Adv. Behav. Res. Ther. Vol. 5, pp. 141–202, 1983.
0146-6402/83 $0.00+.50

THE FACTORS IN PERSONALITY QUESTIONNAIRES AMONG NORMAL SUBJECTS

Paul Kline and Paul Barrett
Department of Psychology, University of Exeter

INTRODUCTION

Reference to the most recent edition of the Mental Measurements Year Book (Buros, 1978) highlights what is the major difficulty with personality questionnaires: just what do they measure. In Buros can be found a huge number of personality inventories, and these are only those that have been published. In addition research papers in personality make use of unpublished questionnaires. Many of the authors of these tests claim in the manuals that their sets of factors are definitive or embracing of the most variance. Some tests, e.g. the M.M.P.I. were not constructed by factor analytic techniques, but by criterion-keyed methods, items being chosen if they could discriminate criterion groups. Given, such a psychometric tower of Babel, it is clearly necessary to attempt to clarify the field.

This is the aim of this monograph. We intend to demonstrate from our own research and from a critical analysis of other research in the field, that the variables measured by this plethora of tests, despite their disparate labels, are not in fact distinct and that in reality a small and well validated set of factors can be extracted from these tests. From this it follows that the majority of personality inventories are not highly valid and can be ignored. In addition we shall show that with a small number of tests these psychologically meaningful factors can be efficiently measured.

CAUSES OF THE CONFUSION

There are many causes for this confusion of variables and these must be discussed so that the force of the research to be reported can be grasped.

(1) The Problems with Criterion-Keyed Tests

Kline (1979) has fully examined the difficulties involved in the construction of criterion-keyed tests. Apart from the practical difficulty of establishing adequate criterion groups, there is the more important problem of the psychological meaning of scales constructed by this method. Items are selected if they can discriminate among groups. However, groups may differ along more than one dimension. Thus any scale so constructed may easily measure several dimensions. Multi-dimensional scales are misleading when used as sources for quantification because two identical scores can be differently composed. If a scale consists of items measuring two factors, for example a score of 10 can reflect 5 + 5, 6 + 4 and so on. This is a serious objection, and is the reason why unifactorial scales are to be preferred.

An implication in the above argument is that there is no obvious psychological meaning to a criterion keyed scale. All that can be said of it is that it discriminates a specific group. From this information (because groups differ in many ways), it is not possible to identify what the scale is measuring especially, as is the case, usually, when any meaningful differences between the groups are not known. We say as is usually the case because, if the differentiating factors are known it is sensible to measure those factors and not rely on criterion-keyed test construction.

In summary, therefore, we can say that even if criterion-keyed tests are shown to be efficient screening devices, then this by no means indicates what the scales are measuring. In all probability such scales are little more than a hotchpotch of items and it would not be surprising if such scales, despite their often forceful labels, in fact measured a number of common factors. The only way such an argument can be shown to be true or false is, of course, to submit such scales to factor analysis preferably by items rather than scales, and to actually see on what factors the items load. This is one body of research which we shall examine later in this monograph.

(2) The Causes of Different Factors Emerging from Different Scales which have been Constructed by Factor Analysis or Analogous Methods

Although our argument above explains why criterion-keyed tests may measure similar factors, it fails to account for the fact that a large number of different factors are claimed to be measured by tests constructed by factor analysis. This point that different factor analyses yield different factors has led many psychologists (e.g. Heim, 1975) to question the value of factor analysis as a useful statistical method for psychology.

Kline (1979) and Cattell and Kline (1977) have examined this problem of the multiplicity of factors (which actually is more apparent than real) in considerable detail and it is possible to summarise the arguments quite briefly. One possibility is that the factors in different tests are really the same but they have been labelled or even interpreted differently. For example, this is certainly the case with extraversion, Eysenck and Eysenck's (1975) title, and exvia as the factor is identified by Cattell (e.g. 1973). When we come to examine the findings this is a possibility which must be always considered.

The second reason for the emergence of apparently different factors is more common. It results from the use (and misuse) of different methods of factor analysis, especially different rotational procedures, and from the consequent failure of many factor analyses to reach simple structure. On rerotation (Carroll, 1981) or in subsequent research where simple structure is obtained (Cattell and Kline, 1977) many of these differences disappear. This being the case it is clearly necessary to discuss both why and how simple structure should be obtained together with the requirement of sound factor analyses.

The Necessity for Simple Structure

Simple structure, as first defined by Thurstone (e.g. 1947) can be considered to be obtained in essence when each factor has a few high loading variables with the remaining subset loading as nearly as possible to zero. The rationale for requiring that simple structure be reached is elegant. Since there is an infinity of mathematically equivalent solutions to any factor analysis and since each solution can be regarded as an hypothesis to account for the observed correlations between the variables, then by the law of parsimony, the simplest solution is the one chosen. Thurstone's aim of reaching simple structure, as he defined it, is the endorsement of Occam's razor.

However, there is a further and perhaps more important advantage in simple structure solutions to factor analyses. This lies in their replicability and their ability to produce factors which reflect truly underlying dimensions, when these are known. It has been shown by Cattell (1973, 1978) that in factor analyses of data where the dimensions are clear (from other evidence) simple structure solutions isolate these dimensions, as in Thurstone's box problem, or with artificial exemplar plasmodes. Thus according to Cattell, and there is little disagreement here among the leading factor analysts, simple structure is the aim of the factor analyses. The technical requirements for adequate analyses, therefore, which we discuss below are those that enable simple structure to be obtained. In addition we set out some further requirements that ensure, as far as possible, that factors are not inadvertently missed.

An Adequate Methodology for Replicable Factor Analyses

As we have indicated, there is now considerable agreement as to how simple structure should be obtained. Our recommendations are based upon the work of Harman (1976), Cattell (1978) and Carroll (1981). We are not asserting that our suggestions are the only viable methods: rather that these and other closely related procedures will more likely reach simple structure. The further any methods diverge from those proposed here, the less likely it is that they can reach a simple, replicable solution.

Sampling of Variables

Where factor analyses are used to explore a field, as often is the case in Personality on account of its complexity, then it is obviously necessary, to avoid missing factors, to sample the full range of variables. Cattell over the years at least attempted to do this when he rated subjects on all descriptive words (the personality sphere) and found the main dimensions before beginning to construct tests (see Cattell (1957) for a full description of this monumental effort). Whether he was successful in this has been called into question by Howarth (1976), but the principle remains. For factors to emerge with any clarity it is generally wise to target at least three variables to load on them.

More formally, the inequality derived by Thurstone (1947, p.239), Lederman (1937), and Harman (1960, pp. 69–73) concerning the number of independently determinable factors can be rearranged to solve for n the number of variables:

$$n \geqslant [(2r+1) + \sqrt{8r+1}]/2 \tag{1}$$

this inequality may be viewed as an indicator of the minimum number of variables that must be included in a correlation matrix in order to determine r common factors. Notice that to determine one factor, at least three variables are required. To determine five factors, at least nine variables are required. In general, overdetermination of factors is a good policy, favouring better interpretation and replicability.

Sampling of Subjects

In exploratory factor analyses it is important to sample a wide range of subjects. If subjects are homogenous in respect of a variable, there will be insufficient variance for it to emerge with the clarity it should. Thus intelligence may not emerge well in a sample of honours students. Obviously such strictures do not apply if the factor analysis aims to uncover the personality structure within a particular group, but this is a different question.

The Numbers of Observations and Variables

It is mathematically desirable to factor cross product, covariance, or correlation matrices computed from data matrices where the number of observations N is greater than the number of variables n. If the observed data matrix is square ($N = n$) or oblong ($N > n$), then the number of determinable components (the rank of the matrix) in the association matrix will be less than n, the number of variables. Of course, such initial matrix rank reduction is artificial, due to measurement error rather than to meaningful data reduction. Although there is some evidence that analysis of $N < n$ association matrices can produce somewhat similar results to those from $N > n$ matrices, it is not a methodology to be recommended. Nunnally (1978) provides an excellent summary of the arguments against numerically (and statistically) undersampling the variable domain. (On a purely practical level, factoring such matrices introduces problems of negative principal component eigenvalues, and the computing of generalised inverse matrices for other methods of component and common factor analysis).

As to the matter of just how many observations should be made on each variable, it is a matter of dispute among factorists of repute as to how great this ratio should be for reliable factor loadings. Nunnally (1978) has argued that 10:1 is the necessary ratio. The other extreme is supported by Guilford (1956) who claims that 2:1 is satisfactory. Cattell (1978) and Gorsuch (1974), for example, fall between these figures.

In fact, there is no rationale given for the size of this ratio by these writers; most refer to their considerable experience with factor analysis. Barrett and Kline (1981a) carried out an empirical examination of this problem using the Eysenck Personality Questionnaire (EPQ) and Cattell's 16 Personality Factor (16PF) items. With a high ratio of observations (subjects) to variables,(20:1) they obtained a factor structure so clear that it formed a sensible baseline against which to judge smaller samples. In fact, they found that the main factors emerged clearly with ratios as low as 2:1 and that with ratios of around 3:1 there was virtually no difference from the original large sample. The influence of the actual number of subjects on the results was also studied. Here it was shown that a large N, presumably because it reduces the standard error of the correlations, makes for more reliable results. Certainly, factorings on samples of less than 100 subjects must be treated with great caution. However, it must be stressed that these were data with a remarkably clear structure and, with factors of lesser clarity, it is possible that different findings might be obtained. Nevertheless it does seem that Nunnally has adopted a criterion that is unnecessarily stringent. Certainly ratios should not drop below 2:1 for reliable factors. If they do so replication is essential with a new sample.

Component or Common Factor Analysis

Perhaps the four most popular methods of factor analysing data are:

1. Principal Components Analysis (PCA)
2. Image Component Analysis (ICA)
3. Principal Factor Analysis (PFA)
4. Maximum Likelihood Factor Analysis (MLFA)

The first two methods are those most popularly associated with component factor analysis models. The model form underlying orthogonal component analysis is:

$$z_{ji} = a_{jl}\, F_{li} + a_{jz}\, F_{2i} + \ldots\ldots\ a_{jm}\, F_{mi} \tag{2}$$

where m = the number of factors
n = the number of variables
$m = n$
N = the number of observations
z_{ji} = the standardised value of an observation i on variable j
a_{jm} = the weight of each variable j on the factors F_1 to F_m (In this model, they are the factor loadings given in the pattern matrix).
F_{mi} = the value or 'score' on each factor for observation i

In matrix notation this becomes:

$$Z = AF \tag{3}$$

where Z = an $(n \times N)$ matrix composed of the elements z_{ji}
A = an $(n \times m)$ matrix composed of the elements a_{jm}
F = an $(m \times N)$ matrix composed of the elements F_{mi}

$$R = \mathrm{AA}' \tag{4}$$

where R = an $(n \times n)$ symmetric matrix with ones in the main diagonal.

Both PCA and ICA account for all the variance in an association matrix. For example, PCA extracts both error and specific variable variance in addition to the 'common' variable variance. While PCA calculates components directly from an observed data association matrix, ICA calculates components from a Gramian rank reduced image variable association matrix. Each image variable observation being the predicted data value from regressing that variable's observations against the remaining n–1 variable observations. The variable anti-image is the error of estimate of the multiple regression predicted values. The subsequent association matrix, either covariance or correlation, is then known as the image association matrix. The reason for this somewhat extensive pre-factoring calculation is to ensure that all variance explained by the factors is common (shared by all other variables), rather than including specific variable variance as with PCA. By specifying the common parts of variables so explicitly, the problems of estimating common variance by any other methods are greatly alleviated.

The methods PFA and MLFA are those most popularly used for common factor analysis. The model form for orthogonal common factor analysis is:

$$z_{ji} = a_{j1} F_{1j} + a_{j2} F_{2i} + \ldots\ldots a_{jm} F_{mi} + u_{jk} Y_{ki} + e_{ji} \tag{5}$$

where $m < n$

k = $m + 1$ to n unique factor subscript
z_{ji} = the standardised value of an observation i on variable j
a_{jm} = the weight of each variable on the common factors F_1 to F_m
F_{mi} = the value or 'score' on each common factor for observation i
u_{jk} = the weight of each variable on the unique factors $Y_{m+1 \text{ to } Y_n}$
Y_{ki} = the value or 'score' on each unique factor for observation i
e_{ji} = the error associated with each observation i on variable j

In matrix notation this becomes:

$$Z = AF + UY + E \tag{6}$$

where Z = an $(n \times N)$ matrix composed of the elements z_{ji}
A = an $(n \times m)$ matrix composed of the elements a_{jm}
F = an $(m \times N)$ matrix composed of the elements F_{mi}
U = an $(n \times n\text{–}m)$ matrix composed of the elements u_{jk}
Y = an $(n\text{–}m \times N)$ matrix composed of the elements Y_{ki}
E = an $(n \times N)$ matrix composed of the elements e_{ji}

Defining U_n as a matrix of order $(n \times n)$ of unique factor variable variances, E_n as the matrix of order of $(n \times n)$ of error variances for each variable, then assuming that the common factor variable variance is completely extracted, both U_n and E_n will be diagonal matrices.

Thus the correlation matrix is then represented as:

$$R = AA' + U_n\, U_n' + \mathrm{E}_n\, E_n' \tag{7}$$

which can also be written as:

$$R = AA' + U_n^2 + E_n^2 \tag{8}$$

In practice, the uniqueness and error is extracted in a single variable specificity matrix. There are no realistically exact methods for partialling measurement error from specific variable variance.

PFA, like ICA above, attempts to solve the same problem of preanalysis specific variance elimination. However, this particular elimination process owes more to heuristics rather than a coherent theoretical model. In general, this method leads to the rank reduction of a matrix at the cost of producing negative variance eigenvalues and indeterminate factor solutions. The indeterminacy issue arises directly from the problem of inaccuracy of specific variance estimation.

MLFA, in contrast to all three factoring methods above, is based upon a radically different methodology. Under an assumption of a given number of m common factors, MLFA is applied in order to obtain estimators of the universe factor loadings from the sample of N observations on the n variables. Because of the statistical background and theory from which this method derives, subsequent tests of significance can be applied to determine the adequacy of the hypothesis regarding the number of factors. (The test variable is distributed as a Chi Square variate.) The basic factor model underlying this particular analysis, however, still differentiates between common and unique factors. Unlike the other methods, MLFA does not seek implicitly to maximise the variance extracted for each factor in turn, rather it attempts to maximise the likelihood of occurrence of each variable loading on each factor.

Thus, most factorists would agree that the problem of factor extraction is identifying those factors that 'truly' represent common variance rather than those which represent specific variable and/or error variance. Various arguments have been offered over the years criticising or eulogising particular factor extraction techniques. For example, Lee and Comrey (1979) strongly criticise PCA with regard to the 'unrealistic elevation of the amount of common factor variance analysed . . .' (p.301). Carroll (1981) still argues that PFA should be used because in certain artificial examples, differences between the factors of PCA and PFA can be shown to occur. However, given Harman's (1976) assertion that with a matrix of more than ~ 25 variables, the differences between PFA and PCA are trivial, the arguments as to the 'best' method seem also to be trivial. Factor analysis in psychometrics is not a one-off mathematically exacting procedure as might be expected when used on simple variables such as the dimensions of a box, balls bouncing on billiard tables, or cups of coffee. Because of the very nature of the variables under examination, slight over or underestimation of communality and numbers of factors is most probable. Hence the resolving of such 'expected' errors by progressive replication of factor studies, and the seeking of external criteria as additional explanatory evidence in confirming or complementing the existence of such factors.

This is the methodology that yields reliable factorial data. No doubt many artificial examples can be produced where each method of factoring fails abysmally. However, until the crucial artificiality conditions can be shown to be present in even a minority of real data matrices, this 'evidence' is at best specious. The golden rule for any psychometric factor analytic study must be to replicate and externalise. That is, make sure that factor results are stable over at least another independent sample of subjects (albeit from the same subject domain). In addition, relate the 'discovered' variables to those not only already isolated within the research and/or applied fields, but also adjudged relevant and reliable. Whether these variables are behavioural or factor analytically derived ones is not so important at this stage. What is important is that the study does not stand in total isolation, simply a mathematical solution in a relative void of confirmatory evidence.

How Many Factors should be Rotated?

This, according to Cattell (e.g. 1973), is one of the crucial issues in accounting for the different findings in the factorisation of questionnaires. If too many factors are rotated, factors tend to split, thus losing clarity and generality; (in Cattell's terminology .. 'bloated specifics'.) If too few are put into the rotation, the resultant factor solution dimensionality is compressed artificially into a few large general factors. Thus clarity is once again lost; (in Cattell's terminology . . . 'grounded secondaries'.) It is, therefore, essential to determine the significant factors to rotate.

Hakstian *et al.* (1982) have recently examined this problem by defining a framework within which factor extraction may be discussed. Instead of simply considering the two factor analysis models presented above in eqs (2) and (5), they follow the approach of Tucker *et al.* (1969) in proposing a 'Middle' factor model. Rather than conceive of either common *or* specific factors, the model postulates *major* common factors, *minor* common factors, and the unique/error factors. Thus eq (6) above becomes:

$$Z = AF_{\text{maj}} + BF_{\text{min}} + UY + E \tag{9}$$

where min = the minor common factor subscript
maj = the major common factor subscript
mn = the number of minor common factors
mj = the number of major common factors
A = an $(n \times mj)$ matrix composed of the weights of each variable on the major common factors
F_{maj} = an $(mj \times N)$ matrix composed of the values or 'scores' of each observation i on the major common factors
B = an $(n \times mn)$ matrix composed of the weights of each variable on the minor common factors
F_{min} = an $(mn \times N)$ matrix composed of the values or 'scores' of each observation i on the minor common factors

NOTE . . . the other matrix variables are as defined above for eqs (5) and (6).

The correlation matrix being defined for eq (8) above becomes:

$$R = \mathrm{AA}' + BB' + U_n^2 + E_n^2 \tag{10}$$

Hakstian *et al.* (1982) state that ... "Use of this model would seem consistent with the view that in reality, the idealised traditional common factor model may be a poor representation for many sets of variables. In addition to the major simple structure factors structuring a domain of interest, the existence would be postulated of a vast number of real influences, which although contributing somewhat to the covariation among a set of variables, are of little consequence and of a random, unstructured form ..." (p.195). This covariation is subsequently reflected in the minor common factors. Thus the middle model defines the factor extraction problem as that of separating the major common factors from all other factors, including those that do account for some trivial commonality between variables.

While the reasoning behind this set of postulates is sound, their practical consequence is unfortunately somewhat ephemeral! Although Hakstian *et al.* propose a set of guidelines as an aid to factor extraction, in reality they add little to current knowledge and debate upon this issue. For example, the Scree test (for identifying the number of PCA components to extract) is based upon the slope of eigenvalues plotted against their extraction order. When the eigenvalues are successively plotted, a falling curved section followed by a straight line (or several) at a much lesser angle to the horizontal is observed. The resemblance of these straight line sections to the screes of rock debris at the base of a mountain led Cattell to propose the name 'scree' test.

Cattell (1966) and Cattell and Vogelmann (1977) present some theory for this test in addition to an extensive empirical test of the method vs the Kaiser–Guttman criterion. Cattell (1978) has suggested four rules for applying the test, stressing that the subjectivity of decision occurs in combining the rules and conditions. Unfortunately, these rules can quite easily be misconstrued by the user. The scree test is in essence a subjective test. In an attempt at reducing the level of subjectivity, Barrett and Kline (1982c) have recently introduced AUTOSCREE, a machine implemented scree test. However, a certain reduced level of subjectivity still remains.

The use of the Kaiser–Guttman (KG) test for PCA requires examination in that the simple logic of retaining factors with eigenvalues > 1 appears to be in error. Its use on ICA would conversely seem to be fully justified given the validity of a crucial assumption and the definition of a gramian matrix. Hakstian and Muller (1973) in their discussion of the KG note ... 'A widespread, but faulty, interpretation of the algebraic work (on bounds of the rank of a reduced correlation matrix) is that concerning Guttman's (1954) weaker lower bound, in which an elegant proof on the universally weakest lower bound to the rank of a gramian reduced correlation matrix, was translated into a rule of thumb concerned with the number of principal components to retain!' Accepting Hakstian and Muller's argument, the only realistic justification for using this test on PCA is Kaiser's (1960, 1965) logic based upon the coefficient α generalisability of factors. Defining coefficient α as:

$$\alpha = \frac{n}{n-1} \quad 1 - \frac{\sum_{i=1}^{n} V_i}{V_t} \tag{11}$$

where n is the number of variables in a test scale, V_t is the variance of the composite test scores, and V_i is the variance of the scores on individual variables i. Mulaik (1972) highlights a crucial assumption made by Kaiser in defining the variance of a test variable (p. 210) 'Since in most psychological problems the variance of a variable has no intrinsic meaning, we can assume that the variances of the n variables are arbitrarily set equal to 1.' This assumption reduces (11) to

$$= \frac{n}{n-1} \quad 1 - \frac{1}{\lambda_m} \tag{12}$$

where λ_m is the eigenvalue associated with a factor m. With this assumption, the test is applicable to both common and component factor analysis.

Hashemi (1981) examined some well known methods of factor extraction in his study of the factor structure of the 16PF test and EPQ. He found in this data and in other established examples from the literature of factor analysis, that two methods seemed to give reasonably consistent and reliable results — the Scree test (Cattell, 1966) and the Velicer test (Velicer, 1976b) which uses the partial intercorrelations between the variables to determine the significant factors. However, these did on occasion disagree and it appeared that each was the better solution in about half the disagreements. It seems sensible where doubt exists to rotate both suggested numbers of factors and to judge from the results which is the more efficient solution. There can be no doubt that ultimately this decision on how many factors to rotate is subjective and that the current methods of factor extraction should be used as guides rather than fixed and rigid rules that must be followed.

FINDING SIMPLE STRUCTURE

Given that our analysis has utilised the right number of factors, there still remains the problem of obtaining a simple structure solution. In practice there is now good agreement as to how this is to be obtained. Before this can be discussed, however, we have to decide whether the solution is to be orthogonal or oblique. We have discussed this fully elsewhere (Kline, 1979, Cattell and Kline, 1977) and shall here summarise the main points.

(1) If simple structure is defined as the position where factors have mainly zero loadings but each a few high loading variables, then oblique rotations are to be preferred since it is highly likely that such a position will necessitate some degree of obliquity among the vectors. If the oblique rotation is such that an orthogonal position can be chosen if it is the best fit (i.e. the vectors are placed regardless of their intercorrelations), then this would seem ideal. So strong is this very obvious argument that the majority of leading factor analysts in fact use oblique rotations in searching for simple structure.

(2) There is a further argument in favour of oblique rotations: one heavily stressed by Cattell (e.g. 1973) and particularly relevant in the field of personality. This is a rational, logical, point namely that it is unlikely, that in the real and complex world of personality, the main dimensions (especially if regarded, as does Cattell, as causal influences) would be uncorrelated. In the analagous field of abilities this is certainly not the case. It seems sensible, therefore, to hypothesise the factors as correlated until this is proved false by the evidence.

(3) Guilford (e.g. Guilford, 1959) takes a different view. He argues that while it is true that individual factors when rotated to the oblique position may be more simple than if an orthogonal rotation be used, nevertheless, overall, the most simple description is provided by a set of uncorrelated rather than correlated factors. That factors are correlated makes interpretation more complex than if they are not, even if individual factors are less well defined.

There is no definite answer to this point. It is a matter of scientific judgement as to whether the arguments for oblique or correlated factors are preferred. In our view the arguments of Cattell seem the more powerful and we accordingly prefer to see simple structure defined as does Thurstone, thus allowing obliquity where necessary.

There is also a further empirical point. If it could be shown that orthogonal dimensions make more psychological sense or correlated more highly with relevant external criteria than do oblique factors, then Guilford's argument would be supported. However, in the field of personality, where admittedly the evidence is complex and imprecise, this is not the case. In the ability field where the findings are better defined, however, the evidence certainly does not favour Guilford. There is little doubt that the second-order factors (which can only arise from correlated factors, of course) are better predictors than Guilford's orthogonal factors (Guilford and Hoepfner, 1971). Furthermore all abilities tend to be correlated as Thurstone (e.g. 1948) found — hence his positive manifold. Indeed Guilford has produced second order factors, based upon correlation between his ability tests rather than his factors — a procedure which we regard as inherently self-contradictory, having one's cake and eating it.

In summary, then, we see little reason to support Guilford's claims concerning orthogonal factors. We still argue that it is essential to rotate to simple structure, allowing factors to take up the oblique position where necessary.

Methods of Rotation

It has been forcefully argued by Cattell (Cattell, 1973; Cattell, 1978; Cattell and Kline, 1977) that a common cause, perhaps indeed the most common, of failing to obtain simple structure lies in the method of rotation.

Over the years there have been a considerable number of different rotational procedures developed, each with their advocates. It would be inappropriate to discuss this problem here in detail (a full discussion can be found in Cattell, 1978) if only because there is now general agreement over how rotation is best carried out, and it is possible to summarise these conclusions.

Simple structure is most usually reached by rotations which maximise the hyperplane count, in essence the number of zero loadings. Hakstian (1971) in a study of various oblique rotational procedures showed that Direct Oblimin (Jennrich and Sampson, 1966) in most conditions was about the most efficient programme although the Harris–Kaiser methods are also powerful (Harris and Kaiser, 1964). Promax (Hendrickson and White, 1964) can be efficient when the simple structure is not too far from orthogonality (for it starts from an orthogonal rotational position) but it must be noted, this does not maximise a hyperplane count.

In brief, an oblique rotational procedure which maximises the hyperplane count of principal factors (or components in a large matrix), selected as significant by the Scree test or other procedure which does not overestimate the number of factors, given a proper sampling of variables and subjects, should yield a replicable simple factor structure. Factors thus revealed should be the main sources of variance in the field.

The above prescription constitutes an effective factor-analytic method which most factorists would agree as efficient. Certainly it is very close to the procedure advocated by Carroll (1981) in his attempts to reanalyse the factor-analytic work on human abilities.

Before we leave the topic of technically adequate factor-analyses a few further points need to be made, although the essentials have been examined. It is probably sensible to check whether simple structure has been obtained and Bargmann (1954) has a method of assessing this, a method which has been further extended by Kameoka and Sine (1978). It is also useful to have a statistical check for factor similarity especially where factors are compared from researches with small samples, when replication of factors is particularly important. Cattell's r_p, the profile similarity coefficient is a possibility here (*vid* Cattell, 1973).

Our reasons for thus delineating technically adequate factor analytic procedures are not academic or even pedagogic. Rather our argument will be that only factors emerging from adequately conducted studies merit consideration. We shall not use these guidelines with obsessional rigidity. Certain deviations, e.g. in sample size, especially when there are replicated findings are not too serious. Poor rotational procedures, however, do usually infirm results.

Before we begin to examine the substantive findings which have arisen from the factor analysis of personality questionnaires, there are certain other methodological issues which so frequently arise in the critical evaluation of research (as distinct from the mere cataloguing of results), that it is better to discuss these first and save needless repetition.

OTHER METHODOLOGICAL ISSUES

Rotation to Target Matrices

Virtually all that we have said about factor analysis is most closely relevant to the use of factor analysis as an exploratory or mapping technique. However, as Eysenck (e.g. 1967) has stressed, it is possible and useful to test hypotheses by way of factor-analysis. Guilford in his studies of human abilities (e.g. Guilford and Hoepfner, 1971) is perhaps the best known exponent of this approach, although the senior author has often used this method in his studies of psychoanalytic personality theory (e.g. Kline and Storey, 1977). Guilford hypothesises the loadings that variables should have on various factors and seeks to approach this pattern as closely as his data allow, using rotational algorithms known as Procrustes solutions. Such analyses which did indeed fit his three faceted model of intellect form the evidential basis for his work on abilities.

Procrustes solutions owe their name to a mythical highway bandit of ancient Greek mythology who tied his victims to an iron bed and stretched or cut off their legs to make them fit its length. The term Procrustean is applied to any harsh or inflexible attempt to force someone or something to fit some preconceived idea or system. Hurley and Cattell (1962) are generally credited with introducing the term Procrustes into factor-analytic problems. A Procrustean transformation is any linear transformation which, given certain specified constraints, seeks to transform a given matrix into a matrix as much as possible like some hypothesised matrix. Procrustean transformations can be used in factor analysis when, instead of trying to rotate factors to optimise some general, abstract criterion such as simple structure, we try to rotate to obtain factors possessing properties as similar as possible to our target (hypothesised) matrix. The elements of the target matrix may be specified exhaustively, (that is all elements are given a target value) or only partially, where the fitting of the transformed matrix to the target matrix is only in terms of the specified elements. In addition, care has to be taken over positing an oblique or orthogonal solution. Finally, the actual values of the hypothesised elements have to be chosen with care. Values may range from a simple classification of: 1 for a positive value, 0 for no value, and −1 for a negative value, to a more numerically exact solution. The former type of target matrix specification is obviously the easiest to fit.

However, these Procrustes rotations to loosely defined target matrices can no longer be regarded as sound evidence for or against hypotheses. This is simply because Horn and Knapp (1973) demonstrated that Procrustes methods were so powerful in attaining target matrices that it was doubtful whether they could reject any hypothesis at all. In their paper they showed that Procrustes rotations could match hypothesised matrices from random data and from data with a structure built into them that was antithetical to the target matrix. Recently Guilford (1974) has attempted to answer this criticism but as was pointed out in their reply (Horn and Knapp, 1974) his case is not reinstated. The Procrustes methodology is too powerful to make hypothesis rejection at all likely, unless the hypothesis can be stated with extreme precision. This effectively means that in the field of personality, where such precision is virtually impossible, Procrustes is ruled out.

Even were this not the case, in our view there is further argument against targeted rotational procedures which makes their use dubious. Even if hypothetical matrices were well matched, unless this hypothetical factor pattern happened to be the simple structure, we would have little confidence in it. If it were simple structure then, of course by definition, that could have been reached with our advocated methods. Our lack of confidence stems from the fact that, as we have argued, the simple structure position is chosen as the most parsimonious solution. On these grounds all others would be rejected. Since personality theory is not precise we can not prefer a solution derived from a dubious theoretical rationale to a parsimonious solution. That is why in our studies of psychoanalytic theory through factor analysis we have compared the theoretical prediction with the simple structure resolution. This argument, of course, means that the attempt to match matrices through targeted solutions can never be worthwhile in itself, even if less powerful techinques than Procrustes are available.

In brief, therefore, we are happy to test hypotheses using factor analyses but prefer to do so using simple structure. There seems no case to be made for trying to reach a target factor matrix.

CONFIRMATORY FACTOR ANALYSIS

At this juncture we must now introduce what many psychometric psychologists, especially in the continent of Europe, regard as the best technique of factor-analysis (e.g. Verster, 1981). This is Maximum Likelihood Confirmatory Factor Analysis (Joreskog and Lawley 1968, Joreskog, 1969). Basically, this procedure is one whereby a set of factor analytic parameters may be specified in advance of the solution, the remaining parameters then estimated by maximum likelihood methods, and finally the actual solution statistically compared with the hypothesised solution in order to determine the goodness of fit. The parameter specification will involve the number of factors, the correlation between factors (if any), the size and direction of individual loadings, and the size and direction of individual variable specific factor loadings. Depending upon the quantity, values, and positions of the specified parameters either in a common factor loading, factor correlation, or specific factor loading matrix, two kinds of solution can be sought: unrestricted and restricted. An unrestricted solution is one that does not restrict the common factor space. By specifying very few fixed common factor loadings or factor correlations, the solution obtained is generally unrestricted by these fixed elements. In an unrestricted solution, no specific factor loadings can be held fixed, since clearly a restriction on the specific factor loading matrix is a restriction of the common factor space. All unrestricted solutions can be obtained by a rotation of an arbitrary unrestricted orthogonal maximum likelihood solution. Generally, all unrestricted solutions are no more than exploratory factor analyses, with simply a little bias introduced to 'guide' the solution towards a very loosely defined target model. A restricted solution, however, imposes restrictions on the whole factor space, and such a solution cannot be obtained by a rotation of an unrestricted solution. This form of solution is obviously the most stringent as regards model fitting; both common and specific factor space is constrained toward a specified solution.

However, as Nunnally (1978) points out, there is a problem with the use of the χ^2 statistical test used within confirmatory analysis. With large samples and matrices it is extremely unlikely to reject similarly specified models such that a plethora of model structures may be fitted 'successfully', requiring the judgement of the investigator to choose the 'best' model: this judgemental decision being based upon the theoretical and practical consequences of accepting any individual model. However, it is clear that to stop fitting additional parameters cannot be decided on a purely statistical basis.

So far, in the field of personality questionnaires, indeed of personality measurement generally, confirmatory analysis has been rarely used so that it is speculative to argue that it will be a better method than the exploratory procedures described above. This is especially true of this field on account of the problem of setting up a theoretically powerful target matrix. So despite the apparent statistical advantages of confirmatory analysis over older methods, we do not consider that confirmatory analysis must always be used. Probably a good application would be to replicate a simple structure solution using confirmatory methods.

The Inter-Item Correlation Coefficient and other Problems in Factoring Items

The factor analysis of personality questionnaires can begin at the item or the scale level. The problem with factoring scales is that second-order factors are likely to emerge. If the scales are truly factorially distinct, they may possibly emerge as factors, each loading on a separate factor. With one scale per factor, this will be a poor solution with factors of little variance — no better than the original scales. If the scales of a questionnaire are examined together with those from other questionnaires, the solution could be more clear. However, in practice, second-order factors accounting for scale intercorrelations are likely to occur. For these reasons the proper analysis of factors in personality questionnaires demands that items rather than scales be factored. However, the factor analysis of item intercorrelations is beset with statistical problems capable of producing considerable error. These we shall briefly discuss.

(a) The dichotomous item

Many personality questionnaires use dichotomous items and the difficulty here lies in the fact that the relevant correlational coefficients are subject to various forms of bias, a serious defect when factor analysis is required. First as Nunnally (1978) points out, tetrachoric coefficients should not really be factored for they are not scalar-products. In any case as Guilford (1956) stresses tetrachoric coefficients have large standard errors (about twice as much as the product moment correlation). This, together with the fact that the correlations are affected by the size of the item splits makes them unsuited to factor analysis. In the past their use has been permissible only because it saved calculating time. There is no excuse today for factoring a matrix of tetrachoric correlations in the study of test items.

Many workers, therefore, use the ϕ coefficient which, unlike the tetrachoric correlation, is designed for non continuous data. Thus strictly to use it in the analysis of personality questionnaires we have to assume that the two item responses are discreet. This is not too serious. However, like the tetrachoric correlation ϕ is affected by item splits and this tends to distort any subsequent factor-analyses.

Thus one reason why the factor analysis of inter-item correlation matrices tends to yield factors of small variance (Cattell, 1973) is the error variance artefactual to these correlation indices.

(b) Other item forms

One possible method of overcoming this problem is to construct items with response formats allowing more variation, e.g. nine-point scales. Comrey

(1970) with his Comrey Personality Inventory has done this and although there are other problems with such rating scales (see Vernon, 1963 for a discussion of these difficulties), this is a viable method for it allows a proper use of the Pearson coefficient. It should be noted apropos of this that three or five point scales are not really long enough to overcome the statistical problems of the Pearson correlation coefficient, so that scales this short still are likely to be compounded with error.

However, since nine-point answer scales are clumsy and make good item writing difficult and encourage response sets of putting the extreme or the middle responses, other ways of overcoming these correlational difficulties have been attempted.

(c) Grouping items

A half way house between factoring scales and factoring items which goes some way to overcoming the problems which we have been examining, is to use as a basis for correlation, scores derived from a group of items. Comrey employs such a method (factored homogeneous item dimensions) in his Inventory. FHID's consist of clusters of items who have been shown to load on the same factor and were conceived to be measures of the same variable. This seems to be a viable method although Cattell has argued that such FHID's are virtually brief scales and thus second-order factors are likely to emerge (Cattell, 1973).

Yet another approach open to the same objection is to use clusters of items with intercorrelations of a given size or more as a basis for the correlation matrix. Cattell (e.g. 1973) has a special clustering method — radial parcel analysis involving the parcelling of items with large inter-item vector cosines in the item factor space. Cattell (1973) in an extensive discussion of this whole problem argues that essentially item factoring and parcel factoring yield similar results but the latter tends to be more replicable, containing, as it does, less error.

Clearly in assessing the results of any studies it is necessary to take into account this aspect of the method — whether items or groups of items were initially factored. Certainly here are abundant sources of error.

Such are the problems involved in the factor analytic study of personality questionnaires. We shall now turn to an examination of results, in studies which have been properly conducted.

THE FACTORS IN FACTORED QUESTIONNAIRES

In our scrutiny of personality factors, it will not be possible or sensible to take into consideration all the factored tests either those published or those described in research articles. Our criteria for the selection of tests are such that good tests are bound to be included.

(1) All measures where the standards of test construction and factor analytic methodology are satisfactory will be included.

(2) All measures where there is good evidence for the validity of the factored variables will be included. This will be done even where the first condition may not be met since this first condition has been introduced only to maximise the possibility of producing valid tests. Needless to say tests meeting only this second condition are few in number.

It is to be noted that this first section only includes factor-analytic tests. We shall examine the factors in tests developed through criterion keying in the next section, since it is logically probable (and empirically the case) that factors in the tests will have been previously described in our study of factored tests.

Before we attempt to set out our list of factors in factored tests, it is necessary to discuss the research design (not the detailed factor analytic method) that needs to be used, if factors are to be identified.

The Research Design for the Identification of Factors

It is necessary to locate factors in factor space relative to each other. Without this proper identification it is not possible. This means that in any adequately designed study marker factors must be included. These markers should be either the best known and accepted factors in the field, and/or other factors thought to be related to those under investigation. This enables comparison across researches to be made and avoids the obvious difficulty that unless this is done, each factor as it emerges from a study, tends to be labelled idiosyncratically by its author or identified as another factor, but of course without evidence. All this is so obvious and banal that readers may require apology for mentioning it. However, in my survey of the psychometric field for its psychological implications (Kline, 1979) it turned out that few factor analysts had actually done this, thus rendering their work of little value.

In practice this means that those factor analyses of individual tests without attempt to locate factors in factor space are of little value in elucidating the meaning of the factors in personality questionnaires. We shall barely consider any such in this paper. At best they can show that a dimension or dimensions underlies a set of items. They cannot show what this is and, of course, it could be a specific to the test itself, a response set such as acquiescence or social desirability for example. Even if it were a factor of more generality, our only means of identification would reside in the item content. This is only face validity and face validity is a notoriously poor indicator of real validity in personality questionnaires (Cronbach, 1970).

Thus most of the research to be discussed in the next section of this paper will concern researches in which factors can be identified by reference to other factors. Of course, the analysis of tests with more than one scale to some extent does this and these need not necessarily be related, for purposes of identification.

Two factored tests stand before all others in the richness of the data that has been accumulated over the years in respect of the meaning of the factors. These are the Eysenck scales, of which the most recent is the EPQ (Eysenck and Eysenck 1975) and the Cattell scales, of which the 16 PF, the adult version, is the most widely used. First we shall examine the factors in the Eysenck Scales.

Factors in the EPQ

The EPQ (1975) is the latest of a long list of personality scales developed by Eysenck and his colleagues at the Maudsley. It purports to measure three factors, extraversion, neuroticism and psychoticism. It differs from the earlier versions (the EPI, Eysenck and Eysenck, 1964) and the Maudsley Personality Inventory (Eysenck, 1956) mainly in the fact that it measures psychoticism in addition to E and N, and in that most of the impulsivity items have been removed from the extraversion scale (Eysenck and Eysenck, 1975). The even earlier Maudsley Medical Questionnaire measured only N. Although most of the research into the psychological meaning of E and N was actually carried out with the earlier versions of Eysenck's tests, we shall assume, as does Eysenck in the EPQ manual (Eysenck and Eysenck, 1975), that the E and N of this test are identical with the previous factors, although, as the reviewers of the EPQ in Buros (1978) (including the present author) were quick to point out, there was a lack of clear evidence on this point. Recently, however, the simple correlation of the EPI and EPQ N and E factors has been reported * and given the reliability of the tests, the assumption of identity is well founded. This is, perhaps, slightly surprising in view of the lack of impulsivity items in the E scale of the EPQ, but there can be no gainsaying the result. Thus the EPQ factors, we argue, represent Eysenck's factors.

The questions concerning the factors in the EPQ can be listed.

1. How many factors account for the EPQ test variance?
2. What are these factors? Are they those expected from the rationale of the test construction?
3. Are they first-order or higher-order factors?
4. How do they relate to the factors in other tests?

*Rocklin, T. and Revelle, W. (1981)

These are, of course, not entirely separate questions and the results which we shall discuss will bear on several if not all of them.

Before examining the factors that have emerged from the properly conducted factorial studies of the EPQ, and to aid understanding of the results it will be helpful to set these questions within their context.

Eysenck (e.g. 1967) has, of course, done far more than develop a series of factored personality tests. These are but a small part of a comprehensive theory of personality. In this theory, extraversion reflects the arousability of the central nervous system, hence its claimed relationship to conditionability; neuroticism, on the other hand, reflects the lability of the autonomic nervous system, while psychoticism, with less confidence (Eysenck and Eysenck, 1976) is held to reflect the androgen level of the individual. As might be expected with variables so firmly rooted in physiology, there is a considerable genetic determination of the population variance. (Eysenck and Eysenck, 1976; Eaves and Eysenck, 1975).

Since the EPQ was specifically designed to measure these variables, together with *L*, a measure of the tendency to endorse the socially desirable response, it is clear that we should expect the items to load on four clear factors with each item, as specified in the test manual loading on one factor alone. Thus a factor analysis of the items in the EPQ, should yield unequivocal results. This is the context of the first two questions.

The question as to whether *E*, *N* and *P* are first or higher-order factors is closely relevant to the rival theories of Cattell (e.g. Cattell, 1973; Cattell and Kline, 1977) and Eysenck. This is a highly complex argument which in the main is not relevant to the specific aims of this paper except that as part of it Cattell asserts that the importance of *E*, *N* (and to a lesser extent *P*) is not denied. However, for Cattell these are higher-order factors accounting for the variance between the primary factors of which he names 23 (Cattell and Kline, 1977). To measure as Eysenck does, it is argued, only the higher-orders, is crude, since for each individual, the factorial composition of his higher-order factor scores differs and this difference has important psychological implications: what Cattell refers to, as depth psychometry (Cattell and Kline, 1977). Eysenck has overemphasized the significance of the higher-order factors by under-factoring (see our discussion of simple structure) which tends to produce higher-order factors at the first order. Eysenck (e.g. 1978) has always asserted in defence that neat as the argument is, the primary factors are too unstable to allow for effective measurement. Thus, study of the EPQ items is highly relevant to this issue as is, and even more so, study of the Cattell item sets.

Finally since almost all factor analysts of any note (Eysenck, 1967; Cattell, 1973; Comrey, 1970 and Guilford, 1959) agree that *E* and *N* are important personality factors, the relation of these to other putative factors obviously deserves careful empirical investigation.

It is in the context of these points that we shall now examine the studies of the factors in the EPQ.

Studies of the EPQ Items Alone

Jackson and Paunonen (1980) in their lengthy survey of recent work in personality assessment conclude that "the data bespeak the difficulty of attempting to relate EPI and EPQ test responses to a general underlying introversion–extraversion construct. Research by Guilford (1977) and Loo (1979) is cited, *inter alia,* as clear support for their gloomy conclusions and they could well have invoked a paper by Helmes (1980) which similarly fails to isolate the *E* dimension.

However, we do not accept their conclusions. We shall first demonstrate that the work by Loo and Helmes is flawed by methodological errors. Then we shall present results which, in our view, demonstrate beyond reasonable doubt that the EPQ does indeed measure three factors, as intended by its authors.

The Work of Loo (1979)

Loo factored the responses of 262 students to the EPQ. Hierarchical factor analyses failed to recover the *E* and *N* items while the *E* items were recovered in a secondary factor — "social extraversion without an impulsivity primary factor." (p.54). At its face value, therefore, Jackson and Paunonen (1980) were correct to argue that this paper failed to confirm the status of *N, E* and *P.*

However, had these authors attempted to evaluate this paper rather than simply report it, they would have discovered the following severe methodological problems which effectively infirm the findings, as Barrett and Kline (1982a) have previously pointed out.

(1) Principal Components rather than principal factors were extracted, although in a large matrix this may not have been too serious.

(2) No rationale is given for extracting significant factors. Sixteen were rotated at the first order and this seems to be a case of over-factoring, since nine factors had only 5 items loading >0.29 on them and one had only 3 items which must be regarded as a specific (items 31, 41 and 75) these being respectively 'Would you call yourself a nervous person?', 'Would you call yourself highly strung?' and 'Do you suffer from nerves?'. A factor analysis which produces such a factor is not normally considered adequate.

(3) It is not specified whether the higher-order rotations were grounded or non-grounded. Without a grounding procedure accurate rotation is difficult.

(4) No mention is made of the correlations between factors.

(5) Although the higher-orders were oblique, it was an orthogonal varimax solution which was interpreted. This is curious at the least.

(6) Associated with overfactoring there was misleading labelling of first-order factors, for example, impulsivity loading on items concerned with hurtful jokes (33), being annoyed by careful drivers (46), and being in debt.

(7) The claim that extraversion is not concerned with impulsivity is hardly surprising since Eysenck deliberately left out these items from the *E* scale.

This study as reported simply fails to meet the criteria for adequate factor analyses. As such its results cannot be taken seriously. Since this study by Loo (1979) is one of the main sources of evidence for the argument that the EPQ does not measure introversion–extraversion, and since we have shown it to be of little value, it means that the support for this claim rests largely on the work of Helmes (1980) which we shall now scrutinise.

The Work of Helmes (1980)

Helmes carried out an extremely detailed item analysis and higher order factor analysis of the EPQ from the responses of 191 Canadian undergraduates. The item analysis (biserial item total correlations and the proportions putting the keyed response to each item) was not promising for the validity of the EPQ. Thus the mean item total correlations for the scales were low (e.g. -0.22 for the *L* scale and 0.08 for the *P* scale) and in addition the *P* scale items were poor discriminators in this sample (as evidenced by extreme response splits) and by the correlation of the *P* scale with the social desirability measure from Jackson's (1974) Personality Research Form.

Obviously, given this item analysis, the factor analyses of the inter-item correlations were unsatisfactory as regards the validity of the EPQ's four dimensions. At neither the first, second, or third order did the four expected factors emerge and, even worse, many items failed to load an interpretable factor. Helmes (1980) agreed with Loo (1974) concluding that ... "it is arguable whether the EPQ originally had the structure claimed for it." This study would appear, therefore, to cast doubt on the importance of *E*, *N* and *P*, as dimensions of personality.

However, examination of the factor-analytic procedures used by Helmes demonstrates that they are by no means consonant with the methods which we have previously suggested to be those agreed upon as most likely to lead to a replicable simple structure. Some of the problems in the two analyses used by Helmes are briefly set out below. (See Barrett and Kline, 1982a, for full details.)

Analysis 1

Four principal components were extracted and rotated using Varimax and Promax procedures. No rationale was used to hit upon the number of factors rotated (other than Eysenckian theory): this is of course critical in reaching simple structure. Both Varimax and Promax have no check on the hyperplane count or whether simple structure was in fact reached. Thus we are not surprised that this part of the investigation failed to replicate the results. It could well be the case that, for example, if six principal components were rotated, four clear factors would have emerged. In addition, Helmes tried to reach the Eysenckian position by using a Procrustes rotation procedure. We have already pointed out the problems with Procrustes rotations (see p. 154) that they can all too easily find their target. Hence little reliance can be placed on these findings, although in fact in this instance the Eysenckian position was not supported.

Analysis 2

Aware of the problems of under-factoring, Helmes, in his second analysis, subjected all principal components with eigenvalues greater than 1 to a promax analysis and then factored the intercorrelations between these factors — second order analysis. The items were then projected on to these second-orders. This was repeated to obtain third order factors, Helmes claiming that he was replicating the methods used by Eysenck and Eysenck (1969) who did obtain from the highly similar EPI items clear *N* and *E* factors, this earlier version of course having no *P* factor.

The criticisms of this part of the study are:

(a) On a point of fact, this is not what the Eysencks did. The projection of variables on to higher order factors after rotation is the method advocated by Cattell and White (1962). It suffers from the disadvantage that higher-order factors have to be located against the lower order factors of which there are insufficient to allow accurate location. Eysenck and Eysenck (1969) in fact, used the Hendrickson and White method (1966) which allows higher-order factors to be located against the original variables. This enables better simple structure and higher hyperplane counts to be obtained. Thus the failure of Helmes to replicate the Eysencks' results in this second-order analysis could easily be attributable to, if only in part, this different method of higher-order factoring. His Cattell-White approach is almost certainly inferior to the procedure used by the Eysencks.

(b) There is another discrepancy from Eysenck's method and this again is not of only pedantic interest, but is one likely to affect the results to a considerable degree. This concerns the number of factors in the rotation. As we have discussed previously, this is a critical point in obtaining simple structure: rotating too many factors tends to split factors, thus destroying the structure; rotating too few factors tending to produce higher-order factors at the first order. Eysenck and Eysenck (1969) rotated 20 factors, an arbitrary figure but one which they argued would allow the Cattell, Guilford and

Eysenck primaries to emerge if they were present in their data. Helmes, in contradistinction rotated all factors with eigenvalues greater than one.

From this alone it is clear that Helmes' work is not a replication of that of the Eysencks. These discrepancies could account for the differences between his findings and the original study. As we argued; following Cattell (1973) some differences in results between factor analysts are attributable simply to different methods.

(c) This second discrepancy raises a further issue. Given that Helmes (1980) did not, despite his claims, follow Eysenck's procedures, it could be the case that his methods were more likely than those of Eysenck to reach simple structure, implying that this study should be given the closest consideration. However, as we have fully discussed in our section on the number of factors to rotate the K–G criterion of eigenvalues greater than one is not satisfactory especially where, as is the case in the factor analysis of items, there is a large number of small factors.

(d) A final point remains. Helmes attempted to reach the Eysenck factors using a Procrustes rotation. As we have fully discussed, Procrustes tends always to reach its target so that results confirmed by this procedure cannot be used to substantiate hypotheses. In its place confirmatory analysis should have been used. However, since even the Procrustes method failed to reveal the structure, it certainly cannot be argued that a more rigorous test would have done so. It does appear that in this sample, at least, *E, N* and *P* did not emerge.

For these technical reasons, therefore, it seems to us that no great weight can be placed upon these results. Certainly Helmes (1980) failed to obtain the factors that would be hypothesised on Eysenckian theory. This, however, as we have shown is not unexpected.

Before leaving this study, one further point deserves mention. It is odd if Eysenck's factors are as clear as he maintains that the item analyses of Helmes failed to reveal common factors running through the item sets. It suggests, as do Helmes' factorial results, in this sample, for reasons unclear, that the EPQ was not measuring these factors efficiently, a conclusion strengthened by the failure of the Procrustes rotation.

The Work of Barrett and Kline (1980)

As part of a large-scale investigation of factors in personality questionnaires, Barrett and Kline (1980b) carried out an extensive factorial analysis of the EPQ items, utilising the methods which we have suggested are most likely to reveal any underlying simple structure.

Samples

Two samples were used: (a) A Gallup adult quota sample (600 males and 598 females) data lent to us for analysis by H. J. Eysenck; (b) 406 University undergraduates (171 males, and 235 females). These subjects were formed into 6 groups, total students, male students, female students, total adults, male adults and female adults and separate factor analyses of the inter-item correlations were computed.

Method

For each group, the ϕ-coefficients between the EPQ items were computed and before factoring, measures of sampling adequacy were obtained (e.g. Cerny and Kaiser, 1977). Principal component analysis was carried out and using both the Scree test (Cattell, 1966) and the K–G criterion, the appropriate number of factors were rotated using the direct oblimin procedure (Jennrich and Sampson, 1966). The hyperplane count determined the position to stop rotation. Full details can be found in Barrett and Kline (1980). Higher-order factoring of the intercorrelations between the primary factors was carried out using the same procedures after the factors had been grounded on the items using the Hendrickson and White method. The significance of the higher-order simple structure was also tested using the Bargmann (1955) test and the tables of Kameoka and Sine (1978) to be found in Cattell (1978). This method seems approximate to those likely to reach simple structure and is difficult to impugn on technical grounds unless it is agreed that only confirmatory analysis can truly test hypotheses. All the factors in the six groups were compared using Pearson correlations and Tucker congruence coefficients. Coefficients >0.75 were regarded as demonstrating an identity between factors.

For the detailed results of the 6 sets of primary and second-order factors, readers must be referred to the original paper. Nevertheless, so clear were the results that they can be easily summarised.

Table 1 indicates that at the second-order in the large adult samples there was virtually perfect recovery of the *N, E* and *P* items. *P* only failed in the females. *N* and *E* were also excellent with the student groups, although *P* was not as successful. The relative failure of *P* may be due to the low subject to variable ratio in the student samples.

However, it is to be noted in discussing the performance of the *P* items that among the students there was a very low mean with the consequent low variance. Student samples do not show (or admit to) *P* behaviours which is hardly surprising since these are found highest in psychotic and criminal groups. It is noteworthy too that the only adult group where *P* failed (the females) also had a low mean and restricted variance.

Table 1. Numbers of E, N, P and L items Loading on the Second Order Factors

Sample	E	N	L	P
Total students	21_1	23_2	17_4	14_3
Male students	18_1	19_2	13_3	13_4
Female students	19_1	18_2	L/P ITEMS	N/L ITEMS
Total Gallup	20_2	23_1	20_4	22_3
Male Gallup	20_2	22_1	20_4	23_3
Female Gallup	18_2	22_1	16_3	13_4

Where subscript denotes factor extraction order. No. of E scale items = 21 No. of N scale items = 23 No. of L scale items = 21 No. of P scale items = 25 Table 1 Reprinted from Barrett and Kline (1980b) with the permission of Pergamon Press.

Nevertheless despite these comments, if we look at the large adult sample it is clear that at the second-order *N, E* and *P* are factors which emerge clear-cut from the EPQ. If we wish to be harshly critical of this test we should argue that more *P* items are necessary of a kind suited to normal populations, to increase the variance of the scale among them. Thus where simple structure has been obtained on a large sample of adults, *N, E* and *P* are clear. There can be little doubt that these three variables can be put on a list of questionnaire factors.

At the first order the results, naturally enough, were not as neat but nevertheless deserve brief mention. Two factors were common to all six groups at the first order — moody irritability and anxious worrying — a split of the *N* factor. In five of the groups extraversion and social desirability were common and in the student male group extraversion was replaced by the more specific sociability. This extraversion factor at the first order was particularly interesting because almost the whole *E* scale loaded on it, thus supporting Eysenck and Eysenck's (1968) claim that *E* can be found at any order of factor analysis, although here, N did split into two at the first order.

Conclusions from this study

There can be little doubt that this study is powerful support for Eysenck's claims concerning the main personality factors: *N, E* and *P* do emerge with impressive clarity.

In our view there is little point in reviewing other work on the EPI and EPQ for the following reasons:

(1) The EPI *E* and *N* are essentially the same as the EPQ *E* and *N* (as the work of Barrett & Kline (1982a) shows), despite the loss of the impulsivity items from the *E* scale of the EPQ. Thus all the EPI findings apply to the EPQ and vice versa.

(2) The methods of the study by Barret and Kline, together with the large

adult sample and the similar student results, make their findings hard to impugn. This is in contradistinction to the methodologically weak studies of Helmes and Loo.

(3) Indeed we only mentioned the work of Helmes and Loo because it was argued in the influential Annual Review of Psychology (Jackson and Paunonen, 1980) that these studies cast doubt on the concept of extraversion as a factor.

However, one issue still remains to be discussed: this pertains to the nature of extraversion. This is particularly pertinent because there is a change in the E scale item content in the EPI and EPQ, and because of work by Guilford (1977) and Eysenck and Eysenck (1977, 1978).

Impulsivity, Sensation Seeking and the Nature of Extraversion

In the EPI, as we have pointed out, extraversion consists of items pertaining to sociability and impulsiveness. As Eysenck and Eysenck (1977) argue this makes good sense since these same authors (Eysenck and Eysenck, 1963) demonstrated that sets of impulsivity and sociability items correlated well (0.47). Carrigan (1960) in her study of extraversion concluded that at that time the undimensionality of extraversion was not convincingly demonstrated, a point strongly made by Guilford (1975, 1977) who claims that extraversion is no more than an illegitimate composite of his R and S factors (variables which we discuss in a later section of this monograph) R being rhathymia (carefree and lively), highly similar to impulsivity, and S being sociability. However, if extraversion and sociability are thus closely related, it is pertinent to ask why the EPQ extraversion scale omits the impulsivity items.

Eysenck and Eysenck (1977) administered sets of impulsivity items and factored the inter-item correlations. The resulting factors were then correlated with the EPQ variables. Impulsiveness breaks down into four factors: narrow impulsiveness, risk taking, non planning and liveliness. These factors were positively but to a small extent (range 0.01 to 0.37, four correlations being beyond 0.3 and four being between 0.2 and 0.29) correlated together. Only liveliness is positively correlated (to a significant degree) with sociability. However, the four sub-factors of impulsivity are all significantly and positively correlated with extraversion but narrow impulsiveness, risk taking and non-planning are more highly correlated with P. Liveliness is the exception here. It is this fact that led the Eysencks to drop impulsivity from the EPQ, because these items would then load on both E and P, "producing positive correlations between what are otherwise orthogonal factors". The Eysencks in concluding this paper argue that both sides of the original controversy were correct. The general factor of impulsivity is related to extraversion and sociability but the narrow impulsivity is far less related to extraversion and not to sociability. Carrigan (1960) and Guilford (1975) were referring to narrow impulsivity, Eysenck was referring to the more general broad factor.

There are several points in this paper worthy of discussion. That impulsiveness breaks into four factors (impulsiveness a former part of extraversion) of course supports the claims of Cattell (1973; Cattell and Kline, 1977) that second-order factors are too broad and that primaries should be examined. The important differences, stressed by the Eysencks concerning the two impulsivities, confirm this view.

Not too much should be made of the differential correlations in this study of variables with extraversion and sociability. This is because the sociability scale consisted of the EPQ sociability item from the *E* scale and correlates with *E*, 0.94. Given the reliabilities of the two scales, these measures must be virtually identical.

The impulsivity items were left out of the EPQ *E* scale because they correlated with both extraversion and psychoticism. However, to do this because it would make otherwise orthogonal factors correlated is curious, indeed reminiscent of Spearman's attempts to purify tests of g. Obviously if variables causing factors to correlate are removed, they will remain orthogonal. Whether they should be regarded as orthogonal when considering their nature is another matter. It seems to us, given that impulsivity items load on both factors that we should regard *E* and *P* as correlated.

In brief this study indicates clearly that impulsivity in its broad sense is related to extraversion (but even more to psychoticism). However, clearly, impulsivity is not a genuinely unitary factor and it is best considered in terms of four oblique primaries. Guilford's (1975 and 1977) claim that it was separate can only apply to a subset of items at best — the narrow impulsivity. Guilford's (1977) paper, therefore, requires no especial discussion, except to note that the rotation to a 'better structure' (Guilford, p. 414) was only a graphical rotation by eye and there was no serious attempt to reach simple structure.

In respect of the light this study throws on extraversion, it seems to us to indicate that impulsivity is to some extent, but not in its entirety, a part of extraversion. However, it is also an aspect of psychoticism, thus suggesting a modest positive correlation between these factors.

The study reported above, as we have argued, supports the notion that three higher-order factors, however ubiquitous, are too broad to define personality in any detail. As we have seen, primary factors of impulsiveness are important. Eysenck and Eysenck (1978) went further along this Cattellian road in a study of impulsiveness and venturesomeness in which 402 male and 787 female adult subjects were administered a 63 item questionnaire (together with the EPQ so that resulting factors could be located relative to E, N and P — an impeccable research design by the criteria adopted in this monograph).

Three factors of venturesomeness (*v*), impulsivity (*im*) and empathy (*Em*) emerged. *V* had positive correlations with *P*, and negative with *N*. *Im* had positive correlations with *E*, *P* and *N*. *Em* had positive correlations with *N*. These three primaries were independent of each other.

It hardly needs pointing out that there is a contradiction in this study and that reported previously (Eysenck and Eysenck, 1977) where impulsivity actually broke down into four factors, while here there is only one. Even if we ignore this discrepancy, the position now remains that Eysenck has identified in addition to his three second-order factors of *N*, *E* and *P*, at least one (and probably three more) impulsivity factor, and empathy factor and a venturesomeness factor. The utility of these concepts, as Eysenck and Eysenck (1978) are careful to point out is as yet unknown. So too we should add, is their stability; it is to be hoped that they replicate more easily than the primary factors of Cattell, for as Eysenck and Eysenck (1969) once argued the problem with primary factors lies in this instability which makes their practical and theoretical value dubious.

One further point about this study deserves careful note. Eysenck and Zuckerman (1978) have also shown that sensation-seeking is a primary factor that is positively correlated to Extraversion. However, this well known Zuckerman scale is not yet a further primary factor. The *V* factor of Eysenck and Eysenck (1978) is in fact the best possible representation of this factor based upon a factorial study of the Zuckerman items with their formats changed together with scales of risk taking and impulsivity. In fact, two factors impulsivity and venturesomeness emerged, the latter loaded on most of the Zuckerman items.

The Eysenckian position can now be easily summarised.

(1) There is no doubt that the EPQ contains three broad second-order factors, Neuroticism, Extraversion and Psychoticism. These emerge with great clarity in studies both of scales and items.

(2) Recently, however, a number of primary factors has been subjected to investigation by Eysenck and his colleagues. This has resulted in the emergence of: (a) an empathy factor; (b) a venturesomeness or sensation-seeking factor; and (c) at least one impulsivity factor.

There is no reason to suspect that this is the complete list. For Eysenck and his colleagues have had no real rationale for investigating these factors rather than others, except in the case of impulsivity which, as we have argued, was causing unwanted correlations between *E* and *P*. For example, empathy was inserted into the study of sensation seeking (Eysenck and Eysenck, 1978) to provide buffer items at least in part. This is in contrast to the work of Cattell whose primary factors were intended to cover the whole personality sphere as defined by rating subjects for all non-synonomous descriptive terms (Cattell, 1957).

Indeed Eysenck may now have begun to tread the primary factor path to purgatory! For as we have argued (Kline, 1979) in the field of abilities, it is ever possible to develop more and more factors by the construction (perhaps by Guilford's three facet model; Guilford, 1967) of tests of homogeneous items. It is desperately necessary to find some way to order the factors. Carroll (1980) indeed in the field of abilities has begun this task by the careful reanalysis of studies using adequate and comparable methodology. Even so new studies are required with marker variables to aid comparison. Eysenck and colleagues may continue discovering more and more factors in their study of scales. Such proliferation will demand composite studies of these variables. Even so resulting factors may be little more than tautologous or bloated specifics. Their value can be evaluated only by their ability to predict real life behaviour and by their relation to external criteria. This is an enormous task and one which I hope readers will recognise as having already been attempted over the last forty years. We refer, of course, to the Illinois school of Cattell. It is ironical indeed that Eysenck now has reached this point. Plus ca change.

Before leaving the factors isolated by Eysenck and his colleagues it is necessary to point out that some of these factors have support for their identification (as should be the case with all questionnaire factors) beyond factor analysis. We have already indicated that *N, E* and *P* are claimed to be related to the behaviour of certain psychophysiological systems and some of the evidence in support of this assertion can be found summarised in Eysenck (1967), Eysenck and Eysenck (1976) and Eysenck (1981). In general it can be said of *N, E* and *P* that, whatever their basis and their names, they do correlate with a wide variety of human behaviours. In as much as this is so, they must be regarded as important personality variables.

The status of the other factors in the Eysenck system, the Eysenck Primaries, is of course far less well investigated, although Zuckerman (1974) has devoted considerable effort to the sensation-seeking scale. However, given the purported physiological basis of *E, N* and *P* and the fact that the other primaries are related to these, a powerful test of the validity of these factors is to investigate the extent to which they are hereditarily determined. With *N, E* and *P* this has long been known: in each of these, as is the case with intelligence, the population variance can be broken into approximately 81% attributable to genetic causes, the rest environmental. (Eaves and Eysenck, 1975, 1976, 1977 quoted by Eysenck and Eysenck, 1976). More recently Fulker (Eysenck, 1981) has carried out a valuable survey of the biometrical-genetical analysis of impulsiveness sensation-seeking and sociability.

Particularly interesting is the work on impulsivity and sociability because, as we saw, originally these were regarded as components of extraversion whereas now impulsivity relates to psychoticism. Although the results are too complex for a full report in this monograph which is not just concerned with the Eysenckian factors, it can be summarised without too much injustice.

(a) the genetic factors contributing to the variance of impulsivity and sociability do not contribute equally to both and the heritability of the interaction (0.72) supports the genetic and environmental determination of the positive covariation of sociability and impulsivity.

(b) The unitary nature of extraversion is evident in the environmental determinants of the trait, although the genetic correlation between sociability and impulsiveness is not so great.

(c) In conclusion Eysenck (1981) argues that about 60% of the reliable variation for sociability, impulsivity and extraversion is due to hereditary causes. Genetical and environmental factors account for the correlation between sociability and impulsiveness — 0.42 (genetical) and 0.66 (environmental). Combining sociability and impulsivity to measure extraversion yields the best method of discriminating between individuals with respect to the genetical and environmental determinants of their responses to these scales. The interaction between subjects and tests has a significant genetic component suggesting that sociability and impulsiveness can be distinguished genetically.

Before going further into the results presented in this paper, one point deserves comment, if the combination of sociability and impulsivity to measure extraversion is so effective, it seems quixotic as we have argued to have removed the impulsivity items from the *E* scale of the EPQ, simply because they loaded on *P*.

In this paper Eysenck (1981) reports on the biometric analysis of the four separate primary impulsive factors based upon the extensive work with almost 600 twins of Eaves *et al.* (1977). In biometric analysis, of course, there is far more than merely apportioning the determinants of variance as genetic or environmental. Models can be tested with variables such as dominance, assortative mating and indeed any other genetic or environmental factor. The results of this study have to be treated with caution because there was a good deal of error variance but nevertheless they indicate that there is a single underlying impulsiveness factor jointly affected by genetic and environmental determinants. There appear to be sex differences in the proportion of genetically determined variance for these variables (which may be due to error) and for most variables these are around 0.5, although for nonplanning the heritability is high 0.88 and 0.91 for males and females respectively suggesting, given the error, that the determination of this factor is wholly genetic.

Generally it can be concluded from this part of the research that impulsiveness is a general factor with a considerable genetic determination, as is the case with its component factors of which one in particular, nonplanning, seems surprisingly unaffected by environmental factors. Certainly, given these findings, it is difficult to argue that these primary factors of sociability and impulsiveness are merely statistical concepts based on the fact that there is an infinity of factor-analytic solutions such that there is no necessity that any factors should correspond to anything in the real world. Such genetical studies, where they show considerable genetic determination, as here, must indicate that the variance concerned reflects some underlying physiological mechanism, whatever this may turn out to be. This is the significance of the results psychologically, not the specific heritability ratios but the fact that there is genetic determination. Those who argue that test variance is specific to tests, are hereby quite refuted.

Finally we come to the biometric analysis of sensation-seeking in which Fulker *et al.* (in press) gave the four sub-scales into which Zuckerman's (1974) sensation-seeking scale breaks down: disinhibition, thrill-seeking, experience-seeking and boredom-susceptibility. About 70% of the reliable variance of the underlying trait appears to be genetic in origin, a result similar to that found for abilities, and in addition shared environmental influences (e.g. the familial factors) appear to play no part. As regards the subscales, it appears that these are under the control of different genes in the two sexes. We shall not here go into further details, since the relevance of these findings to this monograph lies in the fact that biometric analysis has shown a considerable genetic determination for the variables and this, of course, strongly supports the claim that these factors of sensation-seeking are not merely statistical artifacts but do reflect important and physiologically based behaviours.

Eysenck (1981) succinctly summarises this work in his concluding paragraphs. As was established for the secondary factors *P, E* and *N* so for impulsivity and its four subscales and sensation-seeking and its four subscales, biometric analysis shows that a simple model utilising additive genetic variation and *specific* environmental factors (E_1) can account for the population variance. There is no evidence for common environmental factors (E_2) (such as influence of family) producing differences on these scales, a finding contrary to most theories of personality. Eysenck indeed even offers the possible argument that this environmental effect may be constitutional, the environment actually modifying the individual organism e.g. hormonal imbalances during pregnancy. That there should be sex differences in the genetic determination of behaviour (as evinced by the findings with the subscales of impulsivity and sensation-seeking) has also powerful implications for the psychology of personality and, of course, runs counter to the egalitarian feminist zeitgeist.

From the viewpoint of the monograph, we can now argue that:

(1) The EPQ has yielded three scales *E, N* and *P* which are of broad general influence and that;

(2) These can be broken down into primaries; four impulsivity and four sensation-seeking scales which themselves load up on minor secondaries, impulsivity and sensation-seeking;

(3) All these variables show considerable evidence of genetic determination,

while environmental determinants are those not found within the family;

(4) There can be no doubt that these factors are of considerable psychological significance and that the Eysenck questionnaires must be regarded as measuring variables of some importance to the understanding of personality and not as merely psychometric exercises in the construction of homogeneous tests.

FACTORS IN THE GUILFORD TESTS

Guilford, one of the earliest of the factor-analysts of personality and one of the most distinguished, has developed since the mid-thirties a considerable number of factor-analytic personality questionnaires (e.g. Guilford and Guilford, 1934, 1936; Guilford and Zimmerman, 1956). Recently in a handbook to the Guilford–Zimmerman Temperament Survey (Guilford *et al.,* 1976) there has been an attempt to summarise the psychological findings concerning these Guilford Personality factors.

However, before we begin our scrutiny of these factors it is necessary to make a few preliminary points.

(1) There is no definitive list of personality factors in Guilford's work, as Eysenck and Eysenck (1969) point out. Factors tended to be developed as new tests were constructed. We shall restrict ourselves to the Guilford and Zimmerman (GZ) factors as these are the best documented.

(2) As is clear from the recent Handbook, there is relatively little information about the nature of these Guilford factors, compared with what is known about the Eysenck and Cattell factors. This is largely because most of Guilford's research has been directed towards the field of abilities, rather than that Guilford was content to isolate factors in sets of items and leave it at that, as is the wont of many lesser psychometrists.

(3) Guilford is almost alone among reputable factor analysts who are prepared to work with primary factors, in preferring orthogonal rather than oblique factors. This means that, for simple geometric reasons, comparison between Guilford's factors and other oblique sets is not useful. Re-rotation is virtually essential.

Given our strictures on the importance in exploratory factor analysis of obtaining simple structure, the following questions arise with respect to these Guilford factors: (a) Are there good external criteria supporting the identification of the factors? If there are, then whether or not simple structure was obtained becomes of less importance. If the factors can predict real-life behaviour or discriminate meaningfully between groups, then that is good confirmation of their psychological significance. (b) Since these factors are orthogonal, the question arises as to whether, if rotated to simple-structure, they in fact differ from other factor sets, for example those of Cattell in the 16 PF test. If on re-rotation, they align themselves with another set of factors, their raison d'etre has gone, and their labels should be changed. (c) Finally it may be asked whether in fact Guilford's orthogonal solution is not the most simple, in any case.

First we shall list the factors in the Guilford Zimmerman Temperament Survey: G = general activity; R = restraint; A = ascendance; S = sociability; E = emotional stability; O = objectivity; F = friendliness; T = thinking introversion; P = personal relations; M = masculine emotions.

The first point to note here is that there are obvious similarities to some of Eysenck's factors, which is hardly surprising since introversion was originally (see Eysenck, 1947) extracted from Guilford's work by Eysenck. Clearly, *R*, restraint, the opposite of impulsivity and typical of the inhibition of the introvert, is a close relative of Eysenck's primaries. Similarly *T*, thinking introversion, as the name suggests, is a component of extraversion–introversion. *S*, sociability, Eysenck measures both as a primary and as part of extraversion, *F*, friendliness, would also appear related, *E*, emotional stability would appear to be similar to Eysenck's, *N*, neuroticism. Thus an examination of the factors and their description suggests strong resemblances to the work of Eysenck. It would not be unlikely, therefore, that rerotation would indicate that these factors are empirically, as well as conceptually the same.

However, since it is possible that Guilford's orthogonal location of the factors is superior to that of Eysenck and others factorists' oblique rotations (such that if the factors are the same, the other factorists should re-rotate their factors) we must now turn to our first question, and scrutinise the evidence for the psychological meaning of the Guilford factors.

The Meaning of the Guilford Factors

The evidence for this is contained in the Handbook to the GZTS (Guilford *et al.*, 1976) which lists and summarises a very large number of studies carried out with this test. However, as an indicator of the psychological significance of the factors this catalogue of research findings is disappointing. Many of these researches are doctoral dissertations with samples insufficient to allow any conclusions to be drawn. Furthermore, although the Guilford profile scores of many different groups are included the authors make little attempt to integrate the findings or interpret the meaning of any group differences. Thus, although there is much information in this Handbook, on careful scrutiny it turns out to be insufficient for it to be used to give psychological meaning to the factors. Without this it, therefore, becomes necessary to examine our other questions as whether these factors are orthogonal versions of other factor sets and whether better factors (in terms of simple structure) cannot be found.

Relation of GZ Factors to the Eysenck and Cattell Factors

Eysenck and Eysenck (1969) investigated this question in a joint study of the EPI items and selected items (to make the computations possible) from the Cattell and Guilford inventories. These were administered to 600 male and 600 female students. At the primary factor level, the Promax oblique rotation of 20 factors (which was an arbitrary choice and was less than the number of factors which would have been rotated using the criterion of the Scree test — hence the possibility of under factoring) failed to reveal the Guilford factors. This casts doubt on their stability although it must be remembered that only short forms of the factor scales were used. At the second-order the Cattell and Guilford scales did not combine into meaningful factors although the second-order analysis of the Guilford items on their own did produce five factors that, given the oblique solution, were reasonably close to those of Guilford. At the third-order, two large factors emerged, *E* and *N*.

In respect of the Guilford factors, this investigation by Eysenck and Eysenck (1969) fails to confirm their identity and certainly does not suggest that they are orthogonal versions of Cattell's oblique set. However, as Kline (1979) has argued, the results of this analysis have to be treated with some caution not because it is flawed with gross errors of method but because there are a number of small points of procedure which in sum could contribute to this failure to reveal the Guilford factors, the same arguments applying, of course, to the Cattell primaries which equally failed to load their expected factors. These procedural problems (which are fully discussed in our earlier section on this topic) are: (a) the correlation of items rather than parcels or clusters of items. Item factors are usually of small variance; (b) the fact that there were only 8 items per Guilford factor. Any item failures would have certainly rendered factors unstable; (c) the Promax rotation which may have failed to reach simple structure; (d) the possible underfactoring which can lead to the emergence of higher-order factors at the primary level; (e) unities rather than communalities in the diagonals; (f) no check on the hyperplane count and the adequacy of the final solution. It is to be noted that these inadequacies would be likely to affect the primary factor structure, which in this case is the one of interest, rather than the higher-orders.

Thus, while we by no means wish to reject this evidence for the lack of clarity in the Guilford factors, a finding which we do accept, it is, nevertheless, important to be aware that this study by Eysenck and Eysenck (1969) has to be treated cautiously.

This investigation, therefore, certainly fails to support any claims that the Guilford factors as described in the Handbook to the GZTS (Guilford *et al.*, 1976) should be regarded as important personality factors.

However, we must now contrast this work with an earlier study by Cattell and Gibbons (1968) who administered 424 items purportedly measuring 15 of the Guilford factors and 14 of the Cattell factors to a sample of just over 300 undergraduates. Technically this research meets most of the criteria for adequate analysis as described in this monograph. Thus the 424 items were parcelled into 68 variables in an effort to overcome the problems of correlating dichotomous items. The significant factors to be rotated were selected by the Scree test, and, after the Varimax orthogonal solution was shown to fall short of simple structure by the Bargmann test, an oblique rotation using the Maxplane topological procedure (Cattell and Muerle, 1960) hand-adjusted by Rotoplot (Cattell and Foster, 1963) was carried out, a rotation which was shown to have obtained simple structure by the Bargmann test.

Before scrutinising the results of this investigation a few comments on the procedures will be useful. The Varimax analysis is highly important, because it is an orthogonal solution, and hence essential to test the claim that Guilford's factors are orthogonal and in the simple structure position. However, the failure of the Varimax rotation to reach simple structure does not, as Cattell and Gibbons (1968) accept, necessarily reject Guilford's hypothesis because its failure may be attributable to the fact that Cattell items (approximately half the total) would not take up the orthogonal position. Strictly to refute this claim, Guilford's items should have been factored on their own.

Maxplane and Rotoplot were, until the development of more recent analytic rotation procedures, the most efficient methods for obtaining simple structure (as Cattell, e.g. 1973, argued). However, they are cumbersome and difficult to use, especially rotoplot, which is highly dependent upon the skill of the user. Thus Direct Oblimin is a preferable rotation. However, since simple structure was shown to have been obtained it is probably fair to argue that in this study, this point is of little importance. Essentially, therefore, we can regard this research by Cattell and Gibbons (1968) as technically adequate to answer the questions it posed.

The answers, indeed, were surprisingly (in view of the Eysenck and Eysenck (1969) study) clear.

(1) As already indicated the Guilford factors did not emerge from the orthogonal rotation, although this might have been due to the influence of the Cattell primary factors.

(2) In the oblique solution the Guilford factors tended to align themselves with the Cattell either showing themselves equivalent to the Cattell factors, or loading up on more than one Cattell factor. The inference here is, of course, that essentially the Guilford and Cattell factors are the same and that the differences lie in the orthogonality and obliquity of the factor analyses. Since the oblique position gives the simplest structure this solution must be preferred. Hence the Guilford factors essentially disappear. They cannot be regarded as questionnaire factors, in their own right.

The conclusion from these two studies of the Guilford factors must be that it is difficult to consider them as well established orthogonal factors. Oblique rotation seems to account for the variance in the items with greater simplicity, but the factors thus emerging are little different from those of Cattell. In addition without substantial external evidence confirming this identification (either in oblique or orthogonal location), it would be unwise to include them in any list of clearly identified personality questionnaire factors.

FACTORS IN THE CATTELL TESTS

Since the forties, Cattell and his colleagues at Illinois and lately at Hawaii have carried out the most extensive factorial studies of personality undertaken anywhere in the world. In principle their work is of outstanding quality in that they have attempted to locate factors in real-life behaviour, questionnaires and objective personality tests; they have developed the necessary psychometric and factor analytic methods and have attempted to investigate the origin of these factors and to identify them against external criteria. In brief, Cattell's work has been aimed to establish a psychometric model of man (see Kline, 1980). In detail some of this work is, inevitably, less good, e.g. smaller samples than is ideal, and less replication than is demanded when theory building is contemplated.

This vast research has resulted in huge numbers of papers and books so that summarising is difficult. Nevertheless much of the relevant information can be found in Cattell (1957, 1973 and 1981) and Cattell and Kline (1977); our brief synopsis is based largely on these sources.

Basically Cattell has claimed that 16 factors as measured by his 16 PF test (Cattell *et al.,* 1970) account for the variance in normal personality. These *Q* or questionnaire factors can be regarded as embracing all personality because his early studies (e.g. Cattell, 1957) showed that these covered the whole personality sphere as defined by behaviour ratings. Recently Cattell and Delhees (1973) have added seven further factors, and twelve abnormal factors for clinical use have also been developed as described in Cattell and Kline (1977).

There have been many investigations of the factor structure of the Cattell factors from 1960 onwards. There have been divergent results from these studies except that generally the hypothesised Cattell structure has not emerged: the factors that did emerge, however, were usually hard to identify and showed little agreement from study to study of which the best known are: Levonian (1961a, 1961b); Eysenck and Eysenck (1969) which we have described with reference to Guilford's factors; Howarth and Browne (1971); Comrey (1973) and Howarth (1976). This disagreement, however, concerned the primary factors; at higher-orders, two factors emerged with great clarity: anxiety and exvia. As Kline (e.g. 1979) has shown there is no doubt that these are identical to Extraversion and Neuroticism, as isolated by Eysenck. Although the labels are in dispute there is no question that Eysenck and Cattell agree their identity.

There is no point in reviewing the results of these studies of the Cattell factors because Cattell (1973) has argued that these divergent findings are attributable to poor methodology: under or over-factoring due to ineffective procedures for choosing the number of factors to rotate; failure to reach simple structure, usually because ineffectual rotation methods were employed. These are but two of his criticisms and in fact our rules for carrying out adequate factor analyses were largely based upon his strictures. Since it is unclear on theoretical statistical grounds whether or not the objections raised by Cattell would in fact invalidate the findings of these researches we decided to carry out an investigation of the factor structure of the 16 PF which met all the criteria of adequate analyses suggested by Cattell. This research is fully described in Barrett and Kline (1982b) and we shall here summarise the relevant results. Broadly our methods were those that so clearly revealed the factor structure of the EPQ, which we have described earlier in this monograph.

Subjects

Two hundred and forty-one female and 251 male undergraduate university students. This, from the viewpoint of the matrix algebra of factor analysis and the standard errors of correlation coefficients, is certainly a sufficiently large sample to allow us to factor analyse the inter-item correlations of the 16 PF.

Test Form

Form A of the 16 PF test was used (184 items).

Factoring Method

The Pearson product moment correlations between the 184 were subjected to Principal Components analysis. Although the work of Velicer (1976, 1977) and Nunnally (1978) shows that with matrices of more than 20 variables there is little difference between principal component and principal factor-analysis, to overcome any possible objections image analysis was also carried out (Guttman, 1953).

Number of Factors in Rotation

For both types of factor analysis, since there is no consensus as to the best method of factor extraction, three procedures were employed: (a) the Kaiser factor alpha criterion (Kaiser, 1960, 1965); (b) the Velicer test (Velicer, 1976a); and (c) Cattell's Scree test, which was automated on computer. (The agreement between this autoscree and the senior author's scree estimates by eye on 21 examples was high with a mean error in the number of factors less than 1.) This autoscree procedure is fully described in Barrett and Kline (1982c). Significant factors were subjected to an oblique Direct Oblimin analysis.

Factor Validity Coefficients (Cattell and Tsujioka, 1964)

These coefficients were calculated for the factor scales, *A* to *Q*4 and for the factors obtained in our analysis. These coefficients, it will be remembered, are defined as the ratio of mean validity (mean item-factor correlation) to mean homogeneity (mean inter-item correlation). This is a useful index since the production of bloated specific or tautologous factor where items are virtual paraphrases of each other is immediately revealed by a moderate factor-validity coefficient.

Radial Parcel Analysis

Since Cattell (1973) has argued that item factoring can flounder on account of the error inherent in item responses, and that radial parcel analysis can lead to statistically more reliable results, a radial parcel analysis using parcels of both 2 and 4 items was also carried out. Radial parcel analysis is a complex technique (fully described by Barrett and Kline, 1981b), in which essentially the most similar items in terms of their total factor loadings are clustered together.

Rotation of 16 and 19 Factors

Finally before we used any of our tests of factor extraction 16 factors were rotated for obvious reasons and also 19 as suggested by Cattell (1972) in his reply to Eysenck.

We have delineated our methods in this way, because we want to make it clear that all the technical objections against those studies which failed to find the hypothesised Cattell structure have been met. In addition these methods meet our own criteria for adequate analyses. Indeed the only possible cause of failure to find the Cattell factors could be the fact that we only used one form of the test. Ideally we should have used the full item set. However, it must be pointed out that, (a) most users of the 16 PF would only have time to administer one form of the test, and (b) if the structure is not discernible in one set it is of dubious validity to add the items to another set, presumably of a different structure.

Results

As is obvious from our description of the methods, there were far too many analyses to report in detail. Here we shall concentrate on those findings most critical to this monograph. Since there was virtually no difference between the principal component and image analyses, thus supporting Harman's (1976) contention in the case of large matrices, we shall discuss only the principal component analyses.

(a) The Velicer and automatic Scree tests indicated that 11 factors should be rotated and this was done.

(b) As indicated above 16 and 19 factors were also rotated. All these solutions yielded a hyperplane count of around 66%. In fact, none of these solutions revealed the 16 Cattell factors. Only the *I* and *G* factors (tender-mindedness and superego) had factor loadings of any clarity. The other factors were not only composite but impossible to interpret. Since the objection to factoring items (that factors merely reflect item difficulty levels) has been made (e.g. Nunnally, 1978), this possible interpretation was investigated. However, there was no sign of this phenomenon. It had to be concluded from this part of the study that there was no support for Cattell's 16 factors.

(c) Factor validities — Cattell, of course, has stressed the importance of calculating factor validities which are usually very high for his scales. Consequently we calculated these for his scales using again all three solutions. Again the results were disappointing from the viewpoint of establishing the validity of the Cattell scales. For although the factor validity coefficients were not dissimilar to those of Cattell, many of them were lower than 0.6 and the meaning of such low coefficients is by no means clear and in addition certain scales, for example *C*, *O* and *Q*4 had their highest validity on factor 1. Thus even the attempt to compute factor validities for the scales was not successful.

(d) Radial parcel analysis — As a final test the 11 and 16 factor solutions were subjected to radial parcel analysis using both 2 and 4 item parcels. However, inspection of the items showed that they were grouped by factors (that on which the item loaded most high) such that no difference in the results from the item analysis could have been expected. Thus radial parcel analysis certainly failed to confirm the hypothesised 16 PF structure.

(e) Obtaining the statistically best factors — Since the 16 PF form A items have been so widely used and were originally designed to tap factors based upon real-life ratings of behaviour, we decided to ignore the hypothesised Cattell scales and instead attempt to locate the best possible factor scales from the matrix, using as criteria, high validity coefficients, factors with high alphas, and item loadings beyond 0.2. In addition classical item analysis of the scales should show high item-total correlations. First the 11 factor solution was used to obtain a set of items representing 11 scales. However, item analysis of these scales indicated low homogeneity and a subsequent rotated factoring of the intercorrelations of these items yielded factors of low alphas such that the results had to be discarded. However, a different approach was then adopted utilising the factor validities which we have previously mentioned. Scales were selected (which might comprise more than one Cattell scale) based on factor loadings, alpha coefficients and the sharing of maximum validity coefficients across the three (11, 16 and 19 factor) solutions. This yielded 7 new scales:

(1) $C + O + Q4$; (2) $E + H$; (3) $G + Q3$; (4) 1; (5) L; (6) B; (7) $Q1$.

This means, of course, that *A, F, M, N* and *Q*2 were discarded.

The responses to the 16 PF were then rescored on the basis of these 7 scales and subjected to item analysis. The alphas for these seven scales were satisfactory except for scale 1 where the *C* items had little relationship with the total score. However, since *C* is so important a variable in the Cattell set of factors it was decided to retain the items. The 128 items of the seven scales were then subjected to a direct oblimin analysis. However, the resulting seven factors still had low factor validities, so that factor *C* was discarded and the new set of 115 items were factored.

Results

The Direct Oblimin oblique rotation, using seven factors by the criterion of the Scree test and the Velicer test, was quite satisfactory. The factor validities and alpha coefficients of the seven new scales are set out below.

Conclusions

From this study of the 16 PF items, it must be concluded that only seven factors can be found among them and that many of the items do not uniquely load any factor. What these factors are is of considerable interest, although

Table 2.

	C Scale Missing		
Scale	Coefficient Alpha	Factor Validity	Position
O+*Q*4	0.73	0.83	2
E+*H*	0.81	0.83	1
G+*Q*3	0.71	0.87	3
I	0.64	0.86	5
L	0.45	0.86	6
*Q*1	0.40	0.54	4
B	0.44	0.76	7

The hyperplane count for this solution was 55%.

until the results have been replicated considerable caution in their interpretation must be exercised.

Factor 1

$O+Q4$, ergic tension and guilt is clearly an important component of Neuroticism or anxiety. Its emergence at the first order is not unexpected as Eysenck has noted a similar phenomenon with his superfactors N, E and P (Eysenck, 1978).

Factor 2

$E+H$, dominance and adventuresomeness is thus an important component of the other second-order factor extraversion. Note that H resembles our previously discussed sensation-seeking factor of Zuckerman. Thus again these two factors are essentially similar to two in the EPQ.

Factor 3

$G+Q3$, conscientiousness + high self sentiment is particularly interesting. Both these factors are conceived of by Cattell also as drives (ergs and sentiments; Cattell and Child, 1975) and in the terminology of McDougall as master sentiments. As a factor it would appear to be similar but not identical to the obsessionality factor, which will be discussed later in this monograph. It is probably best thought of as a self-sentiment factor.

The other factors are as named in the Cattell test — tough-mindedness (perhaps a normal version of Eysenck's P factor), suspiciousness, radicalism and intelligence. It is noteworthy here that two of these factors, 1 and $Q1$, have been utilised by Eysenck in his study of political attitudes (Eysenck, 1954).

These then are the seven factors that reliably emerged from a factorisation of the Cattell items, using the methods which he has advocated. We must now examine the factors found in other personality questionnaires, although this can be done far more briefly since none of these is supported by any great weight of research.

Before examining these other scales one further point about the 16 PF factors should be made. In the course of our study of this test we factored the scales. The factor pattern was not dissimilar to that claimed by Cattell and at the second order Anxiety and Exvia (N and E in Eysenck's terms) emerged with clarity. This leaves us in an anomalous position — the items in the factor scales appear inefficient but the scales themselves appear to work. We cannot readily explain this finding although one possibility is that the items are failing due to error which sums to zero over the total scale items. This finding could explain the fact that sensible results have been achieved with the 16 PF test in applied psychology in educational, clinical and occupational fields (see Cattell and Kline, 1977) despite the fact that at the item level the factors are far from clear. Nevertheless our conclusion must still remain that there are only seven factors in the 16 PF test.

The Seven Extra Factors and the Abnormal Factors

As regards the seven extra factors and the 12 abnormal factors, as yet there is little independent evidence examining their replicability or their relationship to other factors which is essential for identification until external criterion studies of the factors are carried out. The abnormal factors were obtained in part from the MMPI item pool and this will be discussed in a later section of this monograph.

Saville and Blinkhorn (1981) carried out a detailed psychometric examination of the 16 PF test on two large samples, 2000 adults and 1150 undergraduates, samples whose data had been previously reported (Saville, 1973; Saville and Blinkhorn, 1976). However, their study despite its title has little relevance to our monograph since they factored neither scales nor items, concentrating rather on the ultimately trivial problem of whether suppressor items worked in the scales. This is disappointing given the huge samples and the fact that the two forms of the test were used.

COMREY PERSONALITY INVENTORY (COMREY, 1970)

The Comrey factors must now be examined because, as the manual to the test makes clear, they were developed because their author felt that something was wrong with current factored personality tests in that they purported to measure different factors. In addition, Howarth (1978) reviewing the test in Buros (1978) regards it as one of the best available inventories. Certainly the scales are satisfactorily reliable, although as yet there is little evidence for their validity in terms of external criteria. There is further reason for considering these scales worthy of scrutiny. In an effort to avoid the problems of item unreliability, which is always exacerbated in the factor analysis of the inter-item correlation matrix, sets of items that were empirically (from previous factorings) and semantically similar were forced into factored homogenous item dimensions: these formed the basis of the subsequent factorings.

With little in the way of validity studies of these factors, the only hope of identifying them properly lies in locating them in the personality sphere in relation to the best established factors. This is particularly important because as Cattell (1973) has argued, there are technical problems with Comrey's factor-analytic methods. The use of FHID's can lead to second-order factors emerging from the analysis, the FHID's being effectively short scales. Furthermore the rotation methods specially developed by Comrey are unlikely to reach simple structure so that the solution (in the absence of external evidence) is not entirely trustworthy. This view of Comrey's methods, strongly argued by Cattell (1973) with which we are in general agreement, is not universal, for Howarth (1978), as we have discussed, regards these factors as well supported. Thus, all turns on the empirical work. Are the Comrey factors independent of those factors which we have reviewed or are they essentially identical but rotated to a different position?

Comrey and Duffy (1968) studied the Comrey FHID's and the 16 PF and EPI scales. The factor analysis was unfortunately adjusted by hand to psychologically meaningful position, thus making the research subjective and of dubious replicability. However, despite these difficulties two findings stand out. There was a clear extraversion factor loading on the EPI, *E* and the Cattell exvia primaries. Similarly there was an indubitable *N* factor. Thus the two largest higher-order factors can be found in the Comrey test. The other factors did not align themselves with the Cattell factors, but this is hardly surprising, (a) in view of the hand-rotation, and (b) in view of our discussion of the Cattell primaries. From this research it can be firmly concluded only that *N* and *E* appear in the Comrey Inventory: ubiquitous factors indeed.

Barton (1973) in a research cited by Cattell (1973) carried out a factor analysis of the 16 PF, EPI and Comrey scales, using oblique rotations and other techniques recommended by Cattell and thus fulfilling the technical criteria which we have advocated. Cattell (1973) claimed that this research showed that the Comrey factors were essentially the Cattell second-orders. However, the Illinois results have been difficult to replicate and further studies are required on this.

Comrey (1970) claims in the test manual that the Guilford scales which correspond to the Comrey scales are indeed correlated. However, given the dubious validity and status of the Guilford scales, the implication for the meaning of the Comrey factors is unclear. In the absence of powerful external validity for the Comrey factors, and given their idiosyncratic method of construction, yielding rotations of less than simple structure, the only conclusion in the light of the correlational and factor analytic evidence that can be drawn is that N and E appear in this scale. The psychological meaning of the other items, if any, remains to be worked out.

THE JACKSON PERSONALITY RESEARCH FORM, THE PRF (JACKSON, 1974)

The PRF is regarded by its reviewer in the Buros (1978) Yearbook, as a marvellous example of test construction (Hogan, 1978). Item analytic methods with large samples and obsessional attention to psychometric niceties were used to measure the most important needs from Murray's work described in *Explorations in Personality* (Murray, 1938). Although Jackson considers the PRF to be embedded in theory, one may well question the validity of Murray's theorising and see this as a positive disadvantage. Our reason for scrutinising the factors in the PRF is simple. This is a test with scales of impeccable reliability, and of which the scales are of unknown factorial composition (the PRF being constructed by non-factors-analytic procedures). It is increasingly used, hence its factor structure is of great interest.

However, as we pointed out in our discussion of adequate research designs, the factor structure of the PRF scales on their own is not useful, for these, as the Comrey scales, have no external criterion-evidence for validity. What is needed are factor analyses with other personality scales.

Nesselroade and Baltes (1975) factored the PRF and HSPQ (the adolescent version of the 16 PF test) scales in a study of 1662 adolescents, which utilised technically adequate methods, although for the purposes of this monograph, the item rather than scale factors would have been more useful.

Nesselroade and Baltes (1975) claimed that eight factors would account for the variance in the PRF — conscientiousness, ascendance, independence, aggression, aesthetic-intellectual orientation, social contact and one further nameless factor.

Only *B, C, O* and *Q*4 of the HSPQ did not correlate significantly with the PRF factors, while conversely only five of the 20 PRF scales failed to correlate significantly with the HSPQ. Four of the PRF factors indeed were similar to those of Cattell — ascendance (to exvia), infantile control (to superego), aesthetic orientation (to cortertia) and independence (to avoidance of social contact).

There are severe problems in drawing conclusions from this study despite its large sample size and technical proficiency. The first concerns the meaning of the Cattell factors in the first place. Although Cattell and his colleagues, not unnaturally regard them as a reference set of factors, as we have made clear, they cannot be thus described. For this reason we have not discussed this paper in more detail. What is really required is an item rather than a scale factoring of the PRF with other tests. Nevertheless some conclusions of a more general nature, can be drawn.

The first is that, despite the labels and provenance of the PRF scales, which both suggest that the test lies in the field of dynamics rather than temperament, the substantial intercorrelations with the HSPQ factors demonstrate this not to be the case. It may be that the PRF scales measure dynamic variables also but they certainly lie in the sphere of temperament, as is shown by the low correlations between Cattell's MAT and 16 PF tests (Cattell and Child, 1975).

Despite the problem of the replicability of the HSPQ factors, one finding stands out. The *E* (exvia) factor which we have shown to be virtually ubiquitous, again emerges. So does *G*, which was one of the Cattell factors emerging from our study of the items. The labelling of the other factors is dubious without further evidence from critical criteria, although it does appear that the anxiety factor (*O*+*Q*4) is missing from the PRF.

In brief, therefore, it can be argued that despite the difficulties with this investigation, it is possible to show that the PRF measures temperamental variables and that extraversion almost certainly emerges from it. The superego or obsessional factor is also present. Whether the other factors are largely specifics or rotated to their best position cannot be settled until further studies of the internal validity of the scales is undertaken and until item factorings of the PRF are carried out. Clearly it overlaps the Cattell scales and this is a set of items which demands analysis.

Guthrie *et al.* (1981) factored the PRF in a sample of Philippino students. Six factors accounted for the variance, factors which were similar to those found among comparable American and French samples. Despite the fact that the interpretation of factors in cultures different from those for which the test was originally designed is difficult and despite the fact that these factors are not located against others, the fact that the findings are stable makes them worthy of brief note. The six factors appear to be: impulsivity, endurance, abasement, exhibitionism, nurturance, succourance and avoidance of harm. Of these impulsivity we have already found in our temperamental factors, but the others appear different and, without relation to other scales, identification is not possible.

In conclusion we would argue that the PRF which is psychometrically powerful, contains six factors. There is evidence that there is a measure of extraversion, but generally these factors require proper location in factor space.

FACTORS IN THE DYNAMIC PERSONALITY INVENTORY (GRYGIER, 1961; GRYGIER and GRYGIER, 1976)

The Dynamic Personality Inventory was first devised by Grygier (1961) and attempted to measure the psychosexual personality correlates of various patterns of psychosexual development, although it was stressed in this first experimental manual (Grygier, 1961) that psychoanalytic theory was regarded as stimulus rather than a rigid framework for the production of items.

Kline (1981) has carried out an extensive review of the evidence pertaining to the validity and the factor structure of these scales, in his study of the scientific support for psychoanalytic theorising. Readers must be referred there for the full details of our examination of this test. However, it is possible to summarise the results quite briefly.

(1) Although some effort has been made in the most recent manual to the test (Grygier and Grygier, 1976) to validate some of the scales against external criteria, this does not, in this instance, help us to answer the question as to what are the factors in this test. First because even if the scales were validated, the scales themselves might not be unifactorial. Furthermore, psychoanalytic

theorising is so flexible that it is difficult to be confident that a given result does validate the theory or not.

(2) Thus what are required are factorial studies of the Dynamic Personality Inventory. There are numbers of these cited in the manual by Stringer (1976) but the standards of these investigations fall far short of the criteria demanded of adequate studies. Indeed few of the researches reach simple structure and almost none attempt to locate the DPI factors in personality space. Without this or external evidence of validity little faith can be placed in any factor identification. Indeed there is only one study of the DPI that is anywhere near adequate either in technique or design — that of Kline and Storey (1978). In this research the DPI was administered to 128 subjects (61 male) together with the EPI and the 16 PF tests. This enabled us to locate the DPI factors relative to those of Cattell and Eysenck. Since, however, many of the DPI scales claimed to measure traits far different from those in those tests, in addition we inserted the best established measures of such traits, including tests by Gottheil (1965), Lazare *et al.* (1966) and the present author's measures of oral and anal traits (Kline, 1971, 1978).

Unfortunately with the huge numbers of items involved in this study which would have necessitated a minimum number for an adequate item factor analysis of 2000 subjects, it was only possible to factor the scale scores. In fact, fifteen significant factors were rotated to oblique simple structure using Promax (Hendrickson and White, 1966). While it is true that Promax does not always reach simple structure and that our choice of the Kaiser-Guttman criterion for significant factors is one which we dislike (see p. 150), in this particular study there is little doubt that a technically satisfactory solution was obtained. This is evidenced (a) by the clear emergence of *E* and *N* as predicted and by the usual Cattell 16 PF scale structure, and also by (b) the similarity of the factor structure to a pilot study of the DPI carried out on a small student sample ten years previously (Kline, 1968a).

The first point to note about the results is that the DPI is clearly measuring factors different from those found in the EPQ and 16 PF tests. From the viewpoint of this monograph, however, three of the 15 factors are of interest as clearly being questionnaire factors which have not appeared in other scales and of a sufficient breadth to render them of psychological importance. These factors are set out below:

(1) The obsessional or anal character. This measures obsessional traits (a better name than anal albeit less picturesque, since the hypothesised link with anality has yet to be demonstrated). It loaded on the Grygier *A* scales — loading attention to detail — on *Ai3Q,* and on Cattell's *G* and *Q3* — conventionality and control. This is the typical obsessional character, and is probably a useful measure to include in a test battery. This factor is highly similar to the obsessional factor which we studied in detail in Kline (1968b).

(2) The second factor of interest to this monograph was the fourth factor in

the Promax rotation — one of feminine attitudes and interests. It loaded on the female interest scale of the DPI and on the DPI scales of dependence, liking for children, creativity, and some of the oral scales. This is clearly a factor of conventional female interests (as conceived in the West) and as such it makes sense to include in the measurement of personality in Western culture. It is surprising that no such factor emerged from the Cattell studies, given that their construction was based upon behavioural ratings. No Cattell scale loaded on this factor.

(3) The third factor was the masculine interest factor. Equally conventional as the female factor, it loaded on energy, drive, achievement, fascination with fires and lights (the Icarus Complex) and a liking for adventure. Only factor *N*, shrewdness, loaded on this scale.

Thus what appears from this combined analysis of the DPI with the 16 PF and EPI scales is that three meaningful factors virtually separate from those in the work of Cattell and Eysenck can be found: obsessional traits, masculine attitudes and feminine attitudes. The fact that *N* and *E* emerge clearly with their proper loading in this study demonstrates that technically all is well and we would argue that these three factors should be included in any list of personality questionnaire factors.

This examination of the Grygier factors, important because, as we have seen, they are independent of the factors most usually measured in personality questionnaires brings to an end our study of personality questionnaires. In general there are, remarkable similarities between them despite the different labels. The problem essentially has been to locate these factors relative to each and to find the simple structure that best describes the field.

Before we draw our conclusions concerning the factors in personality questionnaires, a few further papers and issues require a brief examination. Browne and Howarth (1978) carried out a joint factoring of all items from published personality tests, a study which in its design does all that has been advocated in this monograph. Sells *et al.* (1970) also carried out a less comprehensive but similar investigation which demands scrutiny. Finally, we must also consider the work done with the MMPI, a non-factored personality questionnaire which, however, does measure factors, factors however which lie beyond the sphere of normal personality, which lead us into the field of abnormal behaviour and which, therefore, are rather beyond the scope of this monograph.

THE WORK OF BROWNE AND HOWARTH (1976)

Browne and Howarth searched through published personality questionnaires and selected 1726 items which were ostensibly different from each other. Extensive though this research was, as is inevitable there were omissions. For example, the seven extra Cattell factors and the DPI factors were not included. To reduce computation, twenty factors were hypothesised and 20 items were selected relevant to each. These items were then rewritten into the yes/no format and administered to a student sample of 488 female and 510 male students. The inter-item ϕ-coefficients were subjected to oblique and orthogonal solutions (20 factors being rotated). Browne and Howarth (1976) attempted to find second-order factors but the primary factors were correlated to such a slight extent that their pursuit was not considered useful.

Eleven factors were stable across orthogonal and oblique solutions and were regarded as robust by Browne and Howarth, accounting for about 60% of the variance. These were social shyness, sociability, mood swings, emotionality, impulsiveness, persistence, hypochondriasis, dominance, general activity, trait and superego.

Before commenting further on this study we should like to point out how similar is this list to the list of factors that we have been slowly building up through this monograph. The first six factors have certainly already been identified.

However, there are some serious technical difficulties in this research which in our view mean that the results are less than definitive and that the factors cannot be regarded as established until they have been identified against external criteria or located in factor space in researches less open to criticism. Since we have previously set out these problems in some detail before (Kline, 1979) we shall here simply quote those points, for we cannot summarise them further.

"(a) Not all factors were included in the study. Thus it could be the case that the results are far from complete.

(b) The rewriting of the items may have changed their psychological meaning. The phrasing of items in personality tests crucially affects (and in ways not fully understood) the variance (Guilford, 1959).

(c) As has already been discussed the use of ϕ-coefficients can produce factors of item difficulty. Furthermore, the correlations are affected by the polarity of the items (Holley, 1973) and the proportion of subjects putting the keyed items. Item parcels (Cattell, 1973) may be more reliable as a basis for correlations.

(d) The original factor hypotheses which guided the item selection may have been faulty.

(e) In any case item selection in this way is little more than choosing items from their content. As have argued, the face validity of items is not a good guide to their validity.

(f) The extraction of the factors for rotation was subjective. As have argued, the number of factors rotated can affect the solution.

(g) Orthomax is an orthogonal solution. It is highly unlikely *a priori* that orthogonal factors are the best description of the personality sphere.

(h) The fact that both the oblique and orthogonal solutions were similar (in

11 cases they were considered to be the same) suggests that simple structure was not obtained.

(i) In fact, no test of simple structure was made."

Since, in addition to the nine difficulties which we have mentioned, the factors have no external validation, we have not included them in our list of questionnaire factors. This, however, is not meant to write off this study as worthless. It was a notable effort and cross-validation together with further study of the factors could be highly useful.

Howarth (1971, 1976, 1978) has developed over the years a series of personality tests, the HPQ (Howarth Personality Questionnaire), and the APF2, the Additional Personality Factor Inventory. The ten scales of the HPQ are called by Howarth, mainstream factors — factors about which in the literature there is agreement on their labelling and a degree of replicability (Harman and French, 1973). Actually we are dubious of these Harman and French criteria because technically poor studies can produce replicable artefacts. The APF2 factors are of lesser variance, according to Howarth but important for prediction. In addition to this there is an individuality inventory.

The factors in the HPQ are: sociability, anxiety, dominance, conscience, medical hypochondria, impulsiveness, cooperativeness, inferiority, persistence and suspicion. All are acceptably reliable. Factors in the Individuality Inventory are: future orientation, phlegmatic temperament, involvement with others, felt tension (state anxiety) self regard, independence, psychoticism, fate control and dislikes annoyances. APF factors are: fear of social unacceptability, hope, general activity, self-pride, existential realization, individual tolerance, "unusuality", self-actualizing, time anxiety and rigidity.

However, none from these lists of factors can be accepted as established until there is proper location with those factors that are known, or strong external validation. Almost nothing of this research has been attempted although the present writers carried out a small pilot research study of the HPQ and APF2 in which these factors were located relative to *E, P* and *N* of the EPQ (Barrett and Kline, 1980a).

This was a small scale investigation, its technical inadequacies being allayed to a large extent by the psychological clarity of the findings in which *P, E* and *N* emerged clearly, as expected. In this research the three personality tests (EPQ, HPQ, APF2) were administered to 79 subjects and the scale correlations were then subjected to a Direct Oblimin simple structure rotation in which seven significant factors (Scree test) were rotated. As aid to further analysis of the HPQ and APF2 items, point biserial correlations between each item and the *E, N* and *P* scores of the EPQ were also calculated. In this way, despite the rather small sample the meaning of these Howarth scales and their items should be elucidated.

The results were clear cut *N, E* and *P* emerging without a shadow of doubt.

However, the analysis failed to reveal any new factors within these Howarth questionnaires. This is not to deny their utility since many of the scales loaded on the superfactors as they should. A few scales had high loadings on just one factor, but in a study of this sort we cannot distinguish between specific factors and those with some generalisability.

In conclusion it can be argued that without further evidence there is no real support for any further factors, beyond those which we have discussed, in the Howarth inventories. External validation studies are necessary. This is the conclusion, too, drawn by Eysenck (1978) in his reanalysis of the Howarth and Browne data.

THE WORK OF SELLS, DEMAREE AND WILL (1970)

These authors carried out a factor analytic study of 600 items representing the Guilford and Cattell factors. Although there were three papers intended to be published only the first has appeared in which the 600 items were administered to 2550 airmen, and the item intercorrelations were subjected to a principal factor analysis followed by a Varimax and Promax rotation of 15 and 18 factors.

As the authors freely admit, there is a problem immediately with the number of factors to be rotated. They would appear, as both Cattell and Guilford have argued, to have underfactored, thus giving rise to second-orders at the first order. Certainly 18 factors is a small number to extract from so large a matrix even if as Sells *et al.* argue, at this point that the residual variance in the matrix was virtually exhausted. They claim that the Scree test was not feasible but it is difficult to see why not in this case: their extraction of only 18 factors without recourse to a reliable measure of factor significance does cast doubt on the significance of their work.

Despite this problem the results of this study are highly interesting, not so much for the substantive findings of new factors, but rather for what failed to emerge. First it is quite clear that in this analysis neither the Guilford nor the Cattell factors emerged. Of their 18 factors all were heterogeneous in respect of the items loading on them except for one which was loaded on Guilford's artistic interest. Of the scale factors, apart from *AA*, *N* loaded on one factor (16 of its 19 items) although this had other items loading on it.

However, too much should not be made of these findings, first because of the rotation difficulty and the under-extraction of factors and secondly because there was no check on the attainment of simple structure. Again the 18 factors accounted for only 22% of the variance. With such small factors there is always the possibility that what has emerged are simply tautologous factors, items with the same verbal content. The check on this, of course, is location in factor space with other known factors and external validation. The only conclusion to be drawn from the Sells *et al.* (1970) study is that the

Guilford and Cattell factors did not appear, as expected. One possible reason for the low factor loadings is simple numerical inaccuracy caused by the large amount of computation and the limited computer-capabilities at that time.

Mention has been made of the work of Harman *et al.* (1975) in our discussion of the Howarth factors. Harman *et al.* went through all the personality factor research work and listed factors that had been replicated. However, since they paid no attention to the stringent criteria necessary for adequate analyses and since the naming of factors by reference to their item loadings is highly subjective such that the notion of replicability is slippery (unless empirically or conceptually defined, as in this monograph) we have not given serious consideration to their work. As Cattell (1973) has argued, one factor properly rotated and identified is more valuable for insight into the field than a dozen confounded with statistical artifact and error.

THE MMPI AND ABNORMAL FACTORS

The MMPI (Hathaway and McKinley, 1951) was constructed without the aid of factor analysis: items being selected for scales if they were able to discriminate among abnormal groups and normals. This test which originally consisted of 9 scales but from which more than 200 scales can be derived (Dahlstrom and Welsh, 1961) has been widely used and it is clearly pertinent to ask what factors it contains.

However, we cannot in this monograph which is concerned with normal personality factors, go into this problem at all deeply. The identification and verification of clinical abnormal factors is a specialised task and implicit in it are theories of abnormal personality and behaviour. What interpretation is made, what clinical groups are used to validate the identification and interpretation are so bound up with particular clinical theorising, that it becomes a topic in itself. All that can be done here is to summarise briefly what has been found with studies of the MMPI. Factorisations of other clinical scales will have to be ignored.

With its 550 items, the early factorisations of the MMPI had to restrict themselves to scale rather item analysis. Generally, especially where scale factoring was done with normals (e.g. Kline, 1967) factors similar to extraversion and anxiety emerged. Orme (1965) indeed claimed that a general factor of emotionality would account for the results. Finney (1961) found six factors, but these were on small samples relative to the number of scales in the study and there was no check on simple structure, so that little reliance can be placed on the findings.

Wakefield and colleagues (1974) factored the MMPI scale scores of 205 married couples. They related the MMPI factor structure to a theoretical model of the MMPI assuming that it did in fact measure *E, P* and *N.* This ingenious procedure (in some ways superior to an empirical test, since actual scales might inadequately measure the theoretical constructs) appeared to support the claim that *E, N* and *P* account for the variance in questionnaires. However there were inexplicable male and female differences and the delineation of the model is subjective. This work demands replication.

However, the MMPI is not really suited to scale factoring. In the first place the scales are not independent since some items contribute to more than one scale. Furthermore, as Jackson and Messick (1961) have argued the scales are beset by response sets of acquiesence and social desirability.

Cattell and his colleagues were among the first workers who attempted to factor the items in the MMPI (although in fact they used item parcels) and to locate the factors in factor space. Cattell and Bolton (1969) factored these item parcels with the 16 PF scales. They found the 16 PF factors together with four abnormal factors. The MMPI items in these factors together with 16 PF items and items derived from other sources but held to reflect abnormal behaviour were then factored and 16 normal + 12 pathological questionnaire factors emerged — 7 depression factors together with paranoia, psychopathic deviation, schizophrenia, psychasthenia and general psychosis. These factors have been incorporated into the Clinical Analysis Questionnaire by Delhees and Cattell (1975). However, until far more research has been carried out with them among abnormal groups it would be rash to list these as well established factors, although the basic rationale of defining the abnormal personality sphere and attempting to describe it factorially is fine. Unfortunately, as McNair (1978) has argued, the CAQ is not yet validated. Until this is done these interesting factors must await confirmation.

Koss (1979) in a useful review of psychometric studies of the MMPI point out that Overall *et al.* (1973) and Hunter *et al.* (1974) in item factorings of two short forms of the test (which *in toto* included all items) yield six identical factors — somatisation, feminine interests, depression, psychotic distortion, low morale and acting out. No external validity is offered for these factors which are different from those found by Cattell and colleagues, as discussed above.

This brief discussion is sufficient to indicate that as yet despite the nearly 6000 citations to the MMPI, no definitive factor structure has emerged. What is clear, however, is that whatever factors finally are shown best to account for the variance, most of them will have relevance mainly to abnormal behaviour although, as Wakefield *et al.* (1974) have shown, *E, N* and *P* are probably also present. At present certainly the MMPI has not been shown to contain any factors important in the description of normal personality.

CONCLUSIONS

What conclusions can be drawn from this survey of factors in some of the best known personality questionnaires? First it is evident that the factor analysis of personality tests has reached a critical point, a mid-life crisis. In its youth, its aims were clear — to reveal the main dimensions of the field. As it grew and thrived, the achievement of this aim seemed near at hand, but was beset by problems; were primary or secondary factors better? how could simple structure be best reached?

Our study, however, reveals that all is not this simple. The failure of the Cattell primaries, despite the enormous statistical and psychological expertise behind their research and development, to emerge with any clarity and the equal failure of any of the other systems of primary factors to stand scrutiny, would appear to show that primary factors are to be eschewed. They are not stable enough to form a basis for theory or model.

Secondary factors are superior then as descriptions of personality. Certainly neuroticism and extraversion are two virtually ubiquitous factors in all questionnaires, and psychoticism also is clearly important. However, this beautiful simplicity (so simple indeed that many personality theorists feel that it must be inadequate) has broken down. The Eysenck factors contain sub-components of sociability and impulsivity. Factors of empathy and sensation-seeking have been found. In the terms of our metaphor, the subject has regressed to its infancy. Primary factors are emerging and with them the same old questions, how many, how important, how stable.

As we have discussed the study of questionnaire factors requires, (a) that factors be located in factor space, and (b) that factors be verified against external criteria. This second point becomes critical if, in fact, the work of Eysenck in extending his higher-order factors is to continue because in such studies there is always the risk of isolating bloated specifics, homogeneous, tautologous factors correlating with nothing in real life. This is why in previous publications (Kline, 1979; Cattell and Kline, 1977, just for example) we have advocated the Cattell factors so strongly: attempts were made to validate them externally.

Against this background our conclusions can now be better grasped.

(1) Three factors, higher-orders, neuroticism, extraversion and psychoticism do seem to have been properly and reliably identified, although there is less clarity concerning psychoticism. Their psychological nature has been extensively investigated and it makes good sense to regard neuroticism and extraversion as major dimensions of temperament, largely heritable dimensions reflecting psychophysiological mechanisms. Their cross-cultural applicability supports this claim and also supports the notion that there must be more to human temperament than these two dimensions since, superficially at least, national characteristics do appear different.

(2) There seems no doubt that primary factors, more specific than these higher-orders, are necessary for an adequate description of personality. Present work, although it has revealed a number of factors, as we have shown, is confused. Apparently well established factors are failing: new factors are

emerging. However, these need external validation. Without this these factors may be of little importance. At present we can probably support an obsessional trait factor, and a sensation-seeking factor, as primaries supported externally and by factor-analysis. Our other primary factors are probably of small variance and require external validation.

(3) One possible conclusion from a scrutiny of the research is that, in fact, there are a myriad of small factors, each on its own of little import psychologically or in terms of variance. In this case, of course, their value in the study of psychology would be limited unless it can be shown that they combine to form such syndromes as the authoritarian personality or dogmatism or some such. However, a study of second-order factors does not strongly support such an argument.

(4) Thus the meagre haul from 50 years factor analysis of psychometric tests is only about six factors, of which only two can be described with any confidence. We shall now briefly state where, it seems to us, further research must go.

(5) As we have discussed, the road that the Eysencks have freshly begun has been traversed before by those in Illinois and it has been fraught with difficulties. However, new techniques have been developed since then and these may be sufficient to resolve these problems.

(6) Conjoint studies of the best items in the better personality inventories must be carried out. However, to avoid the difficulties of estimating the number of significant factors and other methodological problems of obtaining simple structure, it is suggested that confirmatory analysis be used. We can hypothesise any system or systems of factors we like and put it to the test using confirmatory analysis. Scrutiny of the factors described in this monograph is likely to yield useful clues to formulating the necessary hypothesis. If the worst obtains, we can carry out exploratory simple structure analyses followed by confirmatory studies.

(7) Since this method will still allow the isolation of bloated specifics all confirmed factors must be subjected to external validation, against criterion groups.

(8) When this has been done, specification equations including ability factors and if possible motivational factors, should be drawn up for relevant behaviours e.g. job success, psychotic status and the specification equations should be tested: are the multiple correlations as expected?

(9) All resulting factors should be subjected to biometric analysis, as the findings from this are closely relevant to their proper identification.

(10) Every attempt should be made to tie these factors in to theory in the light of the findings obtained from the studies we have outlined above.

(11) From this it can be seen that we are arguing essentially for statistically and conceptually refined versions of what has gone on before in the study of personality tests. This is, indeed, true. However, in our view this study of the

factors in personality questionnaires should be undertaken only as part of more general multivariate analyses of personality data. Here we can indicate only briefly what this should include.

(a) The multivariate analysis of projective test data. These tests are likely to tap personality dynamics. Remember that multivariate analysis can refer not only to variables as in *R* factor analysis, but also to occasions thus tapping dynamics, *P* analysis.

(b) The multivariate analysis of objective tests, of which there are a huge number. These factors should be related to our questionnaire data.

(c) The multivariate analysis of percept-genetic material as described by Kragh and Smith (1970) which appears to yield measures of defences and other processes.

(d) This leads on to our final point, the process analysis of all the factors in all types of test. Factor analysis *per se* is a static representation of covariance between variables, this covariance measured at a time *t*. There is ultimately no way that a dynamic process may be modelled using a factor model generated from data collected at a single point in time. Rather than simply observe test responses, the challenge is now to begin modelling the processes whereby these test responses originate and to explore the interaction between response sets as indicated by the structural factor models. In addition, the extent to which situations dynamically modify behaviour requires careful elucidation — not in the gross and somewhat distorted approach of the situationalists where all behaviour is considered situationally determined, but more in the way that both Eysenck, Cattell, and the senior author have been maintaining for at least 20 years. Cattell, more than anyone has consistently stated that situations interact with trait process, these statements being ultimately presented as a mathematical model (Cattell, 1980). While Cattell's methods may be somewhat impractical with regard to the implementation of his behaviour/situation model, his underlying reasoning is impeccable.

Tentatively, we would advocate three 'new' methodologies for exploring the problem of dynamic process representation. The first being that already implemented by Sternberg (1977) in his componential process study of ability factors. The second, implemented in a minor way by Colby *et al.* (1971), is that subsumed under the nomenclature 'Artificial Personality'. In a similar way as artificial intelligence attempts to model reasoning and learning processes, so artificial personality may be expected to model behaviours and processes thought of as personality. The third methodology, already under investigation by us, is the representation of behaviour dynamics using a class of multivariate transfer function models generally used within time series analysis and forecasting. Explicit details of these models are provided in Box and Jenkins (1976) and Jenkins (1980).

(12) In conclusion, therefore, we argue that there is still a place for the factor

analysis of questionnaire data, but not as an end in itself. Rather it must be part of a larger quantified attack on the nature and meaning of personality.

REFERENCES

Bargmann, R. (1954) *The Significance of Simple Structure in Factor Analysis.* Physia Verlag, Wurzburg.

Barrett, P. and Kline, P. (1980a) The location of superfactors, P, E and N within an unexplored factor space. *Person. Indiv. Diff.* **1**, 239–247.

Barrett, P. and Kline, P. (1980b) Personality Factors in the Eysenck Personality Questionnaire. *Person. Indiv. Diff.* **1**, 317–333.

Barrett, P. and Kline, P. (1981a) The observation to variable ratio in factor analyses. *Personality Study and Group Behaviour,* **1** 23–33.

Barrett, P. and Kline, P. (1981b) Radial parcel factor analysis. *Person. Indiv. Diff.* **2**, 311–318.

Barrett, P. T. and Kline, P. (1982a) The itemetric properties of the Eysenck Personality Questionnaire: A reply to Helmes. *Person. Indiv. Diff.* **3**, 73–80.

Barrett, P. and Kline, P. (1982b) An item and radial parcel analysis of the 16 PF questionnaire. *Person. Indiv. Diff.* **3**, 259–270.

Barrett, P. T. and Kline, P. (1982c) Factor Extraction: An examination of Three Methods. *Person. Stud. Grp Behav.*, in Press.

Box., G. E. P. and Jenkins, G. M. (1976) *Time Series Analysis, Forecasting and Control.* (2nd Edn) Holden-Day, San Francisco.

Browne, J. A. and Howarth, E. (1977) A comprehensive factor analysis of personality questionnaire items: a test of 20 factor hypotheses. *Multiv. Behav. Res.* **12**, 399–427.

Buros, O. K. (1978) (Ed.) *VIIIth Mental Measurement Yearbook.* New Jersey, Gryphon Press.

Butcher, J. N. (1979) (Ed.) *New Developments in the use of the MMPI.* Univ. Minnesota Press, Minneapolis.

Carrigan, P. M. (1960) Extraversion-introversion as a dimension of personality: a reappraisal. *Psychol. Bull.* **57**, 329–360.

Carroll, J. B. (1981) Ability factors and processes. Paper at the NATO Conference on Cross-Cultural Psychology, Kingston, Ontario.

Cattell, R. B. (1957) *Personality and Motivation Structure and Measurement.* Yonkers, New York.

Cattell, R. B. (1966) The Scree test for the number of factors. *Multiv. Behav. Pres.* **1**, 140–161.

Cattell, R. B. (1972) The 16 PF and basic personality structure: A reply to Eysenck. *J. Behav. Sci.* **1**, 169–187.

Cattell, R. B. (1973) *Personality and Mood by Questionnaire.* Jossey-Bass. New York.

Cattell, R. B. (1978) *The Scientific Use of Factor Analysis in Behavioural and Life Sciences.* Plenum, New York.

Cattell, R. B. (1980) *Personality and Learning Theory.* Volumes I and II. Springer, New York.

Cattell, R. B. and Bolton, L. S. (1969) What pathological dimensions lie beyond the normal dimensions of the 16 *PF*? A comparison of MMPI and 16 PF factor domains. *J. Consult. clin. Psychol.* **33**, 18–29.

Cattell, R. B. and Child, D. (1975) *Motivation and Dynamic Structure.* Holt, Rinehart, Winston, London.

Cattell, R. B. and Delhees, K. H. (1973) Seven missing normal personality factors in the questionnaire primaries. *Multiv. Behav. Res.* 173–194.

Cattell, R. B. and Foster, M. J. (1963) The rotoplot program for multiple, single plane, visually-guided rotation. *Behav. Sci.* **8**, 156–165.

Cattell, R. B. and Gibbons, B. D. (1968) Personality structure of the combined Guilford and Cattell personality questionnaires. *J. Person. Soc. Psychol.* **9**, 107–20.

Cattell, R. B. and Kline, P. (1977) *The Scientific Analysis of Personality and Motivation.* Academic Press, London.

Cattell, R. B. and Muerle, J. L. (1960) The 'Maxplane' program for factor rotation to oblique simple structure. *Educ. psychol. Measmt.* **20**, 569–590.

Cattell, R. B. and Vogelmann, S. (1977). A Comprehensive trial of the Scree and KG criteria for determining the number of factors. *Multiv. Behav. Res.* **12,** 289–325.

Cattell, R. B. and Warburton, F. W. (1965) *Objective Personality and Motivation Tests.* U. Illinois Press, Champaign.

Cattell, R. B. and White, P. O. (1962) The method of confactor rotation illustrated in the oblique case on the cups of coffee plasmode.Adv. pub. 17. *Lab. Person. Assess. Gr. Behav.* U. Illinois.

Cattell, R. B., Eber, H. L. and Tatsuoka, M. M. (1970) *The 16 PF Test.* IPAT Champaign.

Cerny, B. A. and Kaiser, H. F. (1977) A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multiv. Behav. Res.* **12,** 43–7.

Colby, K. M., S. and Hilf, S. D. (1971) Artificial Paranoia. *Artif. Intelleg.* **2,** 1–25.

Comrey, A. L. (1970) *The Comrey Personality Scales.* San Diego. Educational and Industrial Testing Service.

Comrey, A. L. (1973) *A First Course in Factor Analysis.* Academic Press, New York.

Comrey, A. L. and Duffy, K. E. (1968) Cattell and Eysenck factor scores related to Comrey personality factor. *Multiv. Behav. Res.,* 379–392.

Cronbach, L. J. (1970) Essentials of Psychological Testing. Harper Row, New York.

Dahlstrom, W. G. and Welsh, G. S. (1960) *An MMPI Handbook.* Univ. Press, London, Oxford.

Delhees, K. H. and Cattell, R. B. (1971) The dimensions of pathology: proof of their projection beyond the normal 16 PF source traits. *Personality,* **2,** 149–171.

Delhees, K. H. and Cattell, R. B. (1971) *The Clinical Analysis Questionnaire.* IPAT, Champaign.

Eaves, L. and Eysenck, H. J. (1975) The nature of extraversion: a genetical analysis. *J. Person. soc. Psychol.* **32,** 102–112.

Eaves, L. J. and Eysenck, H. J. (1976) Genotype x age interactions for neuroticism. *Behav. Genet.* **6,** 359–362.

Eaves, L. J., Martin, N. G. and Eysenck, S. B. G. (1977) An application of the analysis of covariance structures to the psychogenetical study of impulsiveness. *Br. J. Math. Statist. Psychol.* **30,** 185–197.

Eysenck, H. J. (1947) *Dimensions of Personality.* Routledge, London.

Eysenck, H. J. (1954) *The Psychology of Politics.* Routledge & Kegan Paul, London.

Eysenck, H. J. (1956) The questionnaire measurement of neuroticism and extraversion. *Rev. di Psicol.* **50,** 113–140.

Eysenck, H. J. (1967) *The Biological Basis of Personality.* Springfield, C. C. Thomas.

Eysenck, H. J. (1970) *The Logical Basis of Factor Analysis.* Chapter 20 In : Jackson and Messick (1967) Ed.

Eysenck, H. J. (1978) Superfactors *P, E* and *N* in a comprehensive factor space. *Multiv. Behav. Res.* **13,** 475–481.

Eysenck, H. J. (1981) A biometrical-genetical analysis of impulsive and sensation-seeking behaviour. In: Zukerman, M. (Ed), 1981. *Biological Bases of Sensation-Seeking, Impulsivity and Anxiety.* Erlbaum, London.

Eysenck, H. J. (1981) (Ed.) *A model for Personality.* Springer, New York.

Eysenck, H. J. and Eysenck, S. B. G. (1964) *E.P.I. Man.* University of London Press, London.

Eysenck, H. J. and Eysenck, S. B. G. (1975) *The Eysenck Personality Inventory.* The University of London Press, London.

Eysenck, H. J. and Eysenck, S. G. B. (1976) *Psychoticism as a Dimension of Personality.* Routledge and Kegan Paul, London.

Eysenck, S. B. G. and Eysenck, H. J. (1963) On the dual nature of extraversion. *Br. J. Soc. clin. Psychol.* **2,** 46–55.

Eysenck, S. B. G. and Eysenck, H. J. (1978) Impulsiveness and venturesomeness: their position in a dimensional system of personality description. *Psychol. Rep.* **43,** 1247–1255.

Eysenck, S. B. G. and Eysenck, H. J. (1977) The place of impulsiveness in a dimensional system of personality description. *Br. J. Soc. clin. Psychol.,* **16,** 57–68.

Eysenck, S. B. G. and Zuckerman, M. (1978) The relation between sensation seeking and Eysenck's Dimensions of Personality. *Br. J. Psychol.,* **69,** 483–487.

Finney, J. C. (1961) The MMPI as a measure of character structure as revealed by factor analysis. *J. Counsell. Psychol.* **25,** 327–336.

Fulker, D. W., Eysenck, S. B. G. and Zuckerman, M. A genetic and environmental analysis of sensation-seeking. In Press.

Gorsuch, R.L. (1974) *Factor Analysis.* Saunders, Philadelphia.

Gottheil, E. (1965) An empirical analysis of orality and anality. *J. Nerv. Ment. Dis.* **141,** 308–317.

Grygier, T. J. (1961) *The Dynamic Personality Inventory.* N.F.E.R., Windsor.

Grygier, T.J. and Grygier, P. (1976) *The Dynamic Personality Inventory: Manual.* N.F.E.R., Windsor.

Guilford, J. P. (1956) *Psychometric Methods.* McGraw-Hill, New York.

Guilford, J. P. (1959) *Personality.* McGraw-Hill, New York.

Guilford, J. P. (1967) *Intelligence.* McGraw-Hill, New York.

Guilford, J. P. (1974) Rotation problems in factor analysis. *Psychol. Bull.* **81,** 498–501.

Guilford, J. P. (1975) Factors and factors of personality. *Psychol. Bull.,* **82,** 802–814.

Guilford, J. P. (1977) Will the real factor of extraversion-introversion please stand up? A reply to Eysenck. *Psychol. Bull.* **84,** 412–416.

Guilford, J. P. and Guilford, R. B. (1936) *J. Psychol.* **2,** 109–27.

Guilford, J. P. and Guilford, R. B. (1934) *J. Abnorm. Soc. Psychol.* **28,** 377-399.

Guilford, J. P. and Hoepfner, R. (1971) *The Analysis of Intelligence.* McGraw-Hill, New York.

Guilford, J. P. and Zimmerman. W. S. (1956) *Psychol. Monogr.* **70,** 1–26.

Guilford, J. S., Zimmerman, W. S. and Guilford, J. P. (1976) (Eds) *The Guilford Zimmerman Temperament Survey Handbook.* Sheridan, San Diego.

Guthrie, G. M., Jackson, P. N., Astilla, E. and Elwood, B. (1981) Personality Measurement: Do the scales have similar meanings in another culture? Paper at NATO Conference on *Human Assessment and Cultural Factors.* Ontario, Kingston.

Guttman, L. (1953) Image theory for the structure of quantitative variates. *Psychometrika* **18,** 277–296.

Guttman, L. (1954) Some necessary conditions for common factor analysis. *Psychometrika,* **19,** 149–161.

Hakstian, A. R. (1971) A comparative evaluation of several prominent methods of oblique factor transformation. *Psychometrika* **36,** 175.

Hakstian, A. R., Rogers, W. T., Cattell, R. B. (1982) The behaviour of number-of-factors rules with simulated data. *Multiv. Behav. Res.* **17,** 193–219.

Harman, H. H. (1976) *Modern Factor Analysis,* 2nd Edn University of Chicago Press, Chicago.

Harman, H. H. *et al.* (1975) Reports of a study of factor-referenced cognitive tests and temperament scales. E.T.S., Princeton.

Harris, C. W. and Kaiser, H. F. (1964) Oblique factor-analytic solutions by orthogonal transformations. *Psychometrika,* **29,** 347–362.

Hashemi, K. (1981) *The factor structure of the EPQ and 16 PF tests.* Unpub. PhD Thesis, Univ. Exeter.

Hathaway, S. R. and McKinley, J. C. (1951) *The Minnesota Multiphasic Personality Inventory Manual (Revised).* The Psychological Corp, New York.

Heim, A. W. (1975) *Psychological Testing.* Oxford University Press, London.

Helmes, E. (1980) A psychometric investigation of the Eysenck Personality Questionnaire. *Appl. Psychol. Measmt* **4,** 43–55.

Hendrickson, A. E. and White, P. O. (1966) A method for the rotation of higher-order factors. *Br. J. Math. Statist. Psychol.* **19,** 97–103.

Hogan, R. (1978) The Personality Research Form. In: Buros, O.K. (Ed.)

Holley, J. W. (1973) *The Rorschach.* In: Kline, P. (1973) (Ed.)

Horn, J. and Knapp, J. R. (1973) On the subjective character of the empirical base of Guilford's Structure of Intellect model. *Psychol. Bull.* **80,** 33–43.

Horn, J. and Knapp, J. R. (1974) Thirty wrongs do not make a right: reply to Guilford. *Psychol. Bull.* **81,** 502–504.

Howarth, E. (1971) Howarth Personality Questionnaire 2. University of Alberta Press, Edmonton.

Howarth, E. (1976) Were Cattell's 'Personality Sphere' factors correctly identified in the first instance? *Br. J. Psychol.* **67,** 2, 213–230.

Howarth, E. (1978) The Comrey Personality Inventory. In: Buros, O. K. (Ed.)

Howarth, E. (1980) *Howarth Personality Questionnaire.* University of Alberta Press, Alberta.

Howarth, E. and Browne, J. A. (1971b) An item factor analysis of the 16 PF questionnaire. *Person. Int. J.* **2**, 117–139.

Hunter, S., Overall, J. E. and Butcher, J. N. (1974) Factor structure of the MMPI in a psychiatric population. *Multiv. Behav. Res.* **9**, 283–301.

Hurley, J. R. and Cattell, R. B. (1962) The Procrustes program: producing direct rotation to test a hypothesised factor structure. *Behav. Sci.* **7**, 258–262.

Jackson, D. N. (1974) *Personality Research Form.* Research Psychol. Press, New York.

Jackson, D. N. and Messick, S. (1961) Content and Style in Personality Assessment. *Educ. Psychol. Measmt* **21**, 771–790.

Jackson, D. N. and Messick, S. (1967) *Problems in Human Assessment.* McGraw-Hill, New York.

Jackson, D. N. and Paunonen, S. V. (1980) Personality Structure and Assessment. *A. Rev. Psychol.* **31**, 503–551.

Jenkins, G. M. (1979) Practical Experiences with Modelling and Forecasting Time Series. G. Jenkins and Partners (Overseas), Jersey.

Jennrich, R. I. and Sampson, P. F. (1966) Rotation for simple loadings. *Psychometrika* **31**, 313–323.

Joreskog, K. G. (1969) A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika,* **34**, 183–202.

Joreskog, K. G. and Lawley, D. N. (1968). New Methods in maximum likelihood factor analysis. *Br. J. Math. Statist. Psychol.,* **21**, 85–97.

Kaiser, H. F. (1960) The application of electronic computers to factor analysis. *Educ. Psychol. Measmt.* **20**, 141–151.

Kaiser, H. F. (1963) Image Analysis. In: *Problems in Measuring Change* (Ed.) Harris, C. W. Univ. Wisconsin Press, Wisconsin.

Kaiser, H. F. (1965) Psychometric approaches to factor analysis. *Proceedings of the 1964 Invitational Conference on Testing Problems.* E.T.S., Princeton.

Kameoka, V. and Sine, L. (1978) An extension of Bargmann's tables and a computer programme for testing the statistical significance of simple structure. Quoted by Cattell, 1978.

Kline, P. (1967) Obsessional traits and emotional instability in a normal population. *Br. J. med. Psychol.* **40**, 193–197.

Kline, P. (1971) *Ai3Q Test.* National Foundation for Educational Research, Windsor.

Kline, P. (1973) (Ed.) *New Approaches in Psychological Measurement.* Wiley, Chichester.

Kline, P. (1978) *OOQ and OPQ Personality Tests.* Exeter Department of Psychology, Univ. Exeter.

Kline, P. (1979) *Psychometrics and Psychology.* Academic Press, London.

Kline, P. (1980) Models of Man: the Psychometric model. In: *Models of Man*, Chapman, A. J. and Jones, D. (Ed.). B.P.S., Leicester.

Kline, P. (1981) *Fact and Fantasy in Freudian Theory.* (2nd Edn) Methuen, London.

Kline, P. and Storey, R. (1977) A factor analytic study of the oral character. *Br. J. Soc. clin. Psychol.* **16**, 317–328.

Kline, P. and Storey, R. (1978) The Dynamic Personality Inventory: What does it measure? *Br. J. Psychol.* **69**, 375–383.

Kos, M. P. (1979) MMPI item content: recurring issues. Chapter 1 In: Butcher, J. N. (1979) (Ed.)

Kragh, V. and Smith, G. (1970) Percept-Genetic Analysis. Gleerups, Lund.

Lazare, A., Klerman, G. L. and Armor, D. J. (1966) Oral, obsessive and hysterical personality patterns: an investigation of psychoanalytic concepts by means of factor analysis. *Arch. Gen. Psychiat* **14**, 624–30.

Lee, H. B. and Comrey, A. L. (1979) Distortions in a commonly used factor analytic procedure. *Multiv. Behav. Res.* **14**, 301–321.

Levonian, E. (1961a) A statistical analysis of the 16 Personality Factor Questionnaire. *Educ. Psychol. Measmt,* **21**, 589–596.

Levonian, E. (1961) Personality Measurement with items selected from the 16 PF questionnaire. *Educ. Psychol. Measmt,* **21**, 937–946.

Loo, R. (1979) A psychometric investigation of the Eysenck Personality Questionnaire. *J. Person. Assess.* **43**, 54–58.

McNair, D. M. (1978) The Clinical Analysis Questionnaire. Vol. 1, p. 745–747. In: Buros, O. K. (1978).

Mulaik, S. A. (1972) *The Foundations of Factor Analysis.* McGraw Hill, New York.
Murray, H. A. (1938) *Explorations in Personality.* Oxford Univ. Press, New York.
Nesselroade, J. R. and Baltes, P. B. (1975) Higher-order convergence of two distinct personality systems: Cattell's HSPQ and Jackson's PRF. *Multiv. Behav. Res.* **10,** 387–408.
Nunnally, J. (1978) *Psychometric Theory.* McGraw-Hill, New York.
Orme, J. E. (1965) The relation of obsessional traits to general emotional instability. *Br. J. med. Psychol.* **38,** 269.
Overall, J. E., Hunter, S. and Butcher, J. N. (1973) Factor structure of the MMPI — 168 in a psychiatric population. *J. Consult. Clin. Psychol.* **41,** 284–286.
Rocklin, T. and Revelle, W. (1981) The measurement of extraversion: a comparison of the Eysenck Personality Inventory and the Eysenck Personality Questionnaire. *Br. J. soc. Psychol.* **20,** 279–284.
Saville, P. (1973) The standardisation of an adult personality inventory on the British population. *Bull. Br. psychol. Soc.* **26,** 25–29.
Saville, P. and Blinkhorn, S. (1976) *Undergraduate Personality by Factored Scales.* N.F.E.R., Windsor.
Saville, P. and Blinkhorn, S. (1982) Reliability, homogeneity and the construct validity of Cattell's 16 PF. *Person. Indiv. Diff.* **2,** 325–333.
Sells, S. B., Demaree, R. G. and Will, D. P. (1970) Dimensions of Personality: 1. Congruent factor structure Guilford and Cattell trait markers. *Multiv. Behav. Res.* **5,** 391–422.
Sternberg, R. J. (1979) *The nature of mental abilities. Am. Psychol.,* **34,** 214–230.
Stringer, P. (1970) A note on the factorial structure of the Dynamic Personality Inventory. *Br. J. med. Psychol.* **43,** 95–103.
Stringer, P. (1976) Chapter on the Factor Structure of the DPI in Grygier.
Thurstone, L. L. (1947) *Multiple Factor Analysis A Development and Expansion of Vectors of the Mind.* Univ. of Chicago Press, Chicago.
Velicer, W. F. (1976a) The relation between factor score estimates, image scores and principal component scores. *Educ. Psychol. Measmt,* **36,** 149–159.
Velicer, W. F. (1976b) Determining the number of components from the matrix of partial correlations. *Psychometrika* **41,** 321–327.
Velicer, W. F. (1977) An empirical comparison of the similarity of principal components, image and factor patterns. *Multiv. Behav. Res.* **12,** 3–22.
Vernon, P. E. (1964) Personality Assessment. Methuen, London.
Verster, J. M. (1981) The structure, organisation and correlates of cognitive speed and accuracy. A cross-cultural study using computerised tests. Paper at NATO conference, Human assessment and cultural factors. Kingston, Canada.
Wakefield, J. A. Yom, B. H. L. Bradley, P. E. Doughtie, E. B. Cox, J. A. and Kraft, I. A. (1974) Eysenck's Personality dimensions a model for the MMPI. *Br. J. Soc. clin. Psychol.* **13,** 413–420.
Zuckerman, M. (1974) The sensation-seeking nature. In: Maher, B. (Ed.) *Progress in Experimental Personality Research,* Vol. 7, Academic Press, New York.