British Journal of Psychology (1986), 77, 137–142 *Printed in Great Britain*

Distributions of intelligence and educational level in fathers and sons

P. A. Vroon, J. de Leeuw and **A. C. Meester**

According to the purely hereditary position with respect to the origin of differences between individuals, the correlation between the intelligence levels of single parent and offspring should at least be 0·50 when perfectly reliable tests are used. This is supposed to hold when there is random mating. In the case of assortative mating the correlation rises theoretically to 0·71. From a psychometric and theoretical point of view, parent and child should be tested at the same age with the same test. Data of 2847 father–son pairs meeting these requirements show that the correlation is 0·33. A further analysis of the data indicates that the distribution of the sons' intelligence and attainment is only weakly determined by hereditary and environmental variables. Path analyses show that almost all variance of the sons' intelligence test scores remains 'unexplained'.

Darwin's theory of evolution inspired psychology to measure individual differences. This mental test movement started around 1900. Intelligence tests are widely used for the selection of children in schools and staff in industry and business. Right from the start a controversy arose between the so-called hereditarians and environmentalists (Vroon, 1980). According to hereditarians such as Jensen (1972) and Eysenck (1973) about 80 per cent of the variance of intelligence test scores and related variables such as educational level is caused by genetic variance. On the other hand, environmentalists claim that education and environment are mainly responsible for the observed individual differences. This IQ debate is still going on (Eysenck & Kamin, 1981).

Parents and children have 50 per cent of their genes in common. According to many adherents of the hereditarian viewpoint, intelligence is determined by about 100 specific genes, having equal, small and additive effects (Eysenck, 1973, see also McAskie & Clarke, 1976). On the basis of Mendelian genetics it is predicted (Jensen, 1972) that the correlation between single parent and offspring is 0·50 when there is random mating. Since people tend to marry each other also on the basis of intelligence level and education, there is usually some degree of assortative mating. With random mating the single-parent offspring correlation is predicted to be 0·50. This means that half the variance of the offspring is described by the variability of the parents. It follows that in this case the mid-parent offspring correlation should be 0·71 (McAskie & Clarke, 1976). Assortative mating increases both values. The single-parent offspring correlation approaches the mid-parent offspring correlation when the degree of assortative mating increases. Consequently, in case of some degree of assortative mating the single-parent offspring correlation should be somewhere between 0·50 and 0·71. It follows that with perfect assortative mating all the variance of the parents describes (or explains) those of their children. These predictions hold when intelligence differences are entirely due to genetic factors, and when perfectly reliable tests are used.

As a rule, parent–child resemblance is studied by testing parents and children once. If parents are tested as adults and children as children, often quite different tests are used. Usually this also occurs when parents and children are tested at the same age since there is a considerable time lag between the testings. A particular test takes a sample out of the undetermined assembly of intellectual performances. Consequently, it is not justified to correlate different tests straight away, Since there is no perfect theory of test performance development, it is also misleading to test parents and children with the same test at considerably different ages. When we define resemblance in terms of actual performance, a direct comparison is required between a large sample of parents and children, tested at the same age with the same test (McAskie & Clarke, 1976). Several summaries of these studies show that in many cases the samples are small and rarely representative (Erlenmeyer-Kimling & Jarvik, 1963; Scarr & Carter-Saltzmann, 1982). Moreover, most of the time parents and children have been investigated with different tests and/or at different ages.

There are three studies meeting heavier requirements. Guttman (1974) compared 89 Israeli father–son pairs using the Raven Progressive Matrices Test (a non-verbal test of logical reasoning) and found a correlation of 0·36. De Fries *et al.* (1979) compared 672 father–son pairs of European ancestry and 241 father–son pairs of Japanese ancestry on Hawaii. The Raven correlations were 0·23 and 0·09, respectively. Finally, Park *et al.* (1978) did a similar investigation in Korea with 100 father–son pairs. They found a correlation of 0·33. The findings of these authors indicate that the observed correlation between single parent and child is much lower than the value predicted by the completely genetic model.

Studies on the relationship between these correlations and education or socio-economic class present a confusing and highly controversial picture. For example, in case of high or low SES groups, the correlations between intelligence and attainment may be low because of restriction of range. Methodologically speaking, McAskie & Clarke (1976) are right when they say that parents and children should be tested at the same age with the same test, being relatively free from educational knowledge. An example is the Raven Progressive Matrices test. Moreover, a large and inhomogeneous sample is necessary. We have met these requirements by carrying out a study with the help of the Dutch Ministry of Defence.

Method

In the Netherlands military service is compulsory. The Ministry calls up all males for a medical and psychological examination in the year they reach the age of 18. Since 1945 several intelligence tests have been used in combination. The only test that remained unchanged since that time is the Raven Progressive Matrices test. In its original form it contains 60 items of increasing difficulty. In 1945, an item-test analysis was carried out and the 40 best discriminating items were chosen. This means that the data of the fathers of the sons who are tested now can be traced if the father was called up in 1945 or later.

The summonses for the examination are sent out by a national centre of the Ministry. In 1981 and 1982 a total of 20 000 questionnaires were added to the standard forms. Since about 500 persons received this order each day, all questionnaires were sent out within a total of 40 days. We asked the draftees to state their registration number which contains the year, the month and the date of birth and three other digits, and to mention their most recent educational certificates and current education. Furthermore, we asked if their father was still alive, if he was a part of the family, his profession and registration number or date of birth. In 10 000 questionnaires we also asked respondents to give the educational level of the mother.

We received about 5 000 questionnaires back which were coded and put on computer file. This response rate of 25 per cent is not unusual for this type of study. The raw Raven scores of the sons were looked up, and subsequently the raw scores of the fathers at the time they were tested. Due to incorrectly or incompletely filled-in forms, missing data, doubtful cases, wrong registration numbers leading to false father–son combinations and various other checks, 2847 father–son pairs were left for analysis.

Results

The score distribution of the fathers appears to be symmetric and normal. As far as the sons are concerned there is a relative loss of variance, and the mean is shifted almost eight points higher (from 22 to 30 items out of 40). This dramatic 'increase in intelligence' has earlier been observed in The Netherlands (Dronkers, 1978; Meester & De Leeuw, 1984) and elsewhere (Flynn, 1984), but it remains largely unexplained. The rise may be partly due to better education, but this is incompatible with the current decline of Scholastic Aptitude Test scores (Flynn, 1984).

It is important to find out if the sons in our example are a random sample from the population of all sons tested in the same period. For this purpose, our son distribution was compared with the population of all individuals tested during the period the questionnaires were sent out. A comparison of the son distribution with this population comprising 57 897 individuals shows that lower test scores are somewhat underrepresented in the sample. An explanation might be that many persons with (very) low test scores are not motivated to fill in the questionnaire. However, a statistical test revealed that the distortion is not important ($\chi^2 = 69{,}5$, d.f. $= 40$).

Subsequently, we computed the regression functions applying to the complete bivariate distribution. The regression of sons on fathers increases only slightly. The regression function of fathers on sons shows that there is some instability in the lower part of the scale. Above the mid-point of the scale the regression function is somewhat steeper. According to a chi-square test the linearity of regression had to be rejected for both curves. The non-linearity of the regressions makes the ordinary correlation coefficient a rather doubtful measure. It equals 0·30.

Another statistical analysis is possible if the test scores are grouped into a relatively small number of classes. We grouped the scores into six classes, following a procedure used by the Ministry. The first and last classes each contain approximately 10 per cent of the distribution, the four middle classes 20 per cent each. The raw bivariate distribution cannot be bivariate normal, but it can be 'grouped' or 'strained' binormal. By this we mean that there are two bivariate normal latent variables, i.e. the true test scores of fathers and sons, varying continuously. They are rounded to the observed test scores, but not necessarily by using equal intervals. In this way there are five parameters corresponding with category boundaries for sons, and five parameters for the fathers. Consequently, the grouped binormal model has 11 parameters, 10 for the boundaries and one for the correlation. Correlations estimated on the basis of a grouped binormal model are called polychoric correlations. The polychoric correlation for our data with six classes is 0·32 with a standard error of 0·02. This is a better estimate of the population correlation coefficient because it computes the same parameter on the basis of a more appropriate model.

In order to refine the analysis, we can also apply a correction for selection (restriction of range). Here we take into account that the studied sons are a slightly biased sample from the population of all sons tested in the same period. We assume that responding or non-responding is independent of the score of the father, given the score of the son. To put it differently: we assume that the conditional distribution of the score of the father, given the score of the son, is the same in the responding and in the non-responding groups. Applying our polychoric algorithm to the corrected tables reveals a correlation of 0·34 with standard error 0·02.

We also tested if the polychoric correlation differs when we calculate it over different socio-economic classes of the fathers. For this purpose, we divided the sample into six classes on the basis of the current profession of the father. It turns out that the correlations do not systematically vary over socio-economic class. The correlation between father and son is basically the same, independent of professional level. This 'partial polychoric correlation' is 0·27, with standard error 0·02. Apparently, the size of the correlation is independent of environmental factors such as professional status and education. The model saying that the association between the scores of fathers and sons is the same for all SES

Table 1. Correlations between Raven scores of the fathers (R_f), educational level of the fathers (E_f), profession of the fathers (P_f), Raven scores of the sons (R_s) and educational level of the sons (E_s)

	R_f	E_f	P_f	R_s	E_s
R_f	1·00	—	—	—	—
E_f	0·53	1·00	—	—	—
P_f	0·42	0·60	1·00	—	—
R_s	0·29	0·24	0·18	1·00	—
E_s	0·33	0·43	0·37	0·49	1·00

classes can also be tested with log-linear analysis. We find a chi square of 111·5 with 125 d.f., which is again non-significant. Both polychoric and log-linear methods strongly reject the hypothesis that scores of sons and fathers are independent, given SES.

Path analysis

The strength of associations in a correlation matrix can also be assessed by means of path analysis. We scaled the professional levels of the fathers and the educational levels of the fathers and the sons on six-point scales, using procedures of the Ministry and of the Dutch Bureau of Statistics. The data describing the educational levels of the mothers could not be used because their education was very homogeneous. Most of the mothers had only completed primary school. This results in a serious lack of variance, which would have led to correlations of about zero. This also means that we have no reason to say that there is a considerable degree of assortative mating in our sample.

Consequently, we have five variables: Raven score of the father (R_f), Raven score of the son (R_s), professional level of the father (P_f), educational level of the father (E_f) and educational level of the son (E_s). On the basis of PREHOM the matrix was calculated for optimally scaled variables. Conventional methods to calculate correlations presuppose that in case of ordinally scaled variables the distances between the categories are equal. This assumption is usually incorrect and may lead to sub-optimal estimates of the correlations. PREHOM optimizes the estimated correlation matrix by taking into account that the distances may be not equal (Table 1). Path analysis is a very complex mathematical technique, related to multiple regression analysis (Blalock, 1961). An important application of path analysis is to establish 'causal' inferences about correlations in *ex post facto* research. One of the difficulties, however, is that the word 'cause' has surplus meaning and metaphysical overtones. Furthermore, causal laws cannot or only hardly be demonstrated empirically, but it is very helpful to think causally. In using path analysis, a correlation matrix is analysed into unidirectional and sometimes bidirectional influences. A path coefficient can be read as a correlation coefficient but the amount of explained variance is not straightforwardly known after squaring these coefficients (Saris & Stronkhorst, 1984).

We calculated as many path analyses on the matrix as possible. Paths cannot be added or deleted randomly because a chi square test has to show if the analysis still fits the matrix or not. Many models could not pass this test. The statistically best fitting model is shown in Fig. 1.

We see that the relationships within the population of the fathers and within the population of the sons are much stronger than the relationships between the populations, a finding which is perhaps not very surprising. The Raven seems to be a test that is rather free of specific educational effects. The path from test to education is much stronger than

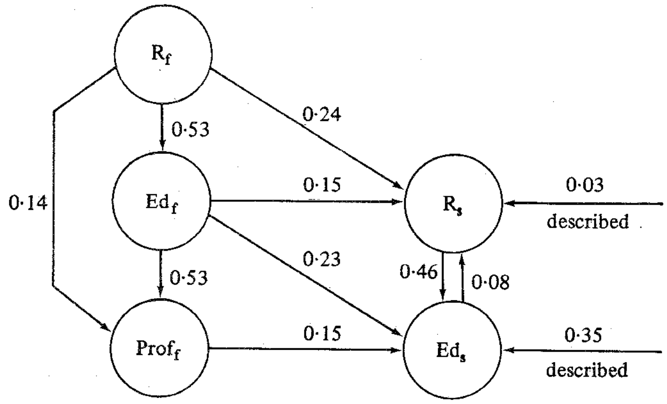

Figure 1. Best fitting path analysis on the correlations between Raven scores of the fathers (R_f), educational level of the fathers (Ed_f), profession of the fathers ($Prof_f$), Raven scores of the sons (R_s) and educational level of the sons (Ed_s). The arrows on the right hand side show the percentages of described variance.

the other way around. On the basis of Fig. 1 we *infer*, strictly speaking, that education (and attainment) is strongly influenced by the relatively stable Raven score. Conversely, there is no reason to assume that education considerably determines the score on the Raven test. In cybernetic terms: there is a small amount of feedback from attainment to Raven score. This interpretation is in line with the well-known finding that the Raven is a relatively education- and culture-free test.

Interestingly, the model as a whole describes a very small amount of variance only, namely 0·35 of the E_s variance and 0·03 of the R_s variance. This seems to mean that neither genetic nor gross educational variables play a very *important* causal role in the distribution of the sons' intelligence. We failed in correcting for the possible unreliability of the Raven: this model had to be rejected on statistical grounds.

Discussion

The present study has several advantages over earlier ones. We have the largest sample ever studied, given that parent and offspring are tested at the same age with the same test. It should be mentioned that, contrary to the predictions of McAskie & Clarke (1976), the sons show less IQ variance than their fathers. We do not know whether this is caused by a slight response bias in the sample.

Of course it should be remarked that the Raven test is just one of the arbitrary ways to measure intelligence in the psychometric tradition. This test correlates about 0·75 with, for example, the well-known and widely used Wechsler Adult Intelligence Scale. Although the matrix test requires no verbal fluency, it relates no more closely to so-called Performance IQ than the Verbal IQ; both correlations are about 0·70.

The apparent absence of assortative mating lowers the maximum predicted single-parent offspring correlation to 0·50. We find 0·34. As far as the path analyses are concerned, the amount of described variance is low. This is partly caused by the fact that we could not study variables of the mothers. Nevertheless, the data suggest that much less can be explained than might be expected on the basis of *substantial* hereditary factors. We have also not much evidence to defend the environmentalist's position, saying that the educational level and the profession of the father is greatly responsible for the IQ and the education of the child.

This family study is unable to disentangle genetic and environmental factors because the sons share family environment as well as genetic factors with their fathers. It is possible that the relatively low correlations between fathers and sons are caused by the fact that environmental variables belonging to the family become less important when children grow older. By and large, the findings indicate that the distributions of intelligence and educational level of fathers and sons cannot be sufficiently explained by current social science theorizing. This conclusion is contrary to the claims of both leading positions in the IQ debate.

References

Blalock, H. (1961). *Causal Inferences in Nonexperimental Research.* Chapel Hill: University of North Carolina Press.

De Fries, J. C., Johnson, R. C., Kuse, A. R., McClearn, G. E., Polvina, J., Van den Berg, S. G. & Wilson, J. R. (1979). Familial resemblance for specific cognitive abilities. *Behavior Genetics*, **9**, 23–43.

Dronkers, J. (1978). De stijging van intelligentiescores. *Hollands Maandblad*, **19**, (363), 15–19.

Erlenmeyer-Kimling, L. & Jarvik, L. F. (1963). Genetics and intelligence: A review. *Science*, **142**, 1477–1479.

Eysenck, H. J. (1973). *The Inequality of Man.* London: Maurice Temple Smith.

Eysenck, H. J. & Kamin, L. (1981). *Intelligence: The Battle for the Mind.* London: Pan Books.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932–1978. *Psychological Bulletin*, **95** (1), 29–51.

Guttman, R. (1974). Genetic analysis of analytical spatial ability: Raven's Progressive Matrices. *Behavior Genetics*, **4**, 273–284.

Jensen, A. R. (1972). *Genetics and Education.* London: Methuen.

McAskie, M. & Clarke, A. M. (1976). Parent–offspring resemblances in intelligence: Theories and evidence. *British Journal of Psychology*, **67**, 243–273.

Meester, A. C. & De Leeuw, J. (1984). Over het intelligentieonderzoek bij de militaire keuringen vanaf 1925 tot heden. *Mens en Maatschappij*, **59**, 5–26.

Park, J., Johnson, R. C., De Fries, J. C., McClearn, G. E., Mi, M. P., Rashad, M. N., Van den Berg, S. G. & Wilson, J. R. (1978). Parent–offspring resemblance for specific cognitive abilities in Korea. *Behavior Genetics*, **8**, 43–52.

Saris, W., & Stronkhorst, L. H. (1984). *Causal Modelling in Non-experimental Research.* Amsterdam: Sociometric Research Foundation.

Scarr, S. & Carter-Saltzmann, L. (1982). Genetics and intelligence. In R. J. Sternberg, (ed.), *Handbook of Human Intelligence.* Cambridge: Cambridge University Press.

Vroon, P. A. (1980). *Intelligence.* Amsterdam/New York: North-Holland.

Received 3 May 1985; revised version received 27 June 1985

Requests for reprints should be addressed to P. A. Vroon, Department of Psychonomics, University of Utrecht, Heidelberglaan 1, 3584, C.S. Utrecht, The Netherlands.

J. de Leeuw and A. C. Meester are at the Department of Data Theory, University of Leiden, Hooigracht 15, 2312 KM Leiden, The Netherlands.