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

Francois Nielsen, University of North Carolina at Chapel Hill

Abstract

The classic ("status attainment") model of educational and occupational attainment suffers from three related shortcomings when used as a tool for comparative or policy-oriented research on social mobility: (1) ambiguity of model parameters as measures of opportunity for achievement vs. ascription; (2) vulnerability to incomplete specification of family background; and (3) confounding of environmental and genetic influences. These issues can be addressed in part by using a ("behavior genetic") model that distinguishes variance components associated with genetic endowment, shared (or common) family environment, and unshared (or specific) environment. Size of the genetic component (heritability) measures opportunity for achievement; size of the shared environment component (environmentality) measures social ascription. A multivariate behavior genetic model of adolescent verbal IQ, grade point average and college plans is estimated using data for six types of adolescent sibling pairs living in the same household: MZ twins, DZ twins, full siblings, half siblings, cousins and non-related siblings. Results show large genetic components, relatively small shared environmental components, and large unshared environmental components for all three outcomes. Parameters of the behavior genetic model can be used to compare mobility regimes across social contexts and the model therefore provides an important tool for comparative social mobility research.

Introduction

Recurring questions in the literature on social stratification and mobility involve evaluating the role of ascription vs. opportunity in educational and socio-economic attainment: whether some societies are more open than others, whether ascription is declining in the course of industrialization, or whether a given subgroup of society enjoys fewer opportunities. The status attainment model introduced by Blau and Duncan (1967) is one of the principal tools used to address such questions in a comparative perspective (Breen and Jonsson 2005). Opportunity for achievement is typically associated with the effects on attainment of variables presumed to reflect inherent individual qualities and effort (e.g., cognitive ability, education); ascription is associated with family background characteristics (e.g., parental education, family SES).

This paper is dedicated to the memory of Bruce K. Eckland, 1932-1999. Data are from Add Health, a program designed by J. Richard Udry, Peter S. Bearman and Kathleen Mullan Harris, and funded by a grant, P01-HD31921, from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald Rindfuss and Barbara Entwisle for assistance in the original design. Data analysis was supported by a small grant from University Research Council at UNC. For help with the data, encouragement and advice I wish to thank Daniel Adkins, Ken Bollen, Tom DiPrete, Guang Guo, Kathi Harker Tillman, Kathie Harris, Rosemary Hopcroft, Mike Shanahan and Dick Udry. Direct correspondence to Francois Nielsen, Department of Sociology, University of North Carolina, Chapel Hill, NC 27599-3210. E-mail francois_nielsen@unc.edu.

© The University of North Carolina Press

Social Forces, Volume 85, Number 1, September 2006
The paper presents a case that the status attainment model is inadequate as a tool for comparing social mobility regimes due to ambiguity of interpretations of model parameters as representing opportunity vs. ascription, vulnerability to bias due to incomplete specification of family background, and confounding of environmental and genetic influences. Some of the shortcomings of the attainment model can be alleviated by using a behavior genetic model of attainment that explicitly distinguishes between genetic and environmental influences on educational and socio-economic outcomes. To illustrate this point I use data on pairs of siblings who differ in their degree of biological relatedness to estimate a model of schooling involving three measures: verbal IQ, grade point average and college plans. The model decomposes the variance in verbal IQ and the educational measures into components corresponding to effects of genetic endowment, shared family environment and unshared (individual-specific) environment.

The paper develops an idea that has been expressed earlier by researchers in several fields, including sociology, but has not taken root in our field. This is the view that behavior genetic models of attainment are an important tool of comparative social mobility research because they provide consistent measures of opportunity and ascription in a system of stratification: ascription is measured as the proportion of variation in attainment attributable to the shared environment (environmentality) and opportunity as the proportion of variation attributable to genes (heritability).

Subsequent sections discuss problems with the attainment model and the potential usefulness of a behavior genetic model of attainment. Later sections discuss the data, develop the empirical model of adolescent schooling, present the results of the analysis, and finally discusses broader implications of the findings for comparative social stratification research. Existing literature is used to provide a "demo" illustrating genetically-aware social mobility research.

**Achievement & Ascription**

Following Blau and Duncan (1967) sociologists have used sets of recursive equations to describe the process of educational and occupational attainment of individuals in a system of stratification. A typical model of status attainment (as this type of research came to be known) might consist of three equations:

\[
\begin{align*}
(1) & \quad R_{sIQ} = f_1(F_{sOcc}, F_{sEd}) \\
(2) & \quad R_{sEd} = f_2(R_{sIQ}, F_{sOcc}, F_{sEd}) \\
(3) & \quad R_{sOcc} = f_3(R_{sEd}, R_{sIQ}, F_{sOcc}, F_{sEd})
\end{align*}
\]

where IQ denotes a measure of cognitive ability, Ed is educational achievement, and Occ is a measure of occupational prestige; R and F refer to the respondent and the respondent’s father, respectively, and the functions are typically specified as linear (Duncan, Featherman and Duncan 1972). The career of the individual is envisioned as a process in which each level of attainment is a function of previous attainment and characteristics of the family of origin.

Blau and Duncan (1967) discovered an empirical pattern that would be often replicated in later research. First, the direct effect of FsOcc on RsOcc (controlling for RsEd) is small, from which they concluded that there is relatively little social
ascription. Second, the direct effect of RsEd on RsOcc is large; from this they concluded that (1) education serves to reproduce inequality (as most of the correlation between FsOcc and RsOcc is indirect, through RsEd), and (2) there is much opportunity for achievement (as the correlation of RsEd with RsOcc is driven by RsEd residuals representing unmeasured individual resources independent of family background). Overall Blau and Duncan interpreted their findings as reflecting a pattern of low ascription in modern industrial society.

Stratification researchers adopted the attainment model as a means of capturing opportunity for achievement (or openness). Opportunity is associated with effects of intermediate achievement variables (such as RsEd); ascription with effects of background variables (such as FsOcc). With these substantive interpretations of the parameters, the status attainment model can be used as a comparative device to evaluate the relative openness of the stratification system in different societies or historical periods (Breen and Jonsson 2005; Ganzeboom, Treiman and Ultee 1991). The attainment model — resonating with powerful themes of social justice — is a centerpiece of the policy-oriented discourse on social stratification. In this normative vein, model parameters are used to contrast extent of meritocracy vs. enduring strength of social ascription (Olneck 1977:151).

The attainment model is unfortunately inadequate as a tool for comparative stratification research or for use in normative debates because of shortcomings related to (1) interpretations of model parameters in terms of opportunity vs. ascription, (2) the possibility of estimation bias due to incomplete specification of family background, and (3) confounding of environmental and genetic influences.

**Interpretation Issues**

Interpretations of the effects of background (such as family SES) or intermediate variables (such as IQ or education) are substantively ambiguous because intermediate variables may reflect social inheritance rather than intrinsic talent or effort. The difficulty in linking normative concepts to model parameters is exposed in sharp relief in the debate surrounding the finding of Herrnstein and Murray (1994) that, controlling for SES of family of origin, cognitive ability (IQ) has a strong effect on educational and socio-economic outcomes. The authors interpret this pattern (together with other evidence) as indicating increasing returns to cognitive ability (and opportunity) in contemporary U.S. society. Critics have responded by questioning the strength of ability effects and/or arguing that IQ scores reflect social inheritance rather than native talent, so that IQ effects reflect social reproduction rather than opportunity (Fischer et al. 1996).

**Model Specification Issues**

The second issue concerns the open-ended specification of family background. Herrnstein and Murray (1994) measure family background with a composite SES index based on parental education and income. Critics point out that the SES composite does not adequately control for all the relevant aspects of family background. Leaving important aspects of family environment out of the model produces specification bias which artificially inflates the apparent effect
of cognitive ability and thus the evidence for opportunity. If all these factors were properly controlled, critics claim, the effect of ability would be reduced or disappear. Re-estimating some of Herrnstein and Murray’s (1994) models with more detailed family background measures does reduce (but typically does not eliminate) the estimated effects of cognitive ability on educational and socio-economic outcomes (Fischer et al. 1996; Korenman and Winship 2000). This exchange illustrates a general problem with the classic attainment model: the task of controlling for family background with measured variables is inherently open-ended as there is no way to guarantee that all relevant aspects of the family environment have been measured and included in the model.

Confounding of Genetic and Environmental Effects

The role of genes in occupational mobility and the problem of confounding environmental and genetic influences were already addressed in the mainstream sociological literature almost 40 years ago. In a remarkable paper Eckland (1967) argued that the social mobility research of his days (then mainly based on mobility tables) was flawed as it assumes, in estimating aggregate mobility, a null model in which sons from any category of origin are equally likely to reach any category of destination. If the abilities to reach certain destinations are in part genetically determined and, as a result, unequally distributed, it follows that sons from certain origins will be more likely to reach certain destinations. Without control for these genetic effects, the resulting asymmetry will be falsely attributed to a lack of perfect mobility. Thus, Eckland claimed, the degree of social mobility cannot be properly estimated without controlling for the association between origin and destination due to genetic transmission of abilities. The implication is that in measuring social rigidity or ascription, any association between occupational achievements of father and son due to genetic causes should be partialled out from the overall association; social ascription is identified with non-genetic (presumably environmental) causes of inter-generational transmission of status (see also Eckland 1979).

In another landmark article, Scarr and Weinberg (1978) presented results from a study of adopted children showing that the correlations between characteristics of adoptive parents and children’s cognitive outcomes are very small, whereas the correlations between these outcomes and cognitive ability of the birth mother are larger, a pattern suggesting that effects of family background variables on achievement of children in biological families are due to genetic causes rather than the environmental mechanisms sociologists surmise. Such results imply that the association between parental education and educational attainment of a biological child may reflect in part genetic transmission of educational abilities from parent to offspring rather than purely environmental influences. Thus the effect of parental education is not a good measure of social inheritance.

The argument underlying both Eckland (1967) and Scarr and Weinberg (1978) is that biological inheritance produces associations between background variables (e.g., parental education and occupation) and respondent outcomes (e.g., IQ and education) that are conceptually distinct from associations due to environmental mechanisms of social reproduction. In the usual attainment model estimated
from respondents raised in their biological families effects of background variables on cognitive, educational and occupational attainments of respondent are confounded with genetic influences.

The Role of Genes

Even though their articles were published in a mainstream sociology journal, Eckland's (1967) and Scarr and Weinberg's (1978) critique of the classical attainment model did not have much impact on the field. A literature investigating the genetic bases of educational and occupational attainment flourished briefly in the 1970s, involving economists and sociologists (see contributions in Taubman 1977). The work of Jencks and colleagues in sociology was a prominent part of that literature (Jencks 1980, 1992; Jencks and Brown 1977; Jencks et al. 1972). The research tradition persevered in economics (Behrman et al. 1980; Behrman, Pollack and Taubman 1995) where it is experiencing a mini-renaissance (Björklund, Jäntti and Solon 2005), but faded in sociology. Meanwhile since the early 1970s behavior genetic research has experienced explosive growth in the fields of mental health, child development and cognitive and personality psychology (Plomin et al. 1997), producing a sophisticated statistical methodology to disentangle genetic and environmental influences on behavior using data on twins, adopted children and other relatives (Neale and Maes, forthcoming; Rodgers et al. 2001; Rowe and Teachman 2001; Shanahan, Hofer and Shanahan 2003). Findings on the substantial role of heredity in cognitive ability, personality and educational and socio-economic outcomes have also accumulated (Rowe 1994). These conclusions on the role of genes in behavior are now finding their way to a broader public (Pinker 2002).

Behavior genetic models partition the variance of a measurable trait (a phenotype) into a component due to genetic inheritance, a component due to the shared (common, between-families) environment of siblings (aspects of the family and the larger rearing environment that tend to make siblings reared together alike), and a component due to the unshared (specific, unique, within-family) environment of a sibling (environmental factors that differ among siblings and tend to make them different). The ratio of each variance component to the total phenotypic variance is termed heritability, environmentality and specificity, respectively. This decomposition of the phenotypic variance is the key to resolving the difficulties of the attainment model discussed earlier.

The realization has emerged in the behavior genetic literature that the behavior genetic model provides a clear conceptual distinction between opportunity for achievement and ascription. The shared environment component, as it reflects the combined impact of such factors as social class, parental network of acquaintances, minority status, neighborhood characteristics and other aspects of the rearing environment that constitute the common experience of siblings in a family and affect their outcomes (such as educational achievement) in a similar way, captures the background characteristics that stratification researchers presumably have in mind when they conceptualize mechanisms of social reproduction and the ascriptive assignment of status. As Rowe (1994:33) writes: "This ratio [of shared environmental variation to total phenotypic variation] has
important policy implications, because it indicates how a phenotype might be changed by altering the rearing conditions of children with poor phenotypes to be like those of children with good ones. The greater the shared rearing estimate, the more change can be expected to follow from changing rearing conditions.” (See also Jencks 1980:734.) The shared environment component of the total variance in outcome can thus be interpreted as a measure of ascription (Heath et al. 1985, Scarr-Salapatek 1971).

The genetic component, reflecting the extent to which individuals are able to achieve their genetic potential for the trait, can be taken as a measure of opportunity for achievement.³ In this interpretation, the phenotype is viewed as resulting from the interaction of genetic endowment with the social environment. Different social environments are viewed as more or less restrictive of the full expression of genetic potential. Favorable environments, permitting fuller expression of potential, are characterized by high heritability. Unfavorable environments, inhibiting expression of native talent, are characterized by low heritability. Heritability measures realization of genetic potential and therefore opportunity for achievement (Guo and Stearns 2002).

The unshared environment represents a combination of measurement error and idiosyncratic environmental influences that affect siblings in different ways; examples are birth order, a childhood disease that affects one sibling and not another, or association of siblings with different peer groups (Jensen 1997, Turkheimer and Waldron 2000).

The problem of incomplete specification of family background that plagues the attainment model is rendered moot in the behavior genetic model, which measures the overall impact of the shared environment in black box fashion, without actually measuring, or even identifying, the variables involved. Thus it is no longer possible to spuriously overestimate the role of an achievement variable such as IQ or underestimate the role of family background by leaving an important environmental characteristic out of the model.⁴ Finally the behavior genetic approach, by distinguishing among variance components due to genes, shared environment and unshared environment effectively disentangles the confounding of genetic and environmental influences that plagues the attainment model. Heritability, environmentality and specificity can be compared across social settings, groups or social systems. Such comparisons potentially inform a comparative sociology of social stratification and mobility, permitting statements on the relative degrees of social openness across societies and historical periods, and across groups and social contexts within a given society. The components are also meaningful conversation pieces in normative debates concerning social inequality.

Data

Data are from the first two waves of AddHealth, a school-based longitudinal study of adolescents in grades 7 through 12 (Udry 1998). Data were collected from 1994 through 1996. Siblings living in the same household were identified, and if necessary, a sibling was added to the sample to complete a pair. Pairs were classified as monozygotic twins (MZ, N = 170), dizygotic twins (DZ, N =
Achievement & Ascription in Education

290), full siblings (FS, N = 702), half siblings (HS, N = 242), cousins (CO, N = 105), and non-related (NR, N = 174). Ns are pairs actually used in the analysis (on determination of zygosity see Rowe and Jacobson 1998). All sibling pairs are used, even though data on an individual may be repeated when an individual is member of more than one pair (on this, see Eaves et al. 1999:67).

Verbal IQ (VIQ) measures verbal cognitive ability; it is the score on a test consisting of 87 items from the Peabody Picture Vocabulary Test (PPVT) (Neiss and Rowe 2000). Grade Point Average (GPA) is the average answer to four questions in each of two waves concerning recent grades in English or language arts, mathematics, history or social studies, and science. College Plans (CPL) is the average answer to questions on college aspiration and expectation asked in both waves. Only self-identified blacks and non-Hispanic whites were included in the analysis and a few outlying observations with VIQ scores below 50 were excluded.

Average VIQ score is about 12 points less for blacks compared to whites, and about 2 points less for females compared to males. There are also significant race and sex differences in GPA, and a smaller sex (but not race) difference in CPL. Such differences may inflate estimates of shared environmental effects in behavior genetic models (Rowe, Jacobson and Van den Oord 1999). To control these differences each variable was standardized within each of four race by sex categories (black female, black male, white female, white male). Zero-order associations of the measures with age are minute, so age was not controlled further.

Model and Methods

The school attainment model is depicted in Figure 1. It is a structural equations model (SEM) containing both observed variables represented by squares and unmeasured, latent variables represented by circles. Measured variables are VIQ, GPA and CPL. Straight arrows between VIQ and GPA and CPL, and between GPA and CPL, represent elements of an educational attainment model relating the three variables in a recursive fashion, so that VIQ is assumed to affect both GPA and college plans; GPA also directly affects CPL. This subset of effects is termed the phenotypic part of the model.

The behavior genetic aspects are embodied in the latent variables. Measured variables for a given sibling are functions of three genetic factors (A$_1$, A$_2$, A$_3$) assumed uncorrelated within a given sibling. Genetic factor A$_1$ affects all three phenotypes VIQ, GPA and CPL; A$_2$ affects only GPA and CPL; A$_3$ affects CPL only. This patterning of the paths is called a Cholesky factorization (Neale and Maes, forthcoming; see empirical examples in Emde and Hewitt 2001). The Cholesky structure is adopted on substantive and methodological grounds further explained in Nielsen (2006).

Each genetic factor for one sibling is correlated with the corresponding factor for the other sibling by a quantity $k$ corresponding to their degree of relatedness (proportion of genes shared by common descent). MZ twins have identical genes, so $k$ is always equal to 1. When mating is random (a strong assumption that is relaxed later) DZ twins and full siblings share half their genes, so $k = .5$. The parameter $k$ is .25 for half siblings, .125 for cousins and 0 for non-related siblings.
Note: VIQ = verbal IQ, GPA = grade point average, CPL = college plans. k fixed to 1.0 (MZ); .5 (DZ, FS); .25 (HS); .125 (CO); 0 (NR).

It is the ability to use genetic theory to specify in advance the association between genotypes of siblings – an idea going back to Fisher (1918) – that gives behavior genetic models the leverage to disentangle genetic from environmental effects.

Latent factors $C_1$, $C_2$ and $C_3$ represent the \textit{shared or common} environment of siblings, which corresponds to the rearing environment, consisting of such variables as social class or family SES, quality of schools in the community, family connections, etc., that affect siblings reared in the same environment in the same ways and thus contribute to make them phenotypically \textit{similar}. The shared environment is also decomposed into a Cholesky structure with three orthogonal factors ($C_1$, $C_2$ and $C_3$). As the shared environment is assumed to affect each sibling in the same way, the correlation of each $C$ factor across siblings is fixed to 1.

Latent factors $E_1$, $E_2$ and $E_3$ represent the \textit{unshared or unique} environment consisting of sibling-specific experiences that contribute to make siblings phenotypically \textit{different} from each other. Such differentiating environmental influences might include parental preference, birth order, influences of different teachers or peers, or a disease affecting one sibling but not the other. In this model the specific environment also includes errors of measurement in the variables. The unshared environment factors constitute a Cholesky structure. The sibling-specificity assumption is implemented by fixing to 0 the correlation of each factor across siblings.
Achievement & Ascription in Education • 201

The model is estimated by deriving mathematically the expected covariance matrix of the observed variables for each type of sibling pairs as a function of model parameters (see Nielsen 2006 for derivation of the model and Mx script). Each matrix has dimension 6 by 6, with rows and columns corresponding to the observed phenotypes for each sibling (i.e., VIQ₁, GPA₁, CPL₁, VIQ₂, GPA₂, CPL₂, where subscripts denote siblings in a pair). The parameters of the model are then estimated simultaneously for the six types of siblings by minimizing the discrepancies between expected and observed covariance matrices according to the maximum likelihood (ML) criterion (Bollen 1989, Loehlin 2004). The SEM program Mx was used (Neale et al. 2003). Each 6-by-6 covariance matrix provides \((6 \times 7)/2 = 21\) statistics (variances or covariances), so there is a total of 126 statistics over the six groups of sibling pairs. The full model contains 21 path coefficients to be estimated (six for each Cholesky factorization plus three for the phenotypic model), so there are 105 df remaining to test the fit of the model.

Results

The model is estimated from covariance matrices rather than correlations (Neale and Maes, forthcoming). However correlations for the six types of siblings are shown in Table 1, to provide a sense of the information on which estimation of the variance components is based. Correlations across siblings for the same variables are shown in bold type. Cross-siblings correlations are about twice as large for MZ twins (.724, .660 and .663 for VIQ, GPA and CPL, respectively) than for DZ twins (.356, .332 and .264). A classic estimator of heritability is twice the difference between the correlations of MZ and DZ twins. Thus one can estimate heritability as 2(.724 x .356) = .736 for VIQ, .656 for GPA and .798 for CPL. These estimates are similar to those found in other studies of cognition-related outcomes for adolescents and young adults (e.g., Plomin and Petrill 1997). At the other extreme of relatedness, correlations for non-related siblings living in the same household are estimates of the pure impact of the shared environment of siblings on the outcomes; their small sizes (.063, .080 and .190) constitute a preliminary hint that the shared environment is not a substantial determinant of these variables in this population, except perhaps for CPL.

The full model is denoted BACE, as it specifies direct paths relating observed variables (contained in matrix B), in addition to a full Cholesky structure associated with each latent component A, C and E. Fit statistics for the BACE model are shown on the first line of Table 2. Unfortunately the B matrix in the BACE model is not identified; the reason is that the ACE part of the model completely accounts for the observed variables and their correlations, so B cannot improve the fit. B can be estimated in a simplified model with no genetic component, a single shared environment factor C, affecting the observed variables and three uncorrelated latent variables E₁, E₂ and E₃ representing the unshared environment of VIQ, GPA and CPL, respectively. (Cᵢ is then equivalent to a pair-specific fixed effect affecting each observed variable.) The model is labeled BC₁E₃ to indicate the shape of the matrices involved, with C reduced to a single column and E to a diagonal matrix.
Table 1: Correlations for Six Groups of Siblings (N = number of pairs)

<table>
<thead>
<tr>
<th></th>
<th>MZ Twins (below diagonal, N = 170)</th>
<th>DZ Twins (above diagonal, N = 290)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>VIQ, GPA, CPL, VIQ, GPA, CPL</td>
<td></td>
</tr>
<tr>
<td>VIQ,</td>
<td>.239</td>
<td>.356</td>
</tr>
<tr>
<td>GPA,</td>
<td>.277</td>
<td>.332</td>
</tr>
<tr>
<td>CPL,</td>
<td>.290</td>
<td>.264</td>
</tr>
<tr>
<td>VIQ,</td>
<td>.724</td>
<td></td>
</tr>
<tr>
<td>GPA,</td>
<td>.182</td>
<td></td>
</tr>
<tr>
<td>CPL,</td>
<td>.325</td>
<td></td>
</tr>
<tr>
<td>Full Siblings (below diagonal, N = 702)</td>
<td>Half Siblings (above diagonal, N = 242)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VIQ, GPA, CPL, VIQ, GPA, CPL</td>
<td></td>
</tr>
<tr>
<td>VIQ,</td>
<td>.183</td>
<td>.310</td>
</tr>
<tr>
<td>GPA,</td>
<td>.295</td>
<td>-.103</td>
</tr>
<tr>
<td>CPL,</td>
<td>.212</td>
<td>.102</td>
</tr>
<tr>
<td>VIQ,</td>
<td>.411</td>
<td></td>
</tr>
<tr>
<td>GPA,</td>
<td>.163</td>
<td></td>
</tr>
<tr>
<td>CPL,</td>
<td>.133</td>
<td></td>
</tr>
<tr>
<td>Cousins (below diagonal, N = 105)</td>
<td>Non-related Siblings (above diagonal, N = 174)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VIQ, GPA, CPL, VIQ, GPA, CPL</td>
<td></td>
</tr>
<tr>
<td>VIQ,</td>
<td>.292</td>
<td>.063</td>
</tr>
<tr>
<td>GPA,</td>
<td>.171</td>
<td>-.066</td>
</tr>
<tr>
<td>CPL,</td>
<td>.061</td>
<td>.099</td>
</tr>
<tr>
<td>VIQ,</td>
<td>.354</td>
<td></td>
</tr>
<tr>
<td>GPA,</td>
<td>.090</td>
<td></td>
</tr>
<tr>
<td>CPL,</td>
<td>.238</td>
<td></td>
</tr>
</tbody>
</table>

Note: Correlations in bold type are for the same variable across siblings: VIQ = verbal IQ, GPA = grade point average, CPL = college plans. Subscripts denote siblings.

The fit statistics for the BC1E2 model are shown on line two of Table 2. The $\chi^2$ is 555.888 for 117 df (p < .001), which is not a satisfactory fit. (With SEMs the goal is to obtain a non-significant model.) To check the significance of B, the model on line three of Table 2 drops B from the model. This results in a $\chi^2$ increase of 89.657 for three degrees of freedom, a highly significant deterioration in fit.
Table 2: Model Comparisons

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
<th>AIC</th>
<th>RMSEA</th>
<th>Test</th>
<th>$\Delta \chi^2$</th>
<th>$\Delta df$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. BACE</td>
<td>153.930</td>
<td>105</td>
<td>.001</td>
<td>-56.070</td>
<td>.042</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. BC, $E_d$</td>
<td>555.888</td>
<td>117</td>
<td>.000</td>
<td>321.888</td>
<td>.112</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. $\phi, E_d$</td>
<td>645.546</td>
<td>120</td>
<td>.000</td>
<td>405.546</td>
<td>.126</td>
<td>3 vs. 2</td>
<td>89.657</td>
<td>3</td>
<td>.000</td>
</tr>
<tr>
<td>4. ACE</td>
<td>153.930</td>
<td>108</td>
<td>.002</td>
<td>-62.070</td>
<td>.041</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. $\phi E$</td>
<td>179.069</td>
<td>114</td>
<td>.000</td>
<td>-48.931</td>
<td>.046</td>
<td>5 vs. 4</td>
<td>25.139</td>
<td>6</td>
<td>.000</td>
</tr>
<tr>
<td>6. $\phi CE$</td>
<td>296.203</td>
<td>114</td>
<td>.000</td>
<td>68.203</td>
<td>.079</td>
<td>6 vs. 4</td>
<td>142.273</td>
<td>6</td>
<td>.000</td>
</tr>
<tr>
<td>7. $\phi CE$</td>
<td>221.945</td>
<td>111</td>
<td>.000</td>
<td>-33.861</td>
<td>.054</td>
<td>7 vs. 4</td>
<td>34.209</td>
<td>3</td>
<td>.000</td>
</tr>
<tr>
<td>8. $\phi CE$</td>
<td>188.139</td>
<td>111</td>
<td>.000</td>
<td>-54.653</td>
<td>.043</td>
<td>9 vs. 4</td>
<td>13.507</td>
<td>3</td>
<td>.004</td>
</tr>
<tr>
<td>9. $\phi CE$</td>
<td>167.437</td>
<td>111</td>
<td>.002</td>
<td>-63.193</td>
<td>.054</td>
<td>10 vs. 4</td>
<td>4.877</td>
<td>3</td>
<td>.181</td>
</tr>
<tr>
<td>10. $\phi CE$</td>
<td>158.807</td>
<td>111</td>
<td>.000</td>
<td>-62.565</td>
<td>.044</td>
<td>11 vs. 4</td>
<td>11.505</td>
<td>6</td>
<td>.074</td>
</tr>
</tbody>
</table>

Note: B = phenotypic paths; A = genetic paths; C = shared environment paths; E = specific environment paths; $A_{\phi}, C_{\phi}, E_d$: off diagonal elements of A, C, or E fixed (independent factors model); $A_{\phi}, C_{\phi}$: lower triangular matrix A, C reduced to single column vector (common factor model). * favored model

Thus the path coefficients relating the three observed variables of the model are significant in the context of this simplified specification of the latent structure with no genetic influences. One reason for the poor overall fit of Models 2 and 3 is that the absence of genetic component implies identical covariance matrices for all six groups of sibling pairs, a pattern that is clearly inconsistent with the data (Table 1).

For the full ACE model, $\chi^2$ is 153.930 with 105 df (p = .001), which at first sight is not a satisfactory fit. RMSEA is a measure of fit that adjusts for degrees of freedom and sample size (McDonald 1989). RMSEA is .042, below the threshold of .05 corresponding to a very good fit. Fixing to zero the shared environment structure C (model AE) causes a highly significant increase in $\chi^2$ of 25.139 for 6 df (p < .001); therefore the shared environment structure cannot be dropped from the model. Fixing the genetic structure A (model CE on line 6) produces a highly significant increases in $\chi^2$ of 142.273 with 6 df (p < .001); thus the genetic structure cannot be excluded either.

Models 7 to 11 test alternative specifications of the internal structure of the A, C and E matrices, specifically whether the lower triangular matrix can be replaced by a simpler model consisting of a diagonal matrix (three uncorrelated factors, each affecting a single outcome) or by a single column (a single factor affecting all three outcomes). Reducing A to a single column (model $A_{\phi} CE$, the
common factor model) or to a diagonal matrix (model AdCE, the independent factors model) leads to unacceptably large $\chi^2$ increases (68.014 and 34.209 with 3 df, respectively). Factors in the A matrix can be thought of as sets of genes. Thus the tests on lines seven and eight show that the genetic structure can neither be reduced to a single set of genes affecting all three outcomes (A,CE) nor to three distinct sets of genes, each affecting a single factor (A$_0$CE). Likewise reducing the shared environment structure C to a diagonal matrix representing three independent factors (model AC$_d$E) results in a significant increase in $\chi^2$ (13.507 for 3 df, $p = .004$). However reducing C to a single column representing a single shared environment factor affecting all three variables (model AC$_1$E) produces a non-significant increase in $\chi^2$ (4.877 for 3 df, $p = .181$). Finally, combining a common factor structure for the shared environment with a diagonal structure for the unshared environment (model AC$_1$Ed) produces a slightly refined model that does not fit significantly worse than ACE ($\chi^2$ increase 11.505 for 6 df, $p = .074$) or AC$_1$E ($\chi^2$ increase 6.628 for 3 df, $p = .085$), and therefore becomes the favored model. AC$_1$Ed is favored over AC$_1$E for its simplicity, despite a slightly larger AIC value: AIC is -62.565 for AC$_1$Ed vs. -63.193 for AC$_1$E.\(^8\) Estimated parameters for the two models hardly differ.

Standardized path coefficients for the favored AC$_1$Ed model are shown in Figure 2. The model represents shared environmental influences acting on VIQ, GPA and CPL as a single latent factor affecting all three outcomes, which one might perhaps identify with a “privilege” factor capturing the cognitive and academic advantage shared by siblings due to their rearing environment. The unshared environment is represented by separate factors, each one affecting a single measure. It is the behavior one would expect if the unshared environment consisted largely of measurement error. This is somewhat surprising since one would have expected that some unshared influences (e.g., perinatal damage affecting one sibling but not the other) would affect all three variables in similar ways. Effects of the genetic factors tend to be the largest (.202 to .738), and effects of the shared environment the smallest (-.041 to .371), with effects of the unshared environment in between (.572 to .609).

Table 3 shows the proportions of the total expected variances of the observed variables that are explained by the latent factors. Heritability (conventionally denoted $h^2$) is the proportion of variance explained by genetic factors. Estimated heritability is high for GPA (.669) and for CPL (.600), and somewhat lower for VIQ (.536). Environmentality (denoted $c^2$) is the proportion of variance explained by shared environmental factors. Environmentality is substantial for VIQ (.137) but almost nil for GPA and CPL (.002 and .03, respectively). Specificity (denoted $e^2$), or proportion of variance due to unshared influences, is substantial for all three measures (.327, .329 and .370 respectively). Recall that the unshared environment here includes residual variance.\(^9\) Table 3 also shows tests of significance for these parameters using ML-based confidence intervals.\(^10\) All estimates are significant in the sense that the ML 95 percent confidence interval does not include zero except environmentality for GPA. Thus one cannot reject the hypothesis that GPA is entirely explained by a combination of genetic and unshared environmental factors, with no significant role for the shared family environment. Environmentality for CPL also comes close to non-significance.
Discussion: Heritability, Environmentality and Comparative Startification Research

Results presented earlier indicate that the three schooling measures are highly heritable, relatively unaffected by the shared environment, and substantially affected by unshared environments. The shared environment seems to affect all three measures as a single latent "privilege" factor, whereas genetic influences are better represented as partially independent sets of genes specific to each outcome. Such findings are far from isolated. The view that cognitive ability and educational success have a substantial genetic basis and are (beyond childhood) little affected by shared environments has become commonplace in mainstream psychology (Brody 1992, Gottfredson [1994] 1997, Neisser et al. 1996, Sternberg and Grigorenko 1997, especially Hunt 1997). As McGue (1997) notes, the scientific opposition to these conclusions now consists of studies using behavior-genetic models tuned to produce lower heritability estimates (e.g., Daniels, Devlin and Roeder 1997; Feldman, Otto and Christiansen 2000). There is also evidence for a role of genes in the determination of earnings (Behrman, Pollack and Taubman 1995; Björklund, Jäntti and Solon 2005), personality traits and some social attitudes (Rowe 1994). These findings have moved Turkheimer (2000) to proclaim...
Table 3: Maximum-likelihood Confidence Intervals for Heritabilities, Environmentalities and Specificities for Favored AC,E, Model

<table>
<thead>
<tr>
<th></th>
<th>95% CI</th>
<th>95% CI</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>h^2</td>
<td>Low</td>
<td>Up</td>
</tr>
<tr>
<td>VIQ</td>
<td>.536</td>
<td>.408</td>
<td>.649</td>
</tr>
<tr>
<td>GPA</td>
<td>.669</td>
<td>.585</td>
<td>.725</td>
</tr>
<tr>
<td>CPL</td>
<td>.600</td>
<td>.493</td>
<td>.677</td>
</tr>
</tbody>
</table>

Note: Heritabilities (h^2), Environmentalities (c^2) and Specificities (e^2).

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

three “laws” of behavior genetics, namely that (1) all human behavioral traits are heritable, (2) the effect of being raised in the same family is smaller than the effect of the genes, and (3) a substantial portion of the variation in complex human behavioral traits is not accounted for by the effects of genes or families. Pinker (2002:372-99) reckons that the three laws may be “the most important discoveries in the history of psychology (p. 372).”

The most controversial implication of findings of a major role of genetic endowment in educational and socio-economic success is a concern that this knowledge could be misinterpreted to make inequality of school outcomes seem “natural, just, and immutable” and to justify opposition to efforts at improving school performance. For this reason Goldberger (1979) suggests abandoning the enterprise of estimating genetic variance components of educational and socio-economic achievement, but this advice is rejected by Taubman (1995a), who points out that behavior genetic models are valuable in providing a measure of opportunity for achievement. As Behrman and Taubman (1995) write: “The share of the observed variation in schooling that is attributable to across-family variability in environment [environmentality] provides a measure of inequality of schooling opportunity.” (p. 250)

Heritability and environmentality estimates refer to a specific population, characterized by an existing range of environments. Thus these estimates give no guidance on how difficult it would be to change the trait through environmental manipulation outside the existing range (Jencks 1980, 1992:92-119). Thus high heritability does not imply that all environmental policies are ineffective (Eckland 1979). Likewise, it is not because socio-economic success has a genetic basis that inequality in socio-economic outcomes is desirable; to conclude otherwise is to commit the naturalistic fallacy (Pinker 2002:162-63). Time will tell how discoveries about the genetic basis of behavior will be received by sociologists, but it seems unlikely that the accumulated evidence for a major role of genes in socio-economic outcomes will be easily overthrown or ignored.

The fact that heritability and environmentality are not fixed properties of a trait but population statistics inherently dependent on the specific social context has intriguing implications. For comparative social mobility research the parameters of the behavior-genetic model may be viewed as macro-social variables describing the nature of social mobility in a system of stratification. Heritability indexes
opportunity for achievement (realization of native potential) and, conversely, environmentality measures inequality of opportunity due to differences in rearing environments (ascertainment or social reproduction). The analysis reported in this paper – describing one society at a particular point in history – represents a single case from a macro-sociological comparative perspective or for purposes of policy-oriented evaluations. The value of behavior genetic models as tools of comparative social mobility research lies in their providing comparable heritability and environmentality estimates for educational or socio-economic outcomes in different social systems or for subgroups within a given social system, allowing comparisons among these macro-social units on a scale of ascription versus opportunity. Within a given social context further comparisons can be made of the ascriptive vs. opportunistic nature of the achievement process for different dimensions or stages of socio-economic attainment (e.g. cognitive measures, education, occupation, income). While such a systematic comparative corpus does not exist at present, one can already combine estimates from a number of published studies to construct a “demo” of what genetically informed comparative social mobility research might look like.

A general pattern emerging from the literature is that realization of genetic potential (and thus heritability) is lower – and environmentality correspondingly higher – in social contexts that are either disadvantaged or less advanced on a social-developmental dimension. The pattern, originally proposed by Scarr-Salapatek 1971, was clearly in evidence in Heath et al.’s (1985) study in Norway which found that heritability of educational attainment is relatively low (41 percent) for both males and females born before 1940, who would have been exposed to the more elitist traditional educational system in that country. For two cohorts of males born later heritability rises to 74 percent and 67 percent, reflecting greater openness of the system; for females, however, heritability remains lower for these two cohorts (45 percent and 38 percent) suggesting that achievement opportunity has increased for males but not for females. A similar scenario emerges from the study of IQ, educational attainment, and occupation in Norway by Tambs et al. (1989). Here heritability is higher for the younger than for the older group, reflecting again less social ascription affecting the younger group. The same inter-generational pattern obtains in the study of educational attainment in Australia by Baker et al. (1996), but the pattern is not as marked in the Norwegian data analyzed by Lichtenstein et al. (1992). Rowe, Vesterdal and Rodgers (1999), using U.S. data on young adults, find heritability to be higher for IQ and educational attainment (64 percent and 68 percent, respectively) than for hourly wages (42 percent) (see also Rowe, Jacobson and Van den Oord 1999). Guo and Stearns (2002) show that heritability of verbal IQ for adolescents in AddHealth is lower when a parent is unemployed than when no parent is unemployed (42 percent vs. 54 percent), and for black as compared to white adolescents (58 percent vs. 72 percent); lower heritability is thus again associated with fewer opportunities in disadvantaged social environments. Finally Turkheimer et al. (2003) find higher $h^2$ and lower $c^2$ for IQ in high-SES environments, compared to low-SES ones.

However cursory, this partial survey of the literature is suggestive of some of the contextual variables on which genetically-informed comparative stratification
research might focus: historical period and country – as these factors proxy for the nature of the mobility regime; relative socio-economic advantage, race, ethnicity and sex – as these characteristics define groups facing more or less opportunity; and age – both as it defines cohorts subjected to different mobility regimes and as it may otherwise affect model parameters.\textsuperscript{13}

**Conclusion**

The status attainment model has been used to evaluate strength of ascription vs. achievement opportunity in systems of stratification, but this use of the model is problematic. Estimating ascription from the size of family background effects and opportunity from the size of cognitive ability and education effects is misleading because attribution of the effects of variables to ascription or opportunity is arbitrary, because the model is vulnerable to misspecification of family background, and because effects of background and achievement variables are confounded with genetic influences. Behavior genetic models estimated from twins and other family data can disentangle ascription from achievement by controlling for genetic influences and by providing a black box estimate of shared environmental influences that does not require explicit measurement of the variables involved. The impact of the shared environment (environmentality) may be interpreted as a measure of ascription, and heritability as a measure of opportunity for achievement. A multivariate behavior genetic model of verbal IQ, grade point average and college plans estimated from data on adolescent sibling pairs who were in grades 7 through 12 of U.S. schools in 1994-95 shows that variation in all three measures of educational attainment has a large genetic component and a relatively smaller shared environmental component. These findings suggest high levels of educational opportunity for adolescents in U.S. schools at the end of the 20th century.

A cursory overview of published estimates of the parameters of behavior genetic models of attainment-related variables illustrates the possibilities of a comparative sociology of stratification systems using heritability and environmentality parameters as fundamental measures of ascription and opportunity for achievement in a system of stratification. In this view heritability and environmentality are not fixed properties of a trait; they are properties of the stratification system that are expected to vary across societies, historical periods and social contexts. Parameters of the genetic model are estimated, not as permanent properties of a trait such as intelligence, but as a descriptive feature of the social context that may tell us something valuable, say, about the way men and women fared differentially in the educational system in Norway after World War II, or how members of disadvantaged social or racial groups are less able to realize their native potential in the United States at the end of the 20th century.

Models estimated in this paper are based on the assumption of polygenic inheritance, i.e., that the phenotype under study is affected by a number of genes, each of which has a relatively small effect on the phenotype (Fisher 1918). Another rapidly developing category of behavior genetic models uses genetic markers to locate on the genome genes with large effects. These methods have produced major findings, such as genes associated with major mental disorders.
Success of the gene-finding methods depends on the existence of genes with large effects on a trait; only then can the effect of a specific gene be detected against the background of other genetic and environmental influences. The possibility of identifying genes with large effects is methodologically attractive because part of the genetic contribution is then associated with an explicitly measured variable (i.e., presence or absence of a given allele or marker); effects of other genes are treated as a residual, latent genetic factor. There is a great deal of finality in the identification of a gene with a major effect on a trait. When the trait in question is a physical or mental disorder, identification of the responsible locus may facilitate development of a treatment.\footnote{1}

Traits of interest in social mobility research, such as cognitive ability, dimensions of personality and measures of socio-economic achievement may – or may not – be affected by genes of major effect that stand out against the polygenic noise of other genes and the additional din of environmental effects. When found, rare genes of large effect may not explain a large proportion of the genetic variance of a trait in the population. Gene-finding methodologies may one day identify genes responsible for educational and socio-economic achievement, thus resolving the latent genetic component into observed DNA measures. However to the extent that genetic influences on these outcomes are truly polygenic, structural equations models based on the assumption of polygenic inheritance like those estimated in this paper may prove useful for a long time to come (Neale and Maes forthcoming).

Notes

1. There is a widespread but incorrect perception among sociologists that Jencks’ work had effectively ruled out any significant role of genes in the attainment process. In contrast Jencks (1992:13) recalls concluding in earlier work that genes exert “a considerable influence” on cognitive test scores.

2. Many behavior genetic models also include a \textit{genetic dominance} component capturing non-linear effects of allele combinations when a trait is affected by genes with dominant vs. recessive alleles.

3. Alternatively, genes can be viewed as an \textit{instrument} (in the statistical sense) for native potential.

4. The model described later does not allow disaggregating the shared environment into measured background characteristics, although other behavior genetic models do (Behrman, Pollack and Taubman 1995; Loehlin, Horn and Willerman 1997; Taubman 1995b; Waldman 1997). In any case, the shared environment component represents an upper bound for the effect of any specific dimension of the rearing environment on the outcome.

5. There is an alternative regression-based behavior genetic methodology appropriate for univariate analysis called DF analysis (DeFries and Fulker 1985, Kohler and Rodgers 2001).
For CPL the DZ correlation (.264) is less than half the correlation for MZ twins (.663/2 = .332), a pattern suggesting a genetic dominance effect. However further data analysis does not support that hunch (Nielsen 2006).

BACE produces the same estimates and the same fit as the ACE model (see line four of Table 2) for the behavior genetic part, but the estimated coefficients of B shift with changes in the initial values of the coefficients, indicating underidentification (Neale and Maes forthcoming).

Akaike’s Information Criterion (AIC) is a measure of model fit adjusted for degrees of freedom such that a smaller (more negative) value indicates a better fit.

One can also estimate the proportion of the covariance between two measures explained by a set of factors. Associations among these three cognitive and educational measures seem largely due to genetic factors (Nielsen 2006). It is also possible to estimate the correlations among the genetic components of the variances in the three outcome variables (Neale and Maes forthcoming). These calculations suggest that the genes responsible for the heritability of the three measures constitute partially non-overlapping sets, a pattern inconsistent with the view that school success reflects a single “academic ability” factor. Rather, different measures of school success might be associated with different sets of psychological traits, each with a partially independent genetic etiology.

ML-based confidence intervals are considered superior to those based on estimated standard errors (Neale et al. 2003, Neale and Miller 1997, Turkheimer et al. 2003).

Nielsen (2006) contains supplementary analyses and discussions on the following issues: (1) similarity of shared environments of DZ twins compared to full siblings and MZ twins compared to DZ twins, (2) alternative model estimates using MZ and DZ twins only, (3) effects of allowing assortative mating on parameter estimates, and (4) testing genetic dominance in the model for CPL.

Although it stands to reason that environmental interventions outside the normal environmental range are more likely to be expensive, impractical or unethical.

A major discovery of behavior genetics in the past two decades is that heritability of cognitive ability increases, and environmentality declines, from childhood to adulthood (Plomin and Petrill 1997).

The same logic would apply to identification of a specific property of the environment (shared or unshared) with a large effect on the trait.

References

Achievement & Ascription in Education


