Achievement and Ascription in Educational Attainment: Genetic and Environmental Influences on Adolescent Schooling

François Nielsen

Supplementary Materials

1. Shared Environments of Twins

The model estimated with the 6 types of siblings assumes that shared environmental effects are the same for twins and full siblings. Since twins share the same womb and are exact contemporaries, it could be argued that the environments of twins are more similar than that of full siblings. An alternative argument is that the environments of twins, who compete for maternal resources in the womb, might be *less* similar than the environments of full siblings who do not compete to the same extent. If the environments of twins are more similar one would expect correlations between twins to be larger than correlations between siblings. In fact Table 1 shows the opposite pattern: correlations for full siblings are systematically *larger* than those for DZ twins (.411>.356 for VIQ, .360>.332 for GPA, .332>.264 for CPL). This pattern is inconsistent with the hypothesis that the environments are more similar for twins than for full siblings, although it is consistent with the alternative hypothesis of greater dissimilarity of the environments of twins because of more intense competition.

Another concern that looms large in critiques of twin studies is the possibility that the greater similarity of MZ twins compared to DZ twins is due in part to greater similarity of the environments of MZ twins (perhaps because their greater physical resemblance causes them to be treated more alike by parents and others). Extensive research has failed to support this hypothesis; the capsule result is that the environments of MZ twins are indeed more alike, but to the extent that environmental similarity is associated with phenotypic similarity it is as a consequence rather than a cause, in the sense that MZ twins are treated more alike because they look and behave more alike, not that the more similar environments make them more alike (Neale and Maes, forthcoming; Rowe 1994).

2. Estimation With MZ and DZ Twins Only

Model compared in Table 2 in the main article have all relatively small p values, even though the best ones have highly satisfactory levels of fit as measured by RMSEA. The small probabilities may be due in part to the strategy of using the 6 different types of sibling pairs for estimation, since this both produces a very large sample and increases the likelihood of heterogeneity of the data in relation to the model. An additional clue to such heterogeneity is that the share of total model χ^2 contributed by the half-sibling (HS) pairs is especially high (39.1) compared to other groups (e.g., 20.9 for DZ twins). HS pairs may be especially heterogeneous because of different exposures of siblings to the current family environment and

different consequences of assortative mating (see later). One would therefore expect the fit to be better (in terms of the p-value of the model) if estimation is restricted to twins (MZ and DZ) instead of all 6 types of sibling pairs.

------ Supplemental Table S1 about here ------
------ Supplemental Table S2 about here ------

I re-estimated the model using twins only (Table 1 and Table 2). When estimated from the twin pairs (MZ and DZ) only, the favored model is AE, which does not contain a shared-environmental component. This time χ^2=30.645 with 30 df, p=.433, a reassuring fit. The favored model yields heritability estimates that are somewhat larger (62% to 69%) than for the favored model AC_1E_d with all sibling types (54% to 67%).

3. Assortative Mating

I carried out additional analyses to investigate the potential impact of assortative mating on the model of educational attainment. *Assortative mating* (or *homogamy*) is the tendency of people to find mates who are similar to themselves with respect to a given trait. The degree of assortative mating is measured as the correlation of a trait between spouses. In humans assortative mating is high for cognitive ability (e.g., IQ) and even higher for educational attainment (Blossfeld and Timm 2003; Mare 1991; Smits, Ultee and Lammers 1998). Deriving the consequences of assortative mating is complicated, as the outcome depends on the exact nature of the mechanisms producing the correlation between spouses (Falconer and MacKay 1996: 174-177). In general assortative mating tends to increases the correlation between genotypic values of full siblings (including DZ twins) for the assortative trait and other traits correlated with it. The reason is that if spouses have similar values for a heritable trait, they are more likely to have the same alleles for genes affecting that trait; thus their offspring are more likely to inherit these alleles than they would if the parents were randomly selected from the population. The correlation k between genotypic values of full siblings, which is .5 under random mating, is thus larger when assortative mating is present. Since assortative mating does not affect the genetic correlation between MZ twins (who already have the same genes), it tends to attenuate estimates of heritability based on comparing correlations for MZ twins and DZ twins (or non twin full siblings).

Using empirical estimates of the correlations of educational outcomes among spouses, it would be possible to estimate the effect of assortment iteratively within the model (see e.g., Chipuer, Rovine, and Plomin 1990), but this solution is not satisfactory for subtle theoretical reasons.[1] Instead I carried out a crude sensitivity analysis with the twins data, re-estimating the ACE version of the model after replacing the value of k=.5 for the genotypic correlation between DZ twins with two higher values k=.6 and

[1] This is because the effect of assortment depends on the heritability of a trait *at the age of mating* (i.e., for the parents of the adolescents in the study) rather than the heritability of the trait in adolescents that the model estimates (Falconer and MacKay 1996: 176, Table 10.6), and the heritability of a trait can vary considerably over the life span (e.g., Plomin and Petrill 1997). Thus the procedure used by Chipuer et al. (1990) of simultaneous estimation of heritability and assortment effects by "manual" iteration of the model is not entirely justified theoretically.

k=.65 derived from reasonable sets of assumptions about heritability and spousal correlations.[2] Estimates for the alternative specifications show higher heritability estimates with higher k values for VIQ: h^2 is .606, .665, and .671 for k = .5, .6, and .65, respectively. There is a smaller increase in h^2 estimates with k for GPA and CPL. Estimated environmentalities become zero when k=.65. Since AE is the favored model with the twin data it appears that allowing for assortative mating does not affect the qualitative conclusions obtained under the assumption of random mating.

4. Genetic Dominance

As mentioned earlier, the fact that the correlation of CPL for DZ twins (.264) is less than half that for MZ twins (.663) suggests a genetic dominance effect. The CPL correlation for full siblings (.332) is not less than half that for MZ twins, however. Genetic dominance consists of non-additive effects of genes due to the fact that some genes have dominant and recessive alleles. Dominance depends on the pattern of alleles. MZ twins, who have the same genes, share the same patterns of alleles and thus the entire dominance effect. Ordinary siblings (including DZ twins) share only .25 the dominance effect, which is the probability that ordinary siblings both inherit the same pattern of alleles. More distantly related individuals do not share any dominance effect. Some studies find dominance effects for variables in the cognitive domain. Dominance and shared environment effects cannot be estimated in the same model. I estimated an ADE model, in which the C component is replace by a D component for all three outcomes. The ADE model can be reduced to a single factor AD_1E model, in which the dominance effect is strongest for CPL (standardized coefficients of D_1: are .151 for VIQ, .029 for GPA, -.371 for CPL). However the best dominance model (AD_1E with χ^2 = 177.245 for 108 df) is not superior to the corresponding AC_1E model.

5. Additional References

Blossfeld, Hans-Peter and A. Timm, eds. 2003. *Who Marries Whom? Educational Systems as Marriage Markets in Modern Societies.* Dordrecht, Netherlands: Kluwer.

Chipuer, Heather M., Michael J. Rovine, and Robert Plomin. 1990. "LISREL Modeling: Genetic and Environment Influences on IQ Revisited. *Intelligence* 14 (1):11-29.

Mare, Robert D. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Review* 56 (1):15-32.

Smits, J., W. Ultee and J. Lammers. 1998. "Educational Homogamy in 65 Countries: An Explanation of Differences in Openness Using Country-Level Explanatory Variables." *American Sociological Review* 63: 264-285.

---

[2] The genotypic correlation between first degree siblings (such as DZ and FS) is calculated as k=.5(1+m) where m is the correlation between the genotypes of parents due to assortment estimated as $m=rh^2$ where r is the phenotypic correlation between parents and h^2 the heritability of the trait at the time of mating (Falconer and MacKay 1996: 176, Table 10.6). The two values for k used in the analysis are derived assuming that k is approximately the same for all three phenotypes, and assuming first r=.33 (r for IQ from Chipuer et al. 1990) and h^2=.6, so k=.6; second that r=.5 and h^2=.6, so k=.65.

Supplemental Table S1 – Model comparisons – MZ and DZ twins only

Model		Fit statistics					Tests		
	χ^2	Df	p	AIC	RMSEA	Test	$\Delta\chi^2$	Δ df	p
1. BACE	28.676	21	.122	-13.324	.041				
2. BAE	30.645	27	.286	-23.355	.024	2 vs. 1	1.969	6	.923
3. BCE	97.718	27	.000	43.718	.111	3 vs. 1	69.042	6	.000
4. ACE	28.676	24	.233	-19.324	.030	4 vs. 1	0.000	3	*
5. AE[a]	30.645	30	.433	-29.355	.007	5 vs. 4	1.969	6	.923
6. CE	97.718	30	.000	37.718	.103	6 vs. 4	69.042	6	.000
7. ACE$_d$	37.790	27	.081	-16.210	.039	7 vs. 4	9.114	3	.028
8. AE$_d$	41.511	33	.147	-24.489	.033	8 vs. 4	12.835	9	.170
						8 vs. 5	10.866	3	.012

Note: B = phenotypic paths; A = genetic paths; C = common environment paths;
E = specific environment paths; E$_d$: off diagonal elements of E fixed.
[a] favored model
* probability incalculable

Supplemental Table S2 – Variance and covariance components: Proportions of
total predicted variance or covariance associated with genetic and specific
environment factors for favored AE model - MZ and DZ twins only

	Genetic factors			Specific environment		
	VIQ	GPA	CPL	VIQ	GPA	CPL
VIQ	.693			.307		
GPA	.798	.652		.202	.348	
CPL	.826	.844	.617	.174	.156	.383

Note: VIQ = verbal IQ; GPA = grade point average; CPL = college plans.