factors can be found from inter-correlated random error, and Francis had loadings for three non-existent factors reproduced in three independent samples.

This paper is not intended to cause arguments on factor analysis, though this might be its effect. It is intended to question the use of unvalidated test instruments based on the uncritical use of factor analysis. A factor analysis may indicate structure but in most designs it has nothing to do with the validity of a test in respect to applied psychology. The measure of a science is its ability to predict, and we must demonstrate a relationship between predictor and criterion variables in order to justify the continued use of a test, in anything other than a research situation. The ethical problems of failure to validate any instrument by the standards of measurement theory and empirical evidence presents grave problems to the practice of psychology, and puts us in danger of engaging in "personality description after the manner of P. T. Barnum" (Meehl, 1956). Until research studies show clear relationships between the 16 PF factors as scored and criterion variables and in so doing demonstrate that the use of these data reduces the error of judgment in selection situations, this questionnaire must be regarded, despite its long history, as a research hypothesis.

REFERENCES

- Adcock, C. J. Review of Peter Saville, *The British Standardisation of the 16 PF*. *New Zealand Psychologist*, 1973, 2, 43-44.
- Adcock, N. V. The 16 PF test factors. Paper presented at New Zealand Psychological Society Conference, 1972.
- Adcock, N. V., Adcock, C. J., and Walkey, F. H. Basic dimensions of personality. I. A.A.P. Congress, Liege, 1971.

- Armstrong, J. S. The derivation of theory by means of factor analysis: Tom Swift and his electric factor analysis machine. *The American Statistician*, 1967, 21, 17-21.
- Ashby, W. R. *Design for a brain*. London: Chapman and Hall, 1960.
- Brody, N. *Personality research and theory*. New York: Academic Press, 1972.
- Bull, P. E. The Sixteen Personality Factor Questionnaire and the Motivational Analysis Test in the prediction of achievement of insurance agents. Dissertation, University of Waikato, 1973.
- Cattell, R. B. Personality theory growing from multivariate quantitative research. In S. Koch (Ed), *Psychology: A study of science*. New York: McGraw-Hill, 1959.
- Cattell, R. B. and Butcher, H. J. *The prediction of achievement and creativity*. Indianapolis: Bobbs-Merrill, 1968.
- Cattell, R. B., Eber, H. W., and Tatsuoka, M. M. *Handbook for the Sixteen Personality Factor Questionnaire*. Illinois: Centre for Personality and Ability Testing, 1970.
- Francis, I. Factor analysis: Fact or fabrication. Paper presented to the Mathematics colloquium, N.Z., 1972.
- Glass, G. V. and Taylor, P. A. Factor analytic methodology. *Review of Educational Research*, 1966, 36, 566-587.
- Hesketh, B. New Zealand Institute of Personnel Management survey on testing in industry. 1973.
- Humphreys, L. G., Ilgen, D., McGrath, D., and Montanelli, R. Capitalization on chance in rotation of factors. *Education and Psychological Measurement*. 1969, 29, 259-271.
- Meehl, P. E. Wanted a good cookbook. *American Psychologist*, 1956, 11, 263-272.
- Saville, P. *The British standardization of the 16 PF*. Illinois: Centre for Personality and Ability Testing, 1972.
- Schaie, K. W. On the equivalence of questionnaire and rating data. *Psychological Reports*, 1962, 10, 521-522.