

TEST-RETEST RELIABILITIES OF A NEW PERSONALITY QUESTIONNAIRE FOR CHILDREN

By SYBIL B. G. EYSENCK AND H. J. EYSENCK*

(*Institute of Psychiatry, University of London*)

SUMMARY. Different groups of children, varying in age, were tested and retested after varying intervals (one month, six months, twelve months) with a new Personality Questionnaire (PQ) containing scales for the measurement of the new factor of psychoticism (P), as well as of E, N and L. It was found that the new scales had retest reliabilities very similar to those established for the JEPI, and if anything a little higher. It was concluded that with suitable precautions the new questionnaire would be no less useful than the old one, as far as reliability is concerned.

INTRODUCTION

NOT very much is known about the test-retest reliability of personality inventories for children, particularly at the primary school level; this study was designed to throw some light on this question. The questionnaire used (called simply the Personality Questionnaire, or PQ) was designed to incorporate a new dimension of personality, called provisionally 'psychoticism' or P-factor, into the framework of extraversion and neuroticism; the PQ also contains a Lie scale (L scale). The original conception of the P factor, and the construction of the first sets of questions for its measurement, have been described in detail (H. J. Eysenck and S. B. G. Eysenck, 1968); later studies by S. B. G. Eysenck and H. J. Eysenck (1968, 1969) have given details about subsequent developments of the adult version. The original version of the children's PQ has been described by S. B. G. Eysenck and H. J. Eysenck (1969), and subsequent developments by Eysenck, Easting and Eysenck (1971) and H. J. Eysenck and S. B. G. Eysenck (1971). A detailed discussion of the genetic model on which the conception of the P factor is based has been given by Eysenck (1973), together with evidence regarding its heritability. In view of the close genetic connection between psychosis and psychopathy, it was thought that high P scores would be characteristic not only of psychotics (Verma and Eysenck, 1973) but also of criminals, and this has been shown to be so by S. B. G. Eysenck and H. J. Eysenck (1970, 1971) and H. J. Eysenck and S. B. G. Eysenck (1971). P correlates closely with certain types of sexual behaviour (Eysenck, 1972) and attitudes (Eysenck, 1970, 1971). In addition, there are several unpublished studies linking P with a variety of experimental procedures, from vigilance to evoked potentials on the EEG.

It is important to realise that in spite of its clinical connotation the P factor (like N) claims to apply to the *normal* child (or adult), with clinically diagnosed psychotics or neurotics only the extreme end of the distribution. P, like N and E, is based essentially on factor analysis of item intercorrelations; the constitution of the factor makes it essentially a test of the original hypothesis expressed in the choice of items for inclusion. Several successive sets of items were in fact factor analysed, as outlined in the publications mentioned above; the final

* The new questionnaire described in this publication has not yet been published, but research workers intending to use the new scale may purchase any desired quantities from the printing firm whose address will be furnished on request. We are indebted to the following for help in securing the data on which this paper is based: G. Easting, M. Morton and J. Allsopp.

version consists of an 18-item P scale, a 24-item E scale, a 20-item N scale, and a 20-item L scale. The items are almost identical with those published in previous versions, changes being slight and in line with the general character of the factor as outlined previously. High P scorers are seen to be solitary ; not caring for other people ; troublesome, not fitting in ; cruel, inhumane ; lacking in feeling, insensitive ; hostile to others, aggressive ; liking odd, unusual things ; dis-regarding danger, foolhardy ; liking to make fools of others, upsetting them. What emerges from these admittedly subjective interpretations of questionnaire responses is a fairly congruent picture of an odd, isolated, troublesome child ; glacial and lacking in human feelings for his fellow-beings and for animals ; aggressive and hostile, even to near-and-dear ones ; trying to make up for lack of feeling by indulging in sensation-seeking ' arousal jags ' without thinking of the dangers involved. As one might have expected from these descriptions, and from the adult connection between P and criminality, it has been found in children that high P scorers are characterised by bad behaviour in school, and by antisocial conduct outside it (Allsopp, 1972).

RESULTS

In order to ascertain the test-retest reliability of the scales of this test, several different sets of children were tested twice, with either one month, six months, or 12 months, intervening between occasions. Table 1 shows the values obtained by Allsopp (1972) with groups of white, and smaller groups of coloured girls. The white and coloured girls are kept separate because most of the latter had some difficulties with the English language, this not being their native tongue. Table 2 shows similar values for boys, tested at another school, by another investigator. The values in Table 1 are typically higher than those in Table 2, centring around .8 or even a little higher ; those in Table 2 centre around .75 or thereabouts. This slight difference may be due to the fact that for the girls the teacher involved was also the investigator ; for the boys the investigator was not their teacher. If the difference requires any explanation, small as it is, this is perhaps a more likely one than any reference to sex differences. There is no consistent trend for values to increase with increasing age, except perhaps for L, and possibly for N. On the whole, these values are if anything higher than those obtained with the JEPI (S. B. G. Eysenck, 1965).

TABLE 1
TEST-RETEST RELIABILITIES : ONE MONTH'S INTERVAL.

Age	n	P	E	N	L
White girls					
11	31	.83	.84	.65	.78
12	42	.75	.82	.86	.82
13	43	.76	.79	.82	.92
14	45	.75	.88	.86	.89
Coloured girls					
11	17	.70	.87	.65	.95
12	20	.74	.73	.79	.92
13	38	.89	.75	.68	.92
14	8	.78	.55	.85	.88

TABLE 2

TEST-RETEST RELIABILITIES : ONE MONTH'S INTERVAL.
Boys.

Age	n	P	E	N	L
12	58	.68	.83	.71	.59
13	84	.68	.75	.74	.79
14	48	.67	.77	.81	.79

Table 3 shows values for retests done after six months, for boys and girls separately. The number of 15-year-old girls is small, and the values in that row may be disregarded. Values are rather variable along both rows and columns, which makes it somewhat difficult to come to any general conclusion. Retest reliabilities are not lower for P than for the older established scales, and seem to centre around .6 or so if we disregard the youngest group; values for the 7-year-olds would seem to be rather lower, suggesting that perhaps many of them did not understand the questions too well. The 8- and 9-year-olds are a little below the older children, but not sufficiently so to make it inadvisable to test them; for purposes of group comparison, and for experimental purposes, these values are not too low. Boys seem to be more reliably measured than girls as far as P is concerned; this is in line with the higher scores and variances typically obtained by boys on this factor. E is more reliably measured in girls; for N and L the differences are very small. On the whole, it is doubtful if we can claim on the basis of these figures that there are any marked sex differences; different teachers might account for any differences that appear.

TABLE 3

TEST-RETEST RELIABILITIES : SIX MONTHS' INTERVAL.

Age	n	P	E	N	L
Boys					
7	84	.38	.33	.34	.62
8	104	.51	.53	.34	.67
9	87	.43	.53	.74	.75
10	117	.65	.57	.71	.67
11	195	.59	.60	.74	.72
12	198	.71	.60	.70	.59
13	201	.64	.67	.72	.65
14	191	.76	.74	.66	.71
15	47	.81	.54	.77	.84
Girls					
7	88	.53	.44	.51	.65
8	81	.44	.54	.40	.61
9	104	.69	.58	.57	.77
10	102	.48	.73	.74	.74
11	184	.47	.60	.55	.73
12	198	.53	.74	.71	.72
13	176	.65	.76	.74	.66
14	176	.63	.78	.69	.75
15	28	.36	.82	.77	.83

Table 4 shows retest reliabilities for primary school children, the interval being 12 months ; figures for the 10-year-olds are based on very small numbers and may be disregarded. The 7-year-olds tend again to have the lowest values, and are probably handicapped in reading the questions. The other figures are rather variable, but centre around $\cdot 6$, though some values fall well below this figure. Values for P are not any lower than for the better established scales. Considering that we are dealing with primary school children, the size of the reliabilities is not too disappointing, but it should be borne in mind when evaluating such findings as those reported by Eysenck and Cookson (1969, 1970) that the reported correlations between personality and achievement, say, require to be corrected for attenuation before the 'true' correlations can be properly estimated.

TABLE 4
TEST-RETEST RELIABILITIES : 12 MONTHS' INTERVAL.

Age	n	P	E	N	L
Boys					
7	36	$\cdot 53$	$\cdot 37$	$\cdot 48$	$\cdot 34$
8	50	$\cdot 43$	$\cdot 40$	$\cdot 51$	$\cdot 54$
9	60	$\cdot 44$	$\cdot 35$	$\cdot 58$	$\cdot 72$
10	14	$\cdot 32$	$\cdot 59$	$\cdot 64$	$\cdot 69$
Girls					
7	46	$\cdot 41$	$\cdot 50$	$\cdot 41$	$\cdot 53$
8	64	$\cdot 50$	$\cdot 47$	$\cdot 58$	$\cdot 66$
9	64	$\cdot 37$	$\cdot 61$	$\cdot 59$	$\cdot 66$
10	8	$\cdot 90$	$\cdot 58$	$\cdot 64$	$\cdot 51$

The data suggest certain conclusions. (1) Retest reliabilities after the lapse of one month are high enough for secondary school children to use the test with some confidence as measures of present status. (2) Retest reliabilities for the new P scale are not very different from reliabilities for other scales, whatever the test-retest interval. (3) For longer test-retest intervals, reliabilities fall to around $\cdot 6$; such low reliabilities suggest the need for correction for attenuation of correlations with personality variables, and the need for caution in interpreting cross-sectional data ; clearly longitudinal data are preferable. (4) Reliabilities for 7-year-olds are probably too low to place any confidence on scores obtained for children of that age ; 8- and 9-year-olds are in a better position, and although probably inferior to older children with respect to reliability of measurement it does not seem necessary to warn users against working with children of that age, provided numbers are large enough to cancel out chance variations. (5) The reliabilities of the new scales are very similar, and if anything a little higher, than corresponding reliabilities of the JEPI ; the fact that the latter has been used with some success by many investigators suggests that the new scales are sufficiently reliable to be used in the same manner. (6) Clearly much developmental change is taking place in the personality of these young children ; as in the case of the IQ, predictions of terminal status from scores in early childhood would be inappropriate.

REFERENCES

- ALLSOPP, J. (1972). *The Effect of Psychoticism, Extraversion and Neuroticism on Measures of Antisocial Behaviour in Secondary School Girls*. Univ. of Birmingham, unpublished M.Sc. Thesis.

- EYSENCK, H. J. (1970). Personality and attitudes to sex: a factorial study. *Personality*, 1, 355-376.
- EYSENCK, H. J. (1971). Personality and sexual adjustment. *Br. J. Psychiat.*, 118, 593-608.
- EYSENCK, H. J. (1972). Personality and sexual behaviour. *J. Psychosomatic Res.*, 16, 141-152.
- EYSENCK, H. J. (1972). An experimental and genetic model of schizophrenia. In: A. R. KAPLAN (Ed.), *Genetic Factors in Schizophrenia*. Springfield: Thomas.
- EYSENCK, H. J., and COOKSON, D. (1969). Personality in primary school children. 1—Ability and achievement, 2—teachers' ratings. *Br. J. educ. Psychol.*, 39, 109-122 and 125-130.
- EYSENCK, H. J., and COOKSON, D. (1970). Personality in primary school children. 3.—Family background. *Br. J. educ. Psychol.*, 40, 117-131.
- EYSENCK, H. J., EASTING, G., and EYSENCK, S. B. G. (1971). Personality measurement in children: A dimensional approach. *J. special Educ.*, 4, 261-268.
- EYSENCK, H. J., and EYSENCK, S. B. G. (1968). A factorial study of psychoticism as a dimension of personality. *Mult. Beh. Res.*, Special Issue, 15-31.
- EYSENCK, H. J., and EYSENCK, S. B. G. (1971). A comparative study of criminals and matched controls on three dimensions of personality. *Br. J. soc. clin. Psychol.*, 10, 362-366.
- EYSENCK, H. J., and EYSENCK, S. B. G. (1971). The orthogonality of psychoticism and neuroticism. *Percept. and mot. Skills*, 33, 461-462.
- EYSENCK, S. B. G. (1965). *Manual of the Junior Eysenck Personality Inventory*. London: Univ. London Press.
- EYSENCK, S. B. G. and EYSENCK, H. J., (1968). The measurement of psychoticism: a study of factor stability and reliability. *Br. J. soc. clin. Psychol.*, 7, 286-294.
- EYSENCK, S. B. G., and EYSENCK, H. J. (1969). Scores on three personality variables as a function of age, sex and social class. *Br. J. soc. clin. Psychol.*, 8, 69-76.
- EYSENCK, S. B. G., and EYSENCK, H. J. (1969). "Psychoticism" in children: a new personality variable. *Res. in Educ.*, 1, 21-37.
- EYSENCK, S. B. G., and EYSENCK, H. J. (1970). Crime and personality: an empirical study of the three-factor theory. *Br. J. Criminol.*, 10, 225-239.
- EYSENCK, S. B. G., and EYSENCK, H. J. (1971). Crime and personality: item analysis of questionnaire responses. *Br. J. Criminol.*, 11, 49-62.
- VERMA, R., and EYSENCK, H. J. (1973). Severity and type of psychotic illness as a function of personality. *Br. J. Psychiat.* (in press).

(Manuscript received 7th November, 1972)