PERGAMON

Personality and Individual Differences 26 (1999) 175–186

PERSONALITY AND INDIVIDUAL DIFFERENCES

Rejoinder to: The Eysenckian personality structure: A 'Giant Three' or 'Big Five' model in Hong Kong?[1]

Paul Barrett*

The State Hospital, Department of Psychology, Carstairs, Lanark ML11 8RP, Scotland, U.K.

Received 16 February 1998

Abstract

In a response to the Ng, Cooper, and Chandler (1998) paper on the structure of the JEPQ in Hong Kong children, three issues were discussed that cast doubt upon the conclusions reached by Ng et al. These issues concern firstly, the subjective adoption of factor names established in one questionnaire to describe factors of items drawn from a different questionnaire. Secondly, the lack of sample adequacy verification procedures, and thirdly, the forming and testing of hypotheses for which insufficient data and variables existed to permit scientifically adequate tests to be made. Further, using some new empirical analyses of the U.K. reference sample EPQ and JEPQ datasets, in addition to using the original Eysenck and Chan (1982) JEPQ Hong Kong dataset, it was shown that the Ng et al. 4 factor solution shared less similarity with the original Hong Kong factors than it did with the U.K. JEPQ factors. Psychometric indices were also reported that underlined the robust qualities of the U.K. 4 factor model for the EPQ series of questionnaires, *given* the exploratory factor analytical model within which the EPQ questionnaires were designed.

Key words: JEPQ; EPQ; Five Factor Model; Personality; Psychometrics

1. Introduction

The paper by Ng et al. (1998) requires a considered response, addressed not so much toward their results, but more toward the investigative approach that they have adopted, given their specific hypotheses. I think it is important to focus on the investigative approach given that what the authors have attempted in their paper is becoming an increasingly common approach to structural psychometrics amongst many 'applied' personality questionnaire analysts. Let me take each of their hypotheses in turn.

* E-mail: pbarret@liv.ac.uk

[1] "The Eysenckian personality structure: A 'Giant Three' or 'Big Five' model in Hong Kong" appeared in *Personality and Individual Differences* Vol. 25, No. 6. pp. 1111–1131 (1998).

PII: S0191–8869(98)00107–X

1.1. Hypothesis 1. Is the Eysenckian System really a Giant Three model plus a Lie scale, as proposed by Eysenck or is it, rather, a Big Five factor model?

This hypothesis has been addressed variously by many investigators, as Ng et al. note in the introduction to their paper. However, I find Cattell's recent statement in a 1995 paper particularly illuminating, "there are no proofs that various five factor theorists are actually mutually finding the *same* five factors". Ng et al. adopt the same practice as many investigators who seem quite unable to avoid the temptation of naming factors based solely on the basis of subjective interpretation of item content. Cattell's statement brings home the fallacy of this procedure *when trying to locate factors into a pre-existing explanatory framework*. If an investigator wishes to determine whether a factoring of data is producing equivalent factors as given by a target structure, then the only scientifically justifiable solution to this problem is to include the target questionnaire or substantial numbers of marker items or item parcels from the target factor structure, along with the new questionnaire items to be located in that factor space, and *then* attempt to make quantifiable statements about the proposed similarity and meaning to be attributed to the factors. Using composite scales as markers is not optimal, as the level of item detail is lost when using a scale score. Yes, this demands more test-taking time of respondents, and yes, it requires sensitive multivariate configurational analysis methods, but if the application of psychometrics is to conform to that of general scientific methodology, then there can be no compromise. If investigators wish to 'fit' a Big Five factor model, then the first step is to define the model they wish to fit, not by verbal description or semantic analysis, but by mathematical description. For example, Costa and McCrae's NEO has considerable published data on its definitive factor structure (loadings, facet correlations, and factor relationships). These mathematical constraints define the model that is produced from the items, and it is these same constraints that an investigator must utilise in any investigation where they wish to demonstrate meaningful equivalence. A secondary method of 'scale equivalence' might be to simply correlate the NEO scale scores with another test's scale scores, looking for identity (allowing for unreliability) between both the paired scale correlations, and perhaps between the correlation matrices between both within-test correlation matrices. The correlation pattern testing hypothesis procedure of Steiger (1980) would be a suitable method for this latter analysis.

Ng et al. proceed to test their hypothesis solely upon the basis of their semantic understanding of a generic five factor model. However, their sample size is very low, only 202 respondents. This is not an optimal sample size for these kinds of factor analysis, given the implications (Guadagnoli and Velicer, 1988) of the previously known low communality solution of the JEPQ questionnaire. Also, as Barrett and Kline (1981) demonstrated, it is possible to approximately recover the 4 factor 90-item EPQ factor structure in adults, using just 100 respondents, but this was not by any means perfect. Further, this is a cross-cultural study. From detailed examination of the Eysenck and Chan (1982) data, it is clear that some perturbation in the loading structure for the 4 EPQ scales might have been expected in any new sample. In short, their sample size is unlikely to be sufficient to adequately answer their rather adventurous hypothesis above, in that any model misfit might be due either to cultural specificities and/or insufficient numbers of respondents required to define clearly the proposed factor structure.

However, this brings me to my next point, Ng et al. make no attempt to quantifiably determine the fit between their 4 factor solution and that expected from the original English data, using either the U.K. JEPQ factor structure or a 1,0 target factor matrix that corresponds to the U.K. scorekey. It would seem desirable to first examine the Hong Kong data for rotational fit to the target U.K. data, before attempting to conclude that it does not fit. This is especially important if, given their hypothesis above, they also wish to make statements as to the factorial structure of the JEPQ in Hong Kong. Fortunately, Ng et al. provide the varimax rotated factor matrix for their combined sex data. This allowed me to compare their factor matrix to the original U.K. JEPQ combined sex reference sample factor structure, as well as implement an orthogonal procrustes target fit solution (akin to the validimax procedure proposed by McCrae and Costa, 1989) using the U.K. scorekey as the target model 1,0 keying. The U.K. combined sex sample consists of 444 females and 510 males (954 in total) with a mean age of 11.91 and *SD* of 1.44 years. For clarity of reading, all matrices were computed from scored JEPQ data, using the U.K. scorekey, with items subsequently re-ordered by scale into the order, P, E, N, and L. The matrix provided by Ng et al. was converted into this format, and the mathematical sign of items reflected as appropriate to the score key.

The comparison matrices reported below use the modified Kaiser–Hunka–Bianchini (1971) congruential fit procedure (Barrett, Petrides, Eysenck, and Eysenck, in press), reporting congruence coefficients calculated between the maximally congruent Hong Kong and target factor patterns.

The results in Table 1 indicate that even under optimal target-fit procedures, the Hong Kong structure is not a particularly close fit to the U.K structure. The results in Table 2 indicate that it is the P factor that is the least comparable factor in the Hong Kong data. Explaining why this might be the case is problematic as indicated above, due to the confounding of low sample size in addition to the cultural aspects. However, since I also had available the Eysenck and Chan 1982 original Hong Kong JEPQ dataset, with a joint-sex sample size of $n = 1075$ (698 males and 377 females), I was able to conduct a factor comparison analysis that would indicate to what extent the Ng et al. results might be due to low sample size rather than any systematic cultural effect. To achieve this, the first comparison was made between Ng et al. varimax factor matrix as the target, and the Eysenck and Chan dataset factors (Table 3). The expectation here is that the two datasets should be very similar, given the samples are from the same cultural domain. Then, the U.K. reference sample varimax matrix was compared to the Eysenck and Chan dataset factors (Table

Table 1
Congruence coefficients between the maximally congruent Hong Kong factor solution compared with the U.K. JEPQ target 4 factor varimax solution (**P** = Psychoticism, **E** = Extraversion, **N** = Neuroticism, and **L** = Social Desirability). The mean solution cosine for this matrix is 0.84—indicating a less than optimal match between the two solutions. The columns of the matrix represent the Hong Kong factors, the rows are the U.K. target factors

	P	**E**	**N**	**L**
P	**0.87**	0.15	−0.05	0.24
E	0.17	**0.89**	0.08	0.12
N	−0.04	0.07	**0.81**	−0.08
L	0.25	0.11	−0.09	**0.79**

Table 2
Congruence coefficients between the maximally congruent Hong Kong factor solution compared with the U.K. scorekey 1.0 target matrix (**P** = Psychoticism, **E** = Extraversion, **N** = Neuroticism, and **L** = Social Desirability). The mean solution cosine for this matrix is 0.76—indicating a poor match between the two solutions. The columns of the matrix represent the Hong Kong factors, the rows are the target factors

	P	**E**	**N**	**L**
P	**0.61**	0.15	0.10	−0.18
E	0.12	**0.81**	−0.06	−0.01
N	0.10	−0.08	**0.81**	−0.10
L	−0.18	−0.01	−0.10	**0.82**

4). Finally, the U.K. score key target factor pattern was compared with the Eysenck and Chan dataset factors (Table 5).

What these three sets of results indicate is that the Ng et al. dataset is less similar overall to the original Eysenck and Chan dataset factors than it is to the U.K. reference sample data matrix (mean solution cosine of 0.82 vs 0.87). This is of interest as it might be expected that same-country samples would be more similar to one another than to the U.K. data. However, given 14 years or so have elapsed between the respective Hong Kong samples, it is also possible that social changes, as well as cultural changes have produced such a difference.

These results make it more probable that the Ng et al. sample is probably of insufficient size to permit an unambiguous exploration of where structural changes might be taking place. As the Eysencks have always maintained, such cultural analyses should also be carried out on each

Table 3
Congruence coefficients between the maximally congruent Eysenck and Chan (1982) Hong Kong factor solution compared with the Ng et al. (1998) Varimax rotated 4 factor target matrix (**P** = Psychoticism, **E** = Extraversion, **N** = Neuroticism, and **L** = Social Desirability). The mean solution cosine for this matrix is 0.82—indicating a poor match between the two solutions. The columns of the matrix represent the Eysenck and Chan Hong Kong factors, the rows are the target Ng et al. factors

	P	**E**	**N**	**L**
P	**0.90**	0.20	0.18	−0.10
E	0.23	**0.79**	0.32	−0.10
N	0.19	0.30	**0.82**	0.05
L	−0.11	−0.10	0.05	**0.78**

Table 4
Congruence coefficients between the maximally congruent Eysenck and Chan (1982) Hong Kong factor solution compared with the U.K. reference sample Varimax rotated 4 factor target matrix (**P** = Psychoticism, **E** = Extraversion, **N** = Neuroticism, and **L** = Social Desirability). The mean solution cosine for this matrix is 0.87—indicating a relatively poor match between the two solutions. The columns of the matrix represent the Eysenck and Chan Hong Kong factors, the rows are the target U.K. reference sample factors

	P	**E**	**N**	**L**
P	**0.81**	0.24	0.00	0.31
E	0.23	**0.93**	0.01	0.13
N	0.00	0.01	**0.91**	−0.09
L	0.31	0.13	−0.10	**0.84**

separate sex dataset. It is of interest to note that the male and female Eysenck and Chan datasets have a mean solution cosine of only 0.90. In contrast, the U.K. males vs females JEPQ factor comparison mean solution cosine is 0.94.

Assuming for one moment that the Ng et al. sample size was adequate, and that they obtained identical results to those reported in their paper, then moving to a 5 factor solution still seems quite arbitrary, given the arguments concerning the identification of a 5 factor model. Further, given the extremely clear and robust 4 factor solution in the U.K. data, and the experimental, biological, cross-cultural, and genetic evidence supporting these constructs in the adult version of the questionnaire, then any investigator would be most cautious in attempting to posit 5 factors

Table 5
Congruence coefficients between the maximally congruent Eysenck and Chan (1982) Hong Kong factor solution compared with the U.K. scorekey target matrix (**P** = Psychoticism, **E** = Extraversion, **N** = Neuroticism, and **L** = Social Desirability). The mean solution cosine for this matrix is 0.80—indicating a poor match between the two solutions. The columns of the matrix represent the Eysenck and Chan Hong Kong factors, the rows are the target U.K. scorekey matrix

	P	**E**	**N**	**L**
P	**0.72**	0.08	0.10	−0.19
E	0.06	**0.88**	0.02	0.00
N	0.90	0.02	**0.86**	−0.15
L	−0.19	0.00	−0.16	**0.74**

Table 6
Some relevant psychometric indices for both the UK JEPQ and EPQ reference sample, combined sex, datasets

	JEPQ $N = 954$	EPQ $N = 4140$
Coefficient Alphas		
Psychoticism (P)	0.71	0.72
Extraversion (E)	0.72	0.85
Neuroticism (N)	0.82	0.86
Social Desirability (L)	0.85	0.81
Scale Quality (SQUAL)		
Psychoticism (P)	0.56	0.73
Extraversion (E)	0.75	0.86
Neuroticism (N)	0.82	0.89
Social Desirability (L)	0.84	0.86
C-66% Test Complexity Index (TCI)	11.1%	3.3%
Factor Extraction Tests		
Armor's Theta*	4	4
Velicer MAP	4	4
Barrett & Kline AUTOSCREE	4/5	8/9
Rotation Signal-to-Noise (C-ANR)		
Psychoticism (P)	0.67	0.75
Extraversion (E)	0.79	0.88
Neuroticism (N)	0.85	0.89
Social Desirability (L)	0.70	0.84

*The bound value for Armor's Theta (Armor, 1971) was set at 0.5. That is, factors were retained if Theta was above 0.5, and rejected if below. Armor's Theta can be interpreted in the manner of an alpha reliability coefficient for a factor component.

for the EPQ questionnaire model. Mere psychometric manipulation of covariance seems quite arbitrary given some of the empirical and experimental evidence for the casual basis of these constructs. As Barrett and Kline (1980) showed previously, the EPQ can be split into several 'primaries', however, these fold back to 4 factors at the second order. Given we remain within the principal component, classical test theory, and exploratory factor analytical model of analysis, it is of interest to note some relevant psychometric facts about the U.K. JEPQ combined sex sample dataset, and those for the original U.K. EPQ combined reference sample dataset. Table 6 provides some key analytical features of both these datasets. As can be seen from the data in this table, it is very hard to find any quantitative evidence that would support a 5 factor solution for the JEPQ or EPQ.

The formulae and rationale of the various signal-to-noise and complexity indices provided in Table 6 are given in the Technical Appendix. These were originally devised for and reported in

Barrett et al.'s (1996) analysis of the Saville and Holdsworth OPQ questionnaire. They are a class of rather sensitive quantitative parameters that index item complexity, scale overlap, and the clarity with which a scale of items are represented within a multi-scale factor space. They are designed to be used as *psychometric exploratory data analysis* indices in the manner of Turkey's (1977) definition of exploratory data analysis within statistical analysis in general. The only reason they are reported here is that they confirm the clarity of the solution of the EPQ, both in the JEPQ and EPQ adult. The parameter sizes are similar to those found in the adult EPQR. Readers who would like to see the item analysis results and factor patterns upon which these data were computed are invited to access my web site, where these datasets are available for download/inspection (http://www.liv.ac.uk/~pbarrett/epq.htm). The PsWin programs that perform all the calculations made within this rejoinder are also available free from this website at http://www.liv.ac.uk/~pbarrett/programs.htm.

1.2. Hypothesis 2. Given the non-Western character of the population, is E unitary or dual in nature as claimed by a number of personality theorists?

This hypothesis cannot be answered by the data at hand, or the methods chosen for analysis. It is easily possible to 'split' the E factor, either into item parcels (akin to Comrey et al.'s (1968) and Comrey (1984) Factored Homogenous Item Dimensions) as was done by Barrett and Kline (1982) or into two subscales, as was shown by Barrett and Kline (1980). In fact, within the Eysenck Personality Profiler (Eysenck and Wilson, 1991), the *superfactor* Extraversion scale is hypothesised to consist of 7 'extraversion' primaries (although by no means empirically unambiguous; see Eysenck et al. (1992), and Costa and McCrae (1995)). Atheoretical 'factor splitting' is a by-product of the arbitrary application of psychometrics that has no basis in experimental psychology. Without a guiding model or theory, or any kind of causal model to justify a particular partitioning of variance within a covariance/correlation matrix, then the validity of extracting more or fewer factors from a matrix is surely suspect. Certain quantitative indices may be of assistance, but there are no coefficients available that will provide absolutely deterministic evidence of construct number. As is noted in the final paragraph of this paper, there are two other methods of item analysis that might be used for this purpose, but these introduce their own problems. Given the U.K. EPQ version E scale can be split, or not, dependent upon the whim of the investigator, it is concluded here that the Ng et al. hypothesis is actually quite arbitrary, and probably not a cross-cultural issue at all.

1.3. Hypothesis 3. The nature of L.

I agree with the authors that L is probably best considered as a dimension of personality, rather than simply a 'Lie' scale. However, I note that the Eysencks, over the past 15 years or so, have repeatedly referred to this scale as *social desirability*, not simply as a Lie scale. Perhaps the occupational term of 'impression management' is more suitable? I further note that Eysenck, M. W. (1997) uses this scale to help define the 'repressor' individual (see also Weinberger et al. (1979)). Either way, it matters little to the arguments at hand—the JEPQ and EPQ measure of L is relatively clear, factorially well-defined, and replicable across 34 adult cultural datasets (Barrett et al., in press).

1.4. Hypothesis 4. Is P a superfactor of the facets A and C?

Since the authors have not measured A and C, it is curious as to how they could possibly have investigated such a hypothesis. In my response to their hypothesis 1 above, the problem of investigators naming scales as though they had direct evidence of five factor equivalence between their 'new' scales and those provided by say the Costa and McCrae five factor NEO model was noted. However, if I permit myself the same latitude as most investigators who wish to use 5 factors as 'the' factors, then I can take advantage of this position and access a dataset that provides the appropriate scale correlations which might assist in answering such a hypothesis. These data were originally reported in Barrett and Paltiel (1993) and Kline and Barrett (1994), as part of an investigation into the location of two new questionnaires in a comprehensive personality and ability factor space. The data were acquired from 253 adult volunteers who completed a range of questionnaires, two of which were Kline and Lapham's (1991) Professional Personality Questionnaire (PPQ) and the EPQR. The PPQ was constructed to provide a brief but reliable measure of the Big Five factors suitable for us within occupational psychology. The correlations between EPQR-P and the PPQ A and C equivalent scales were -0.12 and -0.25 respectively. Whether these results indicate that P is a 'superfactor' of the facets A and C is largely a matter of subjective preference of the reader as the method of analysis and data at hand do not permit a proper test of this hypothesis. However, what they do indicate is that P is not *the unique opposite pole of A*, as Ng et al. conclude in their study. Of course, the EPQR is the revised P scale, and the data were from adults, so these results can only be taken as an indication of what might be expected within children—assuming of course that the NEO factors are themselves demonstrable in children.

Finally, with the recent advances in psychometric and statistical methodology of Structural Equation Modelling (SEM) and Item Response Test theory, it is of interest to consider that the analyses exemplified within Ng et al., and in my responses above, might now be virtually redundant as continuing methods of psychometric investigation. There are more precise and objective methods (see Wright, 1998) available now to an investigator who is interested in proposing questions that involve dimensionality of constructs, psychological models, and shared measurement variance (Bollen and Lennox, 1991). However, McCrae et al. (1996) have shown that using SEM invariably leads to one rejecting the original target model, as with the NEO Big Five factors. This result should serve as a warning to investigators that 'established' personality models founded upon older exploratory factor analytical methods of analysis perhaps require careful psychometric revision that is both concordant with the available experimental evidence for the validity of the constructs, and with the aim of creating a more mathematically precise partitioning of the so-called 'superfactors' into their component sub-facets.

Technical appendix

Item complexity: C-66%-ISNR

The computational formula is:

$$\text{C-66\%-ISNR} = \left[\frac{ITC_i^2}{ITC_i^2 + \left\{ \sum_{j \neq i}^{N} nsITC_j^2 \bigg/ (N-1) \right\}} \right] \Bigg/ K$$

where: $ITC =$ the keyed-scale item-total correlation i. It is the item-total correlation for an item within the scale in which that item is identified as being a member; $nsITC =$ the correlation between an item and a scale score on which it is assumed not to be associated (non-keyed); $N =$ the number of scales in the test; $K =$ the number of non-keyed scale ITCs $\geqslant$ 66% of the size of the keyed scale ITC.

Thus for each item, this parameter indexes the ratio of squared 'keyed scale' item-total correlation (ITC) to the squares of the non-keyed ITCs, modifying this ratio by dividing it by the number of 'salient' non-keyed ITCs. This correction is required since as the number of scales increases in a test, the effect of one or two high correlations across other scales can be swamped in the calculation of a mean value. 'Salient' is defined as those correlations greater than or equal to two-thirds of the mean ITC for a keyed scale, with a hard lower bound of 0.15 (i.e. if two thirds of the size of the keyed ITC is less than 0.15, then it is set to 0.15). In other words, a subjective decision is made here in deciding that a non-keyed ITC is critical when its size is greater than this value. Essentially this measure treats as a 'signal' the keyed-scale ITC, and 'noise' as the remaining correlations across the non-keyed scales that are at least two-thirds the size of the keyed-scale ITC. The **C-66%-ISNR** parameter varies between 0.0 and 1.0, and is expressed as a %, with 100% indicating maximum possible complexity.

SQUAL

The **S**cale **QUAL**ity index. This parameter is an attempt to provide a single parameter that indexes the *measurement quality* of a scale of items as a whole, taking into account scale-item complexity, the signal-to-noise ratio of the scale, and the disparity of ITCs below the mean ITC within the scale. In essence it is an attempt to capture the many essential psychometric properties of a scale of items as a unitary parameter. The possible values for this parameter vary between 0 and 1.0, with 1.0 indicating perfect scale quality. The formula is:

$$\mathrm{SSNR} = \frac{\left(\sum_{i=1}^{S} ITC_i^2/S\right)}{\left(\sum_{i=1}^{S}/S\right)+\left(\sum_{j=1}^{NS} nsITC_j^2/NS\right)} \qquad \text{C-SSNR} = \mathrm{SSNR}-(\mathrm{SSNR}(K/S))$$

$$\mathrm{CB} = 1.0-\left(1.0-\left(\frac{\sum_{i=1}^{NB} C\text{-}66\%\text{-}ISNR_i < 0.5}{NB}\right)\right)*\left[\frac{NB}{S}\right]$$

$$\mathrm{CR} = 1.0-\left(1.0-\left(\frac{\sum_{i=1}^{NB} ITC_i < BoundValue}{(NL * BoundValue)}\right)\right)*\left[\frac{NL}{S}\right]$$

$$\mathrm{SQUAL} = \text{C-SSNR} * CB * CR$$

where: ITC = the keyed-scale item-total correlation i. The item-total correlation for an item for the scale in which that item is identified as being a member; $nsITC$ = the correlation between an item and a scale score on which it is assumed not to be associated (non-keyed); S = the number of items in a target scale; NS = the number of non-keyed-scale items in the test. i.e. the number of items remaining in the test after those in the target scale are excluded; K = the number of items which correlate $\geqslant$ the specified value of the mean target scale ITC; NB = the number of items whose C-66%-ISNR values is less than 0.5 in a scale; NL = the number of items in a scale whose ITC is less than the mean ITC for that scale; CB = the correction factor for high complexity items; CR = the correction factor for low-ranging ITC disparity; $BoundValue$ = the mean ITC for a scale.

Thus for each scale, the ratio of mean squared ITCs to mean squared non-keyed ITCs is indexed as a Scale Signal-to-Noise Ratio (SSNR). Once again, a correction is applied based upon the identification of salient correlations that are greater than or equal to a specified bound value. For the EPQ analyses, the specified bound value was the mean ITC for each scale. The scale SNR is corrected by treating the salients as of equal 'signal' strength to a keyed item ITC. The logic of this is that if 4 'external' items correlate as highly with the scale score as do the 8 items within the scale, then the C-SSNR parameter would indicate a 50% level of 'noise' in discriminating the scale from the remainder of the test items. Two other correction factors are then applied to the C-SSNR parameter, the first (CB) is a correction based upon the relative size of the C-66%-ISNR coefficients in a scale that are less than 0.5 in size. The correction is then weighted by the number of these 'bad items' in order to provide some degree of sensitivity. The second correction parameter (CR) is also a weighted factor that indexes the relative disparity in ITCs below the mean ITC in a scale. This coefficient is sensitive to low ITCs within a scale that may itself contain many high ITCs. The SQUAL parameter is thus a complex function of signal-to-noise, item complexity, and ITC disparity within a scale. The term 'quality of measurement' has been chosen to best represent the meaning to be attributed to this complex parameter. The parameter is obviously not capable of determining the *utility* of measurement made by a scale of items.

Test Complexity Index (TCI)

This is a summary parameter that attempts to describe the complexity of a test as a single numerical index. It is computed by summing the number of C-66%-ISNR coefficients with values of less than 0.5 (less than 50% 'signal' in an item) and dividing this sum by the number of items in the test. This value is expressed as a percentage and provides another summary parameter indexing the discriminability of test items within a test as a whole. The TCI parameter varies between 0.0 and 1.0, and is expressed as a percentage, with 100% indicating maximum possible test complexity.

Factorial Absolute Signal-to-Noise Ratio (ANR)

This coefficient is based upon Fleming's (1985) measure of the index of fit for factor scales. The scale signal-noise-ratio (SSNR) coefficient, as detailed above, is a direct analogue and use of Fleming's formula, where item-total correlations were defined as the values to be squared. Fleming used factor loadings as the basis for his signal-to-noise ratio. However, in the factor analytic domain, it was decided to modify Fleming's formula by using absolute value loadings rather than squared loadings (the absolute noise ratio). This provides a closer analogue as to how a user interprets columns of factor loadings and is generally more sensitive to the size ratio of salient and non-salient loadings. The coefficient is modified for the same reasons, and using the same formula as for the SSNR parameter above. It must be reiterated that the use of the uncorrected Fleming formula is of little value. As the ratio of the number of items in a scale to the number remaining in the test becomes larger, the sensitivity of the coefficient to index any useful information decreases correspondingly. The correction 'bound' value for the absolute noise ratio parameter uses the conventional lower bound of 0.3 for treating a factor loading as 'significant'. Non-scale items which load equal to or greater than this value on a scale factor are summed as 'significant' non-salients. Thus the *C-ANR* parameter is the direct factorial equivalent of the *C-SSNR* parameter. Finally, an additional correction is made to the *C-ANR* parameter that indexes the number of items within a scale that load less than 0.3 on the scale factor. This correction has to be implemented in order to adjust for the specific case where only some of the keyed items for a scale actually load significantly on a factor. Given no other items load significantly on this factor, it is possible to still maintain a high signal-to-noise ratio even though maybe only half the number of keyed items load above 0.3. Thus, a correction is applied in the same way as that for the SQUAL parameter, using only the CR parameter described above, with loadings replacing the ITC values, and the *BoundValue* replaced with a constant of 0.3. The signal-to-noise parameter is adjusted for non-keyed salient loadings and for the quantity of keyed items loading less than 0.3 on the scale factor. The absolute noise ratio is thus corrected for non-keyed items loading too highly on a factor, as well as correcting for keyed items loading too low. The possible values for this parameter vary between 0 and 1.0, with 1.0 indicating maximum signal to noise. C-ANR values above 0.7 are invariably generated by item factors where most, if not all the keyed items load greater than 0.3 and are the only such loadings on a factor.

Acknowledgement

The author would like to gratefully acknowledge Dr S. B. G. Eysenck for permission to use the Eysenck reference sample data for the analyses reported in this paper.

References

Armor, D. J. (1971). Theta reliability and factor scaling. In H. L. Costner (Ed.) *Socialogical Methodology*. San Francisco: Jossey-Bass.

Barrett, P. T., & Kline, P. (1980). Personality factors in the Eysenck Personality Questionnaire. *Personality and Individual Differences*, *1*, 317–333.

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

Barrett, P. T., & Kline, P. (1982). Radial parcel factor analysis. *Personality and Individual Differences*, *2*, 311–318.

Barrett, P. T., Kline, P., Paltiel, L., & Eysenck, H. J. (1996). An evaluation of the psychometric properties of the Concept 5.2 OPQ Questionnaire. *Journal of Occupational and Organisational Psychology*, *69*(1), 1–19.

Barrett, P. T., & Paltiel, L. (1993). The Concurrent and Structural Validity of the 15 Factor Personality Questionnaire. *International Journal of Recruitment Selection and Retention*, *2*(3), 30–36.

Barrett, P. T., Petrides, K. V., Eysenck, S. B. G., & Eysenck, H. J. (in press). The Eysenck Personality Questionnaire: An examination of the factorial similarity of P, E, N, and L across 34 countries. *Personality and Individual Differences*.

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110*(2), 305–314.

Cattell, R. B. (1995). The fallacy of five factors in the personality sphere. *The Psychologist*, *May*, 207–208.

Comrey, A. L. (1984). Comparison of two methods to identify major personality factors. *Applied Psychological Measurement*, *8*(4), 397–408.

Comrey, A. L. Jamison, K., & King, N. (1968). Integration of two personality factor systems. *Multivariate Behavioral Research*, *3*(2), 147–159.

Costa, P. T., & McCrae, R. R. (1995). Primary traits of Eysenck's P-E-N system: three and five factor solutions. *Journal of Personality and Social Psychology*, *69*(2), 208–317.

Eysenck, H. J., Barrett, P. T., Wilson, G. D., & Jackson, C. (1992). Primary trait measurement of the 21 components of the P-E-N system. *European Journal of Psychological Assessment*, 8, 109–117.

Eysenck, H. J., & Wilson, G. D. (1991). *The Eysenck Personality Profiler*. London: Corporate Assessment Network Ltd.

Eysenck, M. W. (1997). Anxiety and Cognitive Processes. In C. Cooper and V. Varma, (Eds.), *Processes in Individual Differences*. London: Routledge.

Eysenck, S. B. G., & Chan, J. (1982). A comparative study of personality in adults and children: Hong Kong vs England. *Personality and Individual Differences*, *3,* 153–160.

Fleming, J. S. (1985). An index of fit for factor scales. *Educational and Psychological Measurement*, *45,* 725–728.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, *103*(2), 265–275.

Kaiser, H. K., Hunka, S., & Bianchini, J. C. (1971). Relating factors between studies based upon different individuals. *Multivariate Behavioral Research*, 409–422.

Kline, P., & Barrett, P. T. (1994). Studies with the PPQ and the Five Factor Model of Personality. *European Review of Applied Psychology*, *44*, 35–42.

Kline, P., & Lapham, S. L. (1991). The validity of the PPQ: a study of its factor structure and its relationship to the EPQ. *Personality and Individual Differences*, *12*, 631–635.

McCrae, R. R., & Costa, P. T. (1989). Rotation to maximise the construct validity of factors in the NEO personality inventory. *Multivariate Behavioral Research*, *24*, 107–124.

McCrae, R. R., Zonderman, A. B., Bond, M. H., Costa, P. T., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO Personality Inventory: Confirmatory factor analysis versus procrustes rotation. *Journal of Personality and Social Psychology*, *70*(3) 552–566.

Ng, H. S., Cooper, M., & Chandler, P. (1998). The Eysenckian Personality Structure: a 'Giant Three' or 'Big Five' Model in Hong Kong? *Personality and Individual Differences, 25*, 1111–1131.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*(2), 245–251.

Tukey, J. D. (1977). *Exploratory Data Analysis*. New York: Addison-Wesley.

Weinberger, D. A., Schwartz, G. E., & Davidson, J. R. (1979). Low anxious, high anxious, and repressive coping styles: psychometric patterns and behavioral and physiological responses to stress. *Journal of Abnormal Psychology*, *88*, 369–380.

Wright, B. (1998). *Fundamental Measurement for Psychology*. Rasch Measurement Transactions, Memo 64. Chicago: University of Chicago Mesa Press.