Reprinted from Nature, Vol. 314, No. 6013, pp. 734-736, 25 April 1985

Education policy and the heritability of educational attainment

A. C. Heath*, K. Berg†, L. J. Eaves*, M. H. Solaas†, L. A. Corey*, J. Sundet‡, P. Magnus† & W. E. Nance*

* Department of Human Genetics, Medical College of Virginia, Richmond, Virginia 23298, USA
† Institute of Medical Genetics, and ‡ Institute of Psychology, University of Oslo, PO Box 1036, Blindern, Oslo 3, Norway

Many workers assume that genetically determined differences in intellectual ability[1–5] will be influenced little by changes in educational policy or other environmental interventions[6,7]. Others[8,9], however, have suggested that increasing equality of educational opportunity will lead to an increase in the heritability of educational attainment. The resolution of this issue has been delayed until now because of the extremely large sample sizes which would be required[10]. Education data on twins and their parents, from the Norwegian twin panel[11,12], provide a unique opportunity to determine the impact on the heritability of educational attainment of the more liberal social and educational policies introduced in Norway after the Second World War[13]. As reported here, for individuals born before 1940 there is a strong effect of family background on educational attainment, accounting for 47% of the variance, though genetic factors account for an additional 41% of the variance. For females born after 1940 and before 1961, the relative importance of genetic (38–45%) and familial environmental (41–50%) differences changes very little. For males born during the same period, the broad heritability of educational attainment has increased substantially (67–74%), and the environmental impact of family background has correspondingly decreased (8–10%). For males, at least, having well-educated parents no longer predicts educational success, as measured by duration of education, independent of the individual's own innate abilities.

Adult twins from the Norwegian twin panel, a population-based register of like-sexed twins born in Norway between 1915 and 1961 and presently resident there, were asked, by mailed questionnaire, to report the duration of their own education (I, 0-7 yr; II, 8-9 yr; III, 10-12 yr; IV, over 12 years of education) and that of both their parents. Questionnaires were returned by one or both twins from 8,389 families. Each member of 4,608 twin pairs reported his own educational level. Although well-educated twin pairs were overrepresented in the sample of respondents, the sample obtained was more nearly representative of the adult population than the samples obtained in most volunteer studies[11]. Polychoric correlations[14] for educational attainment were computed between male and female monozygotic (MZ) and dizygotic (DZ) twin pairs, between the parents of twins, and between parent and offspring for twins born before 1940, during 1940-49 and during or after 1950 (Table 1). (We report correlations rather than variances and covariances[15], as we are not making any predictions about the effects of social policy on the variance in educational attainment.) Data from

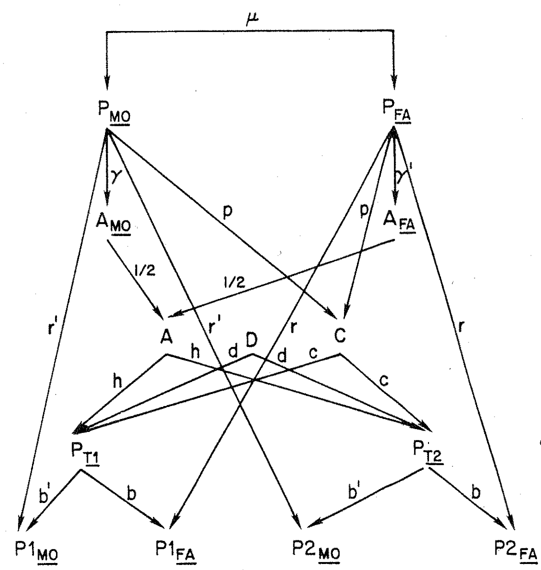

Fig. 1 Path diagram for MZ twins and their parents.

all the twins were used in computing the two-way contingency tables from which polychoric correlations between parent and offspring were calculated. Data for each member of a twin pair were treated as independent observations in computing the parent/offspring contingency tables. Data from only the first twin to report were used to compute the parent/parent contingency tables.

We found that correlations between male and female MZ twins have changed little over the time period studied, implying that the total contribution of familial factors (both genetic and environmental) to variation in educational attainment has remained constant. The change in the correlation between parents (0.85 to 0.79), though statistically significant because of the very large sample sizes, is also very small. There is a slight decline in the correlation between female DZ twins for twins born after 1940 (0.75 to 0.68), but a much more pronounced decline for male DZ twins (0.77 to 0.48); such a decline, relative to the MZ correlation, would be predicted if the influence of genetic differences on educational attainment has become more important in twins born after 1939. A corresponding decline in the parent/offspring correlations also occurs, the change in the parent/son correlations being greater than the change in the parent/daughter correlations.

Combining self-report data from the twins with the education data which they report for their parents allows us to overcome many of the limitations of the classical twin design[15-17]. Provided that the effects of dominance or genotype × age interaction are not too great, this design will resolve the contributions to variation in educational attainment of additive gene action, the environmental effects of parental education, other environmental effects shared by twins or siblings which are unrelated to parental education, and the genetic consequences of assortative mating. We shall not, however, be able to resolve alternative

Table 1 Secular changes in correlations between twins and their parents

	Twins' year of birth 1915-39			1940-49			1950-60		
	n	*r*	s.e.	*n*	*r*	s.e.	*n*	*r*	s.e.
Male MZ twins	259	0.86	0.03	253	0.82	0.03	370	0.85	0.02
Female MZ twins	405	0.89	0.02	342	0.85	0.03	518	0.89	0.02
Male DZ twins	313	0.77	0.04	284	0.48	0.06	463	0.47	0.04
Famale DZ twins	425	0.75	0.03	400	0.68	0.04	576	0.66	0.03
Mother/son	1,652	0.74	0.02	1,456	0.61	0.03	1,809	0.51	0.02
Mother/daughter	2,001	0.72	0.02	1,694	0.57	0.03	2,179	0.55	0.02
Father/son	1,631	0.73	0.02	1,433	0.64	0.02	1,786	0.55	0.02
Father/daughter	1,971	0.74	0.02	1,660	0.79	0.02	2,121	0.60	0.02
Mother/father	2,519	0.85	0.01	2,027	0.79	0.02	2,542	0.79	0.01

Table 2 Correlations between educational levels of twins and educational levels reported by them for their parents

Twin birth-cohort 1915–39

	(1)	(2)	(3)	(4)	(5)	(6)
Male MZ twins ($n = 193$)						
(1)	1.00	0.03	0.07	–	0.06	0.07
(2)	0.86	1.00	0.09	–	0.07	0.06
(3)	0.74	0.60	1.00	–	0.05	0.06
(4)	0.73	0.70	0.95	1.00	–	–
(5)	0.75	0.64	0.87	0.84	1.00	0.02
(6)	0.69	0.71	0.83	0.79	0.93	1.00
Female MZ twins ($n = 310$)						
(1)	1.00	0.02	–	0.06	0.06	0.06
(2)	0.90	1.00	–	0.06	0.06	0.06
(3)	0.80	0.71	1.00	–	–	–
(4)	0.74	0.72	0.98	1.00	0.04	0.04
(5)	0.71	0.65	0.88	0.85	1.00	0.01
(6)	0.67	0.67	0.79	0.86	0.96	1.00
Male DZ twins ($n = 243$)						
(1)	1.00	0.04	0.05	0.06	0.04	0.05
(2)	0.78	1.00	0.06	0.05	0.05	0.04
(3)	0.78	0.73	1.00	0.02	0.02	0.04
(4)	0.71	0.76	0.96	1.00	0.04	0.04
(5)	0.79	0.73	0.92	0.84	1.00	0.02
(6)	0.73	0.79	0.85	0.84	0.95	1.00
Female DZ twins ($n = 309$)						
(1)	1.00	0.04	0.05	0.06	0.05	0.05
(2)	0.78	1.00	0.05	0.05	0.05	0.04
(3)	0.75	0.75	1.00	–	0.03	0.05
(4)	0.65	0.76	0.96	1.00	0.05	0.05
(5)	0.74	0.77	0.89	0.80	1.00	0.01
(6)	0.73	0.80	0.81	0.77	0.96	1.00

Twin birth-cohort 1940–49

	(1)	(2)	(3)	(4)	(5)	(6)
Male MZ twins ($n = 203$)						
(1)	1.00	0.03	0.06	0.07	0.06	0.07
(2)	0.83	1.00	0.07	0.07	0.06	0.06
(3)	0.72	0.63	1.00	0.03	0.05	0.07
(4)	0.61	0.67	0.91	1.00	0.05	0.05
(5)	0.71	0.65	0.81	0.78	1.00	0.01
(6)	0.57	0.66	0.71	0.82	0.96	1.00
Female MZ twins ($n = 284$)						
(1)	1.00	0.03	0.06	0.07	0.05	0.05
(2)	0.84	1.00	0.07	0.06	0.07	0.06
(3)	0.63	0.49	1.00	–	0.04	0.06
(4)	0.61	0.64	0.95	1.00	0.05	0.05
(5)	0.68	0.48	0.83	0.72	1.00	–
(6)	0.70	0.63	0.71	0.78	0.96	1.00
Male DZ twins ($n = 235$)						
(1)	1.00	0.06	0.07	0.08	0.06	0.07
(2)	0.47	1.00	0.07	0.07	0.07	0.06
(3)	0.55	0.56	1.00	0.03	0.04	0.05
(4)	0.49	0.57	0.90	1.00	0.06	0.05
(5)	0.63	0.54	0.85	0.71	1.00	0.02
(6)	0.45	0.60	0.74	0.74	0.91	1.00
Female DZ twins ($n = 319$)						
(1)	1.00	0.04	0.07	0.07	0.06	0.06
(2)	0.72	1.00	0.06	0.06	0.05	0.05
(3)	0.49	0.60	1.00	–	0.05	0.05
(4)	0.39	0.62	0.95	1.00	0.05	0.05
(5)	0.57	0.64	0.76	0.69	1.00	0.01
(6)	0.47	0.65	0.72	0.75	0.96	1.00

Twin birth-cohort 1950–60

	(1)	(2)	(3)	(4)	(5)	(6)
Male MZ twins ($n = 255$)						
(1)	1.00	0.03	0.06	0.07	0.06	0.06
(2)	0.86	1.00	0.07	0.07	0.06	0.06
(3)	0.52	0.53	1.00	0.01	0.03	0.04
(4)	0.48	0.50	0.96	1.00	0.04	0.04
(5)	0.58	0.56	0.82	0.76	1.00	0.01
(6)	0.54	0.56	0.78	0.80	0.97	1.00
Female MZ twins ($n = 361$)						
(1)	1.00	0.02	0.05	0.06	0.05	0.05
(2)	0.89	1.00	0.05	0.05	0.04	0.05
(3)	0.55	0.58	1.00	0.01	0.03	0.03
(4)	0.50	0.58	0.97	1.00	0.03	0.03
(5)	0.61	0.65	0.83	0.77	1.00	0.01
(6)	0.56	0.65	0.77	0.82	0.95	1.00
Male DZ twins ($n = 318$)						
(1)	1.00	0.05	0.06	0.06	0.05	0.06
(2)	0.46	1.00	0.06	0.06	0.06	0.06
(3)	0.49	0.40	1.00	0.03	0.04	0.05
(4)	0.40	0.57	0.86	1.00	0.05	0.04
(5)	0.54	0.41	0.73	0.65	1.00	0.02
(6)	0.44	0.52	0.61	0.76	0.91	1.00
Female DZ twins ($n = 406$)						
(1)	1.00	0.04	0.05	0.05	0.04	0.04
(2)	0.69	1.00	0.05	0.05	0.05	0.05
(3)	0.56	0.46	1.00	0.01	0.03	0.03
(4)	0.48	0.52	0.94	1.00	0.04	0.03
(5)	0.60	0.49	0.83	0.70	1.00	0.01
(6)	0.60	0.56	0.78	0.81	0.91	1.00

Polychoric correlations are given in the lower triangle of each matrix, their standard errors in the upper triangle. (1) First twin's education, (2) second twin's education, (3) mother's education (first twin's report), (4) mother's education (second twin's report), (5) father's education (first twin's report), (6) father's education (second twin's report). –, No estimate of standard error possible due to very low expected cell frequencies.

hypotheses about mechanisms of mate selection and environmental transmission within families[18]. As we have separate reports of parental educational levels by each member of a twin pair, we can check the validity of such data.

Table 2 shows matrices of polychoric correlations, broken down by twin group and birth-cohort, between the educational levels of each twin, and the educational level reported by each twin for each parent; only twin pairs where each twin reported the educational levels of himself and both parents were used in computing these correlations. For some of the polychoric correlations in Table 2 no standard errors are given, because of very low expected frequencies for one or more cells of the two-way 4×4 tables from which the polychoric correlations are derived. Such coefficients may be strongly biased estimates of the true population correlations[14] and so were not used in any data analyses. Agreement between twins is very good, correlations between educational levels reported by each twin for the same parent ranging from 0.86 to 0.98. However, the correlation between a twin's educational level and that which he reports for either parent is consistently higher than the correlation between his own educational level and the educational level reported by his co-twin for the same parent. Fortunately, we can allow for the effects of this bias, by model-fitting.

Figure 1 presents a path model[12,19] for deriving expected correlations between the educational attainments of MZ twins and the educational attainments reported by them for their parents: P denotes true educational attainment (we assume that the twins make no error in reporting their own educational level), and P1 and P2 the educational attainments reported by first and second twins, respectively, for mother (subscript MO) or father (subscript FA). The standardized path regressions of true phenotypic value on additive genetic deviation (A), dominance deviation (D), familial environmental deviation (C) and random environmental deviation (E: omitted from the diagram to simplify presentation) are denoted by parameters h, d, c and e, or for females, in models which allow for sex-dependent effects, h', d', c' and e'. The path regression of additive genetic deviation on true phenotypic value is denoted by the parameter γ' in males, or γ in females, where $\gamma' = h + ac$, $\gamma = h' + ac'$, and a is the genotype/environmental correlation which may be expressed as a function of the other parameters of the model[3]. The path regressions of reported parental educational attainment on true parental educational attainment and on the twin's educational attainment are denoted by b and r for fathers, b' and r' for mothers. Path regressions of offspring familial environmental value on true maternal or paternal educational attainment are denoted by the environmental transmission parameter p. The parameter μ denotes the correlation between the educational attainments of spouses, which is assumed to arise through phenotypic assortative mating[20]. An equivalent diagram can be drawn for DZ twins, and expected correlations derived.

Models were fitted to the polychoric correlations in Table 2 by nonlinear weighted least-squares, using the standard errors of the polychoric correlations as weights, and treating the correlations as though they were independent. Strictly, these correlations are not independent. However, ignoring the correlation between correlations has been found to have little effect on the parameter estimates obtained[21]. Model-fitting yields estimates of model parameters and provides an approximate χ^2 value with which to assess the fit of a given model, and compare its fit with that of other models. The validity of the χ^2 statistic will depend on how closely the sampling distribution of the polychoric correlation coefficients approximates to the multivariate normal distribution. With these very large sample sizes we expect the approximation to be close.

For twins born before 1940, there is no evidence for sex differences in the familial transmission of educational attainment. A purely environmental model gives a good fit to the data ($\chi^2_{44} = 42.18$, $P = 0.55$), although allowing for additive genetic variance gives a significant improvement in fit ($\chi^2_1 = 7.81$,

$P < 0.01$). Parameter estimates obtained for the best model are $h = 0.64$, $c = 0.53$, $\mu = 0.86$, $p = 0.19$, $r = 0.84$ *and* $b = 0.20$, so that 41% of the variance in educational attainment is attributable to additive genetic factors (h^2), 28% to the environmental effects of family background (c^2), and 19% to genotype/environmental covariation ($2hac$), the remaining 13% of the variance being attributable to environmental effects not shared by relatives. Our estimate of the importance of genotype/environmental covariation depends critically on the assumptions made about the mechanisms of environmental transmission and mate selection, which cannot be tested with these data. If we were to make certain alternative assumptions, no genotype/environmental covariation would be predicted[5] and the contribution of the family enviroment would be estimated at 47%.

For twins born after 1939, we find significant heterogeneity of gene expression and environmental effects across sexes. Purely environmental models are rejected with a high level of significance (1940-49: $\chi^2_{50} = 89.44$, $P < 0.001$; 1950-60: $\chi^2_{53} = 169.55$, $P < 0.001$). Models which allow for both additive gene action and environmental transmission from parent to offspring yield negative estimates for the environmental transmission parameter p, which suggests that any environmental effects of parents on their offspring are being masked by genetic dominance in this design. Models which allow for sex-dependent additive gene action, dominance and environmental effects shared by twins, but no environmental transmission from parent to offspring, give the best fits to these data (1940-49: $\chi^2_{48} = 52.61$, d.f. = 48; $\chi^2_{51} = 56.59$, d.f. = 51). Parameter estimates obtained for this model are, for twins born between 1940 and 1949: $h = 0.70$, $h' = 0.67$, $d = 0.50$, $d' = 0$, $c = 0.29$, $c' = 0.64$, $\mu = 0.72$, $r = 0.82$, $b = 0.31$; and, for twins born between 1950 and 1960: $h = 0.52$, $h' = 0.53$, $d = 0.63$, $d' = 0.32$, $c = 0.45$, $c' = 0.71$, $\mu = 0.73$, $r = 0.82$, $b = 0.40$. These estimates imply that genetic factors (including genetic dominance) account for 74% of the variance in educational attainment in male twins born in 1940-49, and 67% in male twins born in 1950-60, but account for only 45% of the variance in female twins born in 1940-49, and 38% in female twins born after 1949. Conversely, environmental effects shared by twins account for only 8% (1940-49) or 20% (1950-60) of the variance in males, but for 41% (1940-49) or 50% (1950-60) in females. As genetic dominance is a major source of variation in IQ[1-3,22], the absence of evidence for dominance in the pre-1940 sample might be interpreted as evidence against a major influence of IQ on educational attainment in individuals born before 1940. However, dominance may be masked by strong environmental transmission in these data.

The results presented here are clearly consistent with the hypothesis that the importance of genetic influences on educational attainment is subject to secular change. Other explanations of our findings seem implausible. Perhaps more pre-war monozygotic male twin pairs have been misclassified as dizygotic, thus inflating the pre-war dizygotic male correlation? The proportion of pre-war male twins who are dizygotic (45.3%) differs little from the proportion of dizygotic twins in the other twin cohorts (44.4-48.8%). Perhaps a 'special twin environment' effect[23] is involved? As there has been a progressive decrease in the influence of parental educational levels on offspring educational attainment (see Table 1), we would expect the difference between male MZ and DZ twin correlations to have declined rather than increased in recent years. The most likely explanation, confirming the hypothesis of Scarr-Salapatek[8], is that increased educational opportunity has led to an increased dependence of educational attainment on innate ability.

This work was supported in part by NIH grants GM30250, HL28922 and HD10291, and by grants from the Norwegian Research Council for Science and the Humanities.

Received 6 November 1984; accepted 8 February 1985.

1. Jinks, J. L. & Fulker, D. W. *Psychol. Bull.* **73**, 311-349 (1970).
2. Jinks, J. L. & Eaves, L. J. *Nature* **248**, 287-289 (1974).
3. Loehlin, J. C. *Behavl Genet.* **8**, 415-426 (1978).
4. Rice, J., Cloninger, C. R. & Reich, T. *Behavl Genet.* **10**, 73-92 (1980).
5. Rao, D. C. *et al. Genet. Res.* **39**, 187-198 (1982).
6. Jensen, A. R. *Genetics and Education* (Methuen, London, 1972).
7. Jensen, A. R. *Educability and Group Differences* (Methuen, London, 1973).
8. Scarr-Salapatek, S. *Science* **174**, 1223-1228 (1971).
9. Scarr-Salapatek, S. *Science* **174**, 1285-1295 (1971).
10. Eaves, L. J. & Jinks, J. L. *Nature* **240**, 84-88 (1972).
11. Magnus, P., Berg, K. & Nance, W. E. *Clin. Genet.* **24**, 103-112 (1983).
12. Heath, A. C. *et al. Behavl Genet.* (in the press).
13. *Educational Statistics* (Central Bureau of Statistics, Oslo, 1978).
14. Olsson, U. *Psychometrika* **44**, 443-460 (1979).
15. Eaves, L. J. *et al. Heredity* **41**, 249-320 (1978).
16. Eaves, L. J. *J. R. statist. Soc. Ser. A* **140**, 324-355 (1976).
17. Martin, N. G. *et al. Heredity* **40**, 97-116 (1978).
18. Heath, A. C. & Eaves, L. J. *Behavl Genet.* **15**, 15-30 (1985).
19. Wright, S. *Ann. math. Statist.* **5**, 161-215 (1934).
20. Fisher, R. A. *Trans. R. Soc. Edinb.* **52**, 399-433 (1918).
21. Rao, D. C. *et al. Behavl Genet.* **7**, 147-159 (1977).
22. Bashi, J. *Nature* **266**, 440-442 (1977).
23. Kamin, L. J. *The Science and Politics of IQ* (Wiley, New York, 1974).