# RESEARCH NOTES

## PERSONALITY AND THE MEASUREMENT OF INTELLIGENCE

H. J. EYSENCK AND P. O. WHITE*

(*Institute of Psychiatry, University of London.*)

SUMMARY. A re-analysis is presented of some data purporting to show that stable children differ from labile ones with respect to the structure of their intellectual abilities. The hypothesis is supported, and additional data are presented tending to show that theories of linear independence between cognitive and non-cognitive areas may have to be supplemented by theories stressing non-linear dependence.

It is usually maintained that intelligence is statistically independent of temperamental factors such as neuroticism and extraversion and the evidence does, indeed, show little cause to doubt lack of correlation between the cognitive and the conative-affective sides of personality. (Cf. Cattell, 1963, for a recent study and discussion). However, it would be unwise to equate *statistical independence* with *lack of interaction* ; most studies reported in the literature have used statistical methods based on product-moment correlations, thus setting orthogonality equal to linear independence, and failing to allow for the possibility of curvilinear regression. This failure to take into account more complex modes of causation may have arisen from the fact that psychometric procedures have been developed very much in isolation, and without connection with the large body of experimental psychology. Eysenck (1957) has argued that the study of temperament and of intelligence can be enriched tremendously by regarding the performance of personality and intelligence tests from the point of view of experimental psychology, considering it as subject to the well known laws of learning theory, and making predictions from these. The usefulness of this approach to the study of personality variables, such as neuroticism and extraversion, has been demonstrated in several publications (Eysenck, 1960, 1964). In this paper, we shall be concerned with a consideration of a similar approach to intelligence test problem solution.

It has been argued (Eysenck, 1957) that the performance of a typical intelligence test may be regarded as an instance of massed practice, in which very similar tasks are attempted repeatedly without the interposition of a programmed rest pause Under these conditions, we would expect reactive inhibition to build up and interfere with the proper execution of the tasks. We would also expect that extraverted subjects, liable as they are to greater accumulation of inhibition, would show work curves different from those produced by introverted subjects, an expectation shown to be verified by two experimental studies at a high level of significance (Eysenck, 1957, pp. 132-133). In another study, Eysenck (1959) predicted that " in the process of solving the sixty problems of the Morrisby Compound Series Test . . . extraverts would show greater reactive inhibition, and consequently falling off in performance during the last quarter of the test as compared with the first three-quarters." The results showed " that extraverts show greater work decrement . . . by taking longer to obtain correct solutions toward the end of the test, as compared with introverts, and by giving up more easily toward the end " (p. 592). (At the beginning of the work, extraverts were significantly quicker than introverts.)

As regards neuroticism, it has become customary to regard this variable as in some ways being synonymous with drive ; this supposition, taken together with the Yerkes-Dodson Law, may be taken to imply the likelihood of a curvilinear relationship between intelligence and neuroticism, extremely high and extremely low values of N being equally incompatible with high scores on intelligence tests. Lynn and Gordon (1961) have reviewed some of the literature on this point, and they have also

* Supported by the Medical Research Council.

reported an experiment of their own which strikingly (and significantly) supported this prediction, although on only a rather small number of subjects. Their findings on extraversion were indeterminate, probably because they purposely used a very short version of the Matrices test, thus making it impossible for any large amount of inhibition to accumulate. Furneaux (1962) has also shown in connection with the prediction of success of university students that simple linear correlations are much less informative than hypothesis-directed investigations into personality-intelligence relations of a more complex character. We may conclude from this brief review that there is ample evidence to suggest that temperamental and cognitive aspects of personality may not be as unrelated as has often been supposed, and that specific hypotheses about their interrelations can be formulated on the basis of modern learning theory and its extension to personality.

One such extension of the traditional approach may be made in the field of factor analytic determination of personality structure. The problem which arises may perhaps be put as follows : When a factor analysis is carried out of personality inventory scales, a number of factors, such as extraversion, neuroticism, etc., usually results (Eysenck, 1960) ; similarly, when a factor analysis is carried out of intelligence test scales, a number of factors such as verbal ability, perceptual ability etc., usually result (Vernon, 1958). These factors are independent, in the first case of intelligence, in the second case of neuroticism or extraversion, as long as we preserve the rule that we are only concerned with linear relations. But we may enquire whether similar factors and relations would emerge if we extracted personality factors from populations differing in intelligence level, or intelligence factors from populations differing in degree of neuroticism, say.

A recent study by Shure and Rogers (1963) has attempted to answer the first question. They administered the eighteen scales of the California Psychological Inventory (CPI) to three student groups differing without overlap in I.Q. level, and then intercorrelated and factor analysed the resulting scores for the three groups separately. They found that while there was considerable overall similarity in the solution, the total factor variance associated with their neuroticism factor dropped by over 30 per cent. in going from the high ability group to the low ability group. (The sum of squared loadings is, respectively, 5·18, 4·64 and 3·48 for the three groups.) No such change was observed in their extraversion factor, the sum of squared loadings being 3·46, 3·76 and 3·17, respectively, for the three groups. While confirmation would, of course, be essential before too much credence can be given to this finding, it would appear that factorial studies of personality may not give invariant results under change of ability level.

The other problem raised is perhaps even more important from the educational point of view ; would factorial studies of abilities be invariant under change of personality composition of the groups under analysis ? It is with this question that this paper is particularly concerned.

The only paper concerned specifically with this problem is one recently published by Lienert (1963). His work is based on 1,003 school children with a mean age of between 15 and 16 ; three-fifths of the children were male. These children were administered thirteen intelligence tests of the Thurstone (1938) type, constituting the so-called Leistungsprüfsystem of Horn (1962a). Also administered was a personality questionnaire modelled after Eysenck's (1953) M.M.Q. by Horn (1962b) which gives a measure of neuroticism and also contains a lie scale. Seventy-seven subjects were excluded from the analysis because they had not completed all the tests or because of unusual lie scale scores. Of the remaining subjects, 259 labile and 262 stable children were selected as constituting the 25 per cent. highest scoring and lowest scoring subjects, respectively, on the neuroticism scale. There were no differences between the groups in age but there were more girls in the labile group. However, Lienert was able to show in a preliminary factor analysis that sex had no effect on the factorial structure of the tests. A product moment correlation of the summed standard scores on the thirteen tests with neuroticism gave a value of –0·16 ; while statistically significant because of the large numbers this is for practical purposes equivalent to a finding of orthogonality between the two variables

Separate matrices of intercorrelations were calculated for the labile and stable subjects, respectively, and split-half reliabilities were calculated for all the tests for the two groups. Reliabilities did not differ, but the average intercorrelation of the tests was slightly and significantly higher for the stable group (·33 as opposed to ·27).

TABLE 1

| | Stable | | | | | Labile | | | |
|---|---|---|---|---|---|---|---|---|---|
| Test: | I | II | III | $h^2$ | Lienert | I | II | $h^2$ | Lienert |
| 1. Discovery of rules (reasoning) ....... | 62 | ·05 | 18 | ·63 | ·67 | ·25 | ·58 | ·47 | ·54 |
| 2. Problems (reasoning) .. | ·63 | ·07 | ·13 | ·63 | ·84 | ·46 | ·40 | ·49 | ·60 |
| 3. Word knowledge (verbal comprehension) .. | ·03 | ·81 | ·08 | ·37 | ·66 | ·88 | –·09 | ·45 | ·77 |
| 4. Word completion (verbal compr. and closure) | ·49 | ·42 | –·03 | ·57 | ·69 | ·27 | ·13 | ·22 | ·14 |
| 5. Word fluency (verbal compr. and fluency) | ·07 | ·84 | ·04 | ·40 | ·90 | ·82 | –·02 | ·45 | ·78 |
| 6. Rotation (spatial orientation) ....... | ·18 | ·06 | ·43 | ·39 | ·45 | –·01 | ·45 | ·25 | ·24 |
| 7. Brick-counting (spatial orientation) ........ | ·07 | ·00 | ·76 | ·43 | ·63 | ·07 | ·71 | ·44 | ·53 |
| 8. Plane counting (spatial orientation) ........ | ·19 | ·01 | ·65 | ·48 | ·65 | –0·2 | ·53 | ·29 | ·32 |
| 9. Hidden figures (spatial orientation and clos.) | ·56 | ·06 | ·21 | ·60 | ·69 | ·32 | ·38 | ·39 | ·36 |
| 10. Hidden pictures (closure) ........... | –·03 | ·27 | ·40 | ·27 | ·62 | ·04 | ·47 | ·29 | ·30 |
| 11. Words (word fluency).. | ·24 | ·38 | ·02 | ·36 | ·56 | ·56 | ·05 | ·34 | ·42 |
| 12. Word beginnings (word fluency and verbal comprehension) .... | ·51 | ·25 | ·05 | ·56 | 56 | ·62 | ·15 | ·44 | ·48 |
| 13. Counting (number) ... | ·22 | ·18 | ·17 | ·34 | ·40 | ·31 | ·02 | ·19 | ·21 |
| | 1·00 | ·33 | ·42 | | | 1·00 | ·32 | | |
| | ·33 | 1·00 | ·17 | | | ·32 | 1·00 | | |
| | ·42 | ·17 | 1·00 | | | | | | |

Factor loadings of stable and labile groups compared on Promax Solution. Also given are original Lienert communality estimates, and Promax intercorrelations between factors. A brief description of each test is quoted from Lienert, and also the test's suggested factor composition.

Next, Lienert carried out a multiple factor analysis following Thurstone's (1947) procedure. It was found that eight factors could be extracted from the stable group and only four from the labile group. Communalities were lower for the labile than for the stable group and specific factors were more important for the labile than for the stable group. After rotation, it was found that three factors could be interpreted for the labile and six for the stable group; the latter were said to be closer to Thurstone's primary factors, whereas the former were much more mixed. These figures suggest strongly that children high and low on neuroticism differ very significantly in the way their mental abilities are structured. This conclusion is so important that a thorough critical analysis of the study seems in order.

The first point of criticism is that too little information is given about the analysis to make detailed evaluation possible. The only reference is to Thurstone's book (1947) which contains a number of different methods of analysis, and it is not possible, for instance, to find out just what criteria were used for the extraction of factors or for the interpretability of factors.

Even more disturbing is the failure of the discussion to agree with the results given. Thus, for instance, Lienert says (page 149) that " factor A is a purely verbal factor because it has substantial loadings only in verbal tests." Inspection of Table 5 (*b*) shows that factor A has the highest loading on a reasoning test, the second highest loading on a word fluency test, the third highest loading on a number test, the fourth highest loading on a space orientation test, and the fifth highest loading on a space orientation test. The sixth highest loading is on a reasoning test. Thus, of the six tests with the highest loading on factor A only one could be interpreted as representing a verbal factor. Factor B is said to be a reasoning factor, having its highest loadings on two tests which, in actual fact, have nearly the lowest loadings on this factor. Altogether, we were unable to make the figures agree with the interpretations, and this must cast doubt on the analysis as a whole, and the conclusions derived by Lienert.

Fortunately, the original matrices were given in the paper, and thus it was possible to carry out a re-calculation based on more modern analytic methods of factor rotation. The method of analysis used by us was Hotelling's principal axes method.

Guttman's (1954) well-known lower bound for the number of common factors indicated the number of factors to be retained. The number is equal to the number of latent roots greater than one in the correlation matrix with unit diagonals. This corresponds identically to Kaiser's (1962) upper-bound for the number of factors with positive generalizability (a term introduced by Cronbach, *et al.* (1963) for the old notion of internal consistency reliability). Three factors were indicated for the stable group and two factors for the labile group. This is in marked contrast to Lienert's solution in which eight factors are retained for the stable group and four for the labile group. The reason for this discrepancy is difficult to assess since Lienert does not indicate his criteria for this decision. Probably, it is largely due to the inefficiency of the centroid method relative to the principal axes method. But, since we have no indication as to the reflection procedures used in the centroid analyses which Lienert presents, the relative efficiency cannot adequately be assessed.

With the number of factors thus fixed, the communalities were estimated by the now standard procedure of iteration by refactoring (Harman, 1960). The method of principal axes was used and after fifteen cycles, all communalities and converged to three decimal places (though most had converged to four or five places). The final communality estimates are presented in Table 1. For comparison, we also present Lienert's estimates in the same table.

The marked tendency towards very much lower communalities for the labile group which Lienert notes is not so apparent in the present analysis. Since Lienert does not indicate his basis for estimation, the reason for the discrepancy cannot be evaluated.

With the communality estimates thus determined, and the number of factors fixed as before the factor loadings were computed for each matrix by the method of principal axes. Kaiser's (1956, 1958) Varimax procedure for analytical rotation to orthogonal simple structure was applied to the principal axes matrices. The Promax (Hendrickson and White, 1964) procedure for analytic rotation to oblique simple structure was applied to the Varimax solutions. The oblique factor loadings for each matrix appear in Table 1, along with the intercorrelations among the primary factors and the test communalities. The principal axes loadings and the intermediate Varimax loadings are not presented here but all relevant matrices are available at the Institute of Psychiatry.

Upon inspecting the patterns of loadings presented in this table, one is not particularly impressed by the clear and unambiguous interpretability of the resultant factors. Indeed, the crisp, clear simple structure usually associated with P.M.A. material is nowhere to be seen. However, oblique rotation has cleaned up the simple structure considerably and tentative hypotheses may be put forth for at least some of the factors. For convenience of reference, $S_1$, $S_2$, and $S_3$ will indicate the respective factors for the stable group, while $L_1$, $L_2$ will indicate those for the labile group.

Factors $S_2$ and $L_1$, seem primarily to involve the use of words. In each case, the four tests with highest loadings were postulated as measures of either the Verbal Comprehension factor (V) or the Word Fluency factor (W). Factor $S_3$ has its three highest loadings on tests postulated as measures of the Spatial Relations (S) factor and no other loadings exceed 0·40. Factors $S_1$, and $L_2$ appear to be rather complex. No very simple interpretation is suggested although the pattern of loadings appear to be rather similar for the two factors. Each is loaded by tests hypothesized as measures of Reasoning (R), Spatial Relations (S) and Closure (C). Additionally, $S_1$ has moderate loadings on putative measures of Verbal Comprehension (V) and Word Fluency.

Our own solution, while differing considerably from Lienert's, does suggest that his main contention is indeed borne out by his data ; the stable group has a more clearly marked structure in the cognitive test field than has the labile group. Three significant factors in the stable group are opposed to two significant factors in the labile group, and as the same standards of selection were employed at all stages, there seems little reason to doubt that these differences are real ones rather than being statistical artefacts. It will, of course, be necessary for this work to be repeated, preferably with a larger selection of tests, before the revolutionary implications of Lienert's work can be accepted ; nevertheless, it would seem likely that personality and intelligence test performance are indeed more closely imbricated than has hitherto been thought likely.

It will have been noted that there is a curious symmetry in the results obtained by Lienert, and those obtained by Shure and Rogers. High ability subjects show higher variance of the N factor than do low ability subjects. High stability subjects show greater organization of abilities than do labile subjects. It would almost appear as if greater stability and ability, respectively, went with greater degrees of organization of ability and stability. It is much too early to speculate about the possible meaning and causes of these relations ; much further research is required before the facts themselves are adequately established to call for explanatory hypotheses. Nevertheless, the theory of linear independence between cognitive and non-cognitive factors may soon have to be supplemented by one stressing non-linear dependence and interrelation.

## References.


Cattell, R. B. (1963). The theory of fluid and crystallized intelligence : a crucial experiment. *J. Educ. Psychol.*, 54, 1-22.

Cronbach, L. J., Rajaratnam, Nageswari, and Gleser, Goldine C. (1963). Theory of generalizability : a liberalization of reliability theory. *Brit. J. Stat. Psychol.*, 16, 137-163.

Eysenck, H. J. (1953). *The Scientific Study of Personality.* London : Routledge and Kegan Paul.

Eysenck, H. J. (1957). *The Dynamics of Anxiety and Hysteria.* London : Routledge and Kegan Paul.

Eysenck, H. J. (1959). Personality and problem solving. *Psychol. Rep.*, 5, 592.

Eysenck, H. J. (1960). *The Structure of Human Personality.* London : Methuen.

Eysenck, H. J. (Ed.) (1964). *Experiments in Behaviour Therapy.* London : Pergamon Press.

Furneaux, W. D. (1962). The psychologist and the university. *Univ. Quart.*, 33-47.

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

Harman, H. (1960). *Modern Factor Analysis.* Chicago : University of Chicago Press.

Hendrickson, A. E., and White, P. O. (1964). Promax : a quick method for rotation to oblique simple structure. *Brit. J. Stat. Psychol.*, 17.

Horn, W. (1962a). *L-P-S Leistungsprüfsystem.* Göttingen : Hogrefe.

Horn, W. (1962b). *S-O-F Stabilitats—und Offenheitsfragebogen.* Göttingen : Hogrefe.

Kaiser, H. F. (1956). *The Varimax method of Factor Analysis.* Unpublished Ph.D. dissertation. University of California.

Kaiser, H. F. (1958). The Varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.

Kaiser, H. F. (1962). Image analysis. In : *Problems in Measuring Change.* Ed. C. W. Harris. University of Wisconsin Press, Madison.

Lienert, G. A. (1963). Die Faktorenstruktur der Intelligenz als Funktion des Neurotizismus. *Ztschr. f. Exp. Angev. Psychol.*, 10, 140-159.
Lynn, R., and Gordon, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.
Shure, G. H., and Rogers, M. S. (1963). Personality factor stability for three ability levels. *J. Psychol.*, 55, 445-456.
Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monogr. No.* 1. Chicago: University of Chicago Press.
Thurstone, L. L. (1947). *Multiple Factor Analysis*. Chicago: University of Chicago Press.
Vernon, P. E. (1958). *The Structure of Human Abilities*. London: Methuen.

## PUBERTY AND TEST PERFORMANCE: A FURTHER REPORT

By J. D. NISBET, R. ILLSLEY, A. E. SUTHERLAND and M. J. DOUSE
(*University of Aberdeen*)

Summary. We reported previously from successive testing of an age-group of girls at ages 7, 9, 11 and 13, that those who reached puberty before age 12 : 3 had a slight but consistent superiority in average intelligence test score. A further testing of a selected sample from this group at age 16 suggests that this superiority diminishes as the late developers reach puberty.

### I.—Introduction.

In a previous paper (Nisbet and Illsley, 1963) we reported on the relation of early onset of puberty to test performance at age 11, on the basis of a follow-up of a year-group of girls from age 7 to age 13. The results presented there agree closely with the findings of Douglas's long-term study under the auspices of the Population Investigation Committee, reported in *The Home and the School* (1964 : Chapter X). Both studies showed that girls who reached menarche relatively early obtained on average slightly higher scores at age 11; but this slight superiority was not associated with onset of puberty, being as marked at age 7 or 8 as at the age of 11.

It is clearly important to extend the age-range of our investigation to cover the later stages of adolescence also. To quote Douglas (*op. cit.*, page 82): "Although the girls who mature early have superior measured ability at 11 years, there is no certainty that they retain this advantage in later school life . . . It is important to know what is happening. If the early developers retain their superiority in performance throughout their school life, there is no reason to allow for maturity when selecting for secondary schools, but if they lose their initial advantage and the late developers catch up, we are selecting inefficiently and should do something about it."

The differences in mean score between early and late maturers are relatively small, compared with the spread of scores in both groups which extends over the full range from highest to lowest. We should not, therefore, exaggerate the importance of the differences for selection at age 11. However, the finding is of interest for the understanding of influences on mental development as measured by tests of reasoning.

The extension of inquiry to include the later stages of puberty has been done in Aberdeen, though only on a selected sample of the original group, namely, those who stayed on at school after age 15 in the senior secondary schools. In our previous study, which stopped at age 13, we were able to identify the early and very early maturers. In this extension, we were able to identify also the late maturers separately from the median group.

### II.—Subjects and tests.

The population of the original inquiry was defined as those girls who were in the second year of secondary schools in Aberdeen in December, 1960. In June, 1963, when the girls were age 16, those in the senior secondary schools were given Moray House Adult Test I. They had done Moray House Advanced Test 10 two-and-a-half years previously, and for the majority age at menarche was already