Which version of the Eysenck Personality Profiler is best?
6-, 12- or 20-items per scale ☆

Leslie J. Francis a,*, Chris J. Jackson b

a Centre for Ministry Studies, University of Wales, Normal Site, Bangor, Gwynedd LL57 2PX, UK
b School of Psychology, University of Queensland, Brisbane, QLD 4072, Australia

Received 9 December 2002; received in revised form 19 January 2004; accepted 12 March 2004
Available online 14 May 2004

Abstract

Data provided by 400 first year undergraduate students were analysed to develop two short forms of the Eysenck Personality Profiler (EPP) in which each of the 22 primary scales is assessed by a 6-item and a 12-item version instead of the usual 20-item per scale measure. In comparison with the 6-item per scale measure, the 12-item version retains more of the characteristics of the long version and seems a good compromise between quality of data and administration time.

Keywords: EPP; Eysenck Personality Profiler; Short EPP; Eysenck

1. Introduction

Originally, Eysenck’s dimensional model of personality concentrated on identifying and measuring a small number of higher order orthogonal factors (Eysenck & Eysenck, 1985). The Maudsley Personality Inventory developed by Eysenck (1959) and the Eysenck Personality Inventory developed by Eysenck and Eysenck (1964b) maintained that individual differences could be most economically and adequately expressed in terms of just two orthogonal higher order factors: extraversion and neuroticism. More recently the Eysenck Personality Questionnaire developed by Eysenck and Eysenck (1975) and the Revised Eysenck Personality Questionnaire

☆ All versions of the Eysenck Personality Profiler (EPP) are available from Chris J. Jackson (chrisj@psy.uq.edu.au).
* Corresponding author. Tel.: +44-124-838-2566; fax: +44-124-838-3954.
E-mail addresses: l.j.francis@bangor.ac.uk (L.J. Francis), chrisj@psy.uq.edu.au (C.J. Jackson).
developed by Eysenck, Eysenck, and Barrett (1985) expanded the model to embrace three orthogonal higher order factors: extraversion, neuroticism and psychoticism. While neuroticism remained almost unchanged in the development from the two dimensional model to the three dimensional model, extraversion emerged somewhat modified (Rocklin & Revelle, 1981).

In a popular book published in the 1970s under the title *Know Your Own Personality*, Eysenck and Wilson (1975) attempted to make the three higher order personality dimensions of extraversion, neuroticism and psychoticism more accessible by defining and attempting to measure seven personality traits underlying each of the three dimensions of personality. Building on this work, Eysenck, Barrett, Wilson, and Jackson (1992) developed the Eysenck Personality Profiler as an instrument capable of measuring 21 personality traits. Three important changes, however, took place in the transition between the instrument proposed by Eysenck and Wilson (1975) and the instrument proposed by Eysenck et al. (1992). First, the number of items used to assess each trait was reduced from 30 to 20, reducing the overall number of items required to measure the 21 traits from 630 to 420. Second, the polarity of some of the scales was reversed in order to make the high scoring poles always consistent with the directions of extraversion, neuroticism, and psychoticism. Third, the scales clustering within the dimensions of extraversion and psychoticism were reassessed. In other words, some of the scales originally located within extraversion were reassigned to psychoticism, while some of the scales originally located within psychoticism were reassigned to extraversion.

According to the Eysenck Personality Profiler, the seven personality traits comprising extraversion were defined as activity, sociability, expressiveness, assertiveness, achievement orientation, dogmatism, and aggressiveness. The seven personality traits comprising neuroticism were defined as low self-esteem, unhappiness, anxiety, dependency, hypochondriasis, guilt, and obsessiveness. The seven personality traits comprising psychoticism were defined as risk-taking, impulsiveness, irresponsibility, manipulation, sensation-seeking, masculinity, and unreflectiveness.

Although a relatively recent addition to the Eysenck family of instruments, a series of publications has already emerged on the application of the Eysenck Personality Profiler. For example, personality profiles have been generated on performing artists (Marchant-Haycox & Wilson, 1992), bikers (Jackson & Wilson, 1993) and physicists (Wilson & Jackson, 1994). The Eysenck Personality Profiler has been employed in studies concerned with group obsessiveness as a moderator of dissimulation (Jackson & Wilson, 1994), the relationship between personality and intelligence (Furnham, Forde, & Cotter, 1998a), test taking style (Furnham, Forde, & Cotter, 1998b), personality and performance correlations at work (Jackson & Corr, 1998), the relationship between personality and work motivation (Furnham, Forde, & Ferrari, 1999), the relationship between test taking styles and neuroticism (Jackson, Furnham, & Lawty-Jones, 1999), and the personality profile of male Anglican clergy in Britain and Ireland (Francis, Jones, Jackson, & Robbins, 2001; Francis, Robbins, Jackson, & Jones, 2000).

Costa and McCrae (1995) explored the relationship between the 21 scales of the Eysenck Personality Profiler and their preferred five-factor solution for personality assessment. Jackson, Furnham, Forde, and Cotter (2000) also explored the structure of the Eysenck Personality Profiler and concluded that there was little difference in the quality of a three factor or five-factor solution. Furnham, Jackson, Forde, and Cotter (2001) explored the relationship between the 21 scales of the Eysenck Personality Profiler and the Myers-Briggs Type Indicator (Myers & McCaulley,
In spite of its apparent usefulness, two significant problems may deter researchers from making full use of the Eysenck Personality Profiler. The first problem is clearly a matter of length. In its present form 21 scales each with 20 items, together with a 20-item lie scale, leads to a test of 440 items. A test of this length is time consuming and may alienate the goodwill of the respondents. The second problem is that little information is available on the reliability of these scales, and where such information is possible, some of the scales do not always perform satisfactorily. For example, in their study of 1148 newly ordained clergy in England, Ireland, Scotland and Wales, Francis et al. (2001) found that, although many of the scales performed very well, alpha coefficients dropped to 0.59 for manipulation, to 0.54 for expressiveness, and to 0.43 for tough mindedness. One possible solution to both of these problems is to examine the performance of the individual scales and to propose a short form by selecting the best items within the current 20-item scales. This would keep the structure of the Eysenck Personality Profiler intact, whilst enhancing its usefulness.

There is a well established history for the development of short forms of Eysenck’s personality measures. On the basis of the Maudsley Personality Inventory, Eysenck (1958) developed 6-item scales. On the basis of the Eysenck Personality Inventory, Floderus (1974) developed 9-item scales and Eysenck and Eysenck (1964a) developed 6-item scales. On the basis of the Revised Eysenck Personality Questionnaire, Eysenck et al. (1985) developed 12-item scales and Francis, Brown, and Philipchalk (1992) developed 6-item scales. On the basis of the Junior Eysenck Personality Questionnaire, Francis and Pearson (1988) developed 6-item scales. On the basis of the Revised Junior Eysenck Personality Questionnaire, Corulla (1990) developed 12-item scales, and Francis (1996) developed 6-item scales.

An earlier attempt to produce a short form of the Eysenck Personality Profiler by Eysenck, Wilson, and Jackson (1996) has employed a somewhat different rationale. Instead of reducing the number of items within each of the 21 scales, Eysenck et al. (1996) reduced the number of primary scales within each of three super-factors of personality. In this instrument extraversion comprises the three traits of sociability, activity and assertiveness; neuroticism comprises the three traits of anxiety, inferiority and unhappiness; and psychoticism comprises the three traits of risk-taking, impulsiveness and irresponsibility. Each of these nine traits is assessed by 20-item scales. See Petrides, Jackson, Furnham, and Levine (2003) for a recent review of its structure.

The wider literature has been critical of short forms for a number of good reasons (Levy, 1968; Smith & McCarthy, 1995; Smith, McCarthy, & Anderson, 2000). The point is well taken that short forms may be inappropriate in clinical assessment procedures when the classification and treatment of individuals is at stake. In respect of survey style studies, however, the case against short forms is much less substantial. What is required of short forms used in this way is clear evidence of their concurrent validity alongside the parent full-form of the scale, and good reliability (both independently and in comparison with the full-form). In this sense short forms need to function as good predictors of scores recorded on the full-form. Recent psychometric defences of short forms have been provided by Scholte and DeBruyn (2001) and by Archer, Tirrell, and Elkins (2001).

Against this background, the aim of the present study is to explore the psychometric properties of the Eysenck Personality Profiler among a sample of undergraduate students and to select the best 6 and best 12 items within each of the 22 scales to produce an instrument of 132 and 264 items,
respectively. We chose to extract 6- and 12-items per scale as our literature review has shown a history of 6-item short scales, and 12 items represents a near-midway comparison between a very short and traditional 20-item versions. Compared with a 20-item per scale measure, we envisage these versions of the EPP to be considerably shorter in terms of administration time.

2. Method

2.1. Participants

The full 440-item Eysenck Personality Profiler (Eysenck et al., 1992; Jackson et al., 2000) was administered to all the incoming students as part of the induction programme within a College of Higher Education in Wales. Almost all the students agreed to participate, generating thoroughly completed questionnaires from 290 females and 110 males. Nearly two thirds (65%) of the respondents were under the age of 20, 27% were in their twenties, and the remaining 8% were aged 30 or over.

2.2. Data analysis

In the present study, principal components analysis and item rest-of-test correlational analysis were employed to identify the best sets of 6 and 12 items to provide a short form for each of the 22 primary scales of the EPP. By these methods, items which recorded low loadings on the first factor of principal component analysis or which had low correlations with the sum total of the other items in the scale were progressively dropped until either a 6- or 12-item per scale solution was achieved.

3. Results

Table 1 assesses the internal reliability of the 20-item scales, the 12-item scales and the 6-item scales in terms of the alpha coefficient (Cronbach, 1951), and also assesses the concurrent validity of the two short forms in terms of their correlation with the full-form and their correlation with each other.

Taking Kline’s (1993) suggestion that the alpha coefficient should reach 0.70 for a normal length scale, three of the 20-item scales failed to reach that baseline, namely expressiveness and dogmatism for the dimension of extraversion, and manipulation for the dimension of psychoticism. It is unlikely that reliable short forms can be generated from unreliable full-forms. The 12-item version of these three scales (expressiveness, dogmatism, and manipulation) also failed to reach the baseline of 0.70. Additionally the scale of hypochondriasis also dropped from 0.73 to 0.69 for the 12-item short form. The concurrent validity of these twenty-one 12-item short form scales and the 12-item short form lie scale is supported by high correlations with the parent 20-item full-form scales in which all correlations are 0.89 or higher.

In contrast, 11 of the 6-items per scale alphas failed to meet Kline’s (1993) threshold criterion of 0.7. These are expressiveness, dogmatism, aggressiveness, hypochondriasis, obsessiveness, risk-
taking, irresponsibility, manipulation, tough mindedness, practical and the lie scale. Aside from the EPP scales which were already showing weak internal reliability in its parent form, the concurrent validity of the 6-item per scale version of the EPP is quite good (all 6-item scales with correlations of 0.73 or more with the parent scale and 0.71 or more with the 12-item scale).

It is also interesting to determine the success of reduced forms of the EPP-S scales. The EPP-S consists of activity, sociability, assertiveness, low self-esteem, unhappiness, anxiety, risk-taking, impulsiveness and irresponsibility. Aside of an alpha for irresponsibility of 0.69, all of our measures of internal consistency and concurrent validity for 6-item and 12-item versions are more than reasonable.

4. Discussion

This study set out to develop and to test two short forms of the 440-item Eysenck Personality Profiler among a sample of 400 undergraduate students. In its long form, our data demonstrate
that 18 of the 21 original 20-item scales met Kline’s (1993) stringent criterion for internal reliability coefficients of at least 0.70. Moreover, wherever the full scale achieved a satisfactory alpha coefficient it was possible to create a 12-item short form which also possessed satisfactory internal reliability and satisfactory concurrent validity in terms of a high correlation with the parent scale. Although there are weaknesses with some of the 12-item scales, these weaknesses are no more serious than weaknesses found in the parent scales. In other words, the 12-item short form provides a satisfactory approximation of the longer form and can be recommended for further use when time constraints render the longer form impracticable. Our findings therefore suggest that the 12-item per scale version of the EPP or EPP-S is very similar to the 20-item per scale version in terms of its basic characteristics.

We believe that there is a reasonable level of item redundancy in the full form of the EPP and that we have presented a 12-item per scale version which has considerable merit. At the same time, it should be noted that the full version of the EPP does generally have higher reliability, does cover a wider domain of behaviours consistent with each of the primary scales and therefore continues to be useful.

The 6-item per primary scale version of the EPP seems weaker than the 12-item version as a result of some poor alpha reliabilities that suggest a lack of internal consistency in the majority of the scales. Having said this however, it is fair to point out that the 6-item per scale version is supported by means of the generally high correlations between this version, the 12-item version and the 20-item parent version. To be fair to the 6-item per scale version, it might be argued that a threshold of 0.7 for judgment that an alpha is reasonable is somewhat harsh. De Vellis (1991) for example argues for a general threshold of 0.65 and an acceptability of 0.6. Clearly, the perspective of De Vellis is helpful in promoting the acceptability of the 6-item per scale version of the EPP.

We were also interested to see how well the primary scales of the EPP-Short could be reduced from 20 items to 12 and six items. Our analysis suggests that the scales which Eysenck et al. (1996) considered important enough to retain in a 10-scale version of the EPP were also of sufficient integrity that they could be reduced down to either a 12- or 6-item version.

In short, this study shows that a 12-item per primary scale version of the EPP retains many of the best features of the 20-item per scale version, but with the advantage of being much shorter. We believe that this abbreviated form of the EPP will be especially valuable when administration time is limited, and a relatively complete description of personality is required. The usefulness of a 6-item per scale version of the EPP is less since a majority of the scales do not reach an alpha of 0.7. Nevertheless, a threshold of 0.7 might be considered harsh and some researchers are happy to set a lower threshold of acceptability. The EPP consists of primary scales with good to excellent psychometric properties and it seems to be just the ones with excellent properties to begin with that can be most satisfactorily reduced down to 12 and perhaps even six items.

References


