



OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

Variation in the Heritability of Educational Attainment: An International Meta-Analysis

Author(s): Amelia R. Branigan, Kenneth J. McCallum and Jeremy Freese

Source: *Social Forces*, September 2013, Vol. 92, No. 1 (September 2013), pp. 109-140

Published by: Oxford University Press

Stable URL: <https://www.jstor.org/stable/43287519>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



Oxford University Press is collaborating with JSTOR to digitize, preserve and extend access to *Social Forces*

JSTOR

Heritability of Educational Attainment

Variation in the Heritability of Educational Attainment: An International Meta-Analysis

Amelia R. Branigan, Kenneth J. McCallum, and Jeremy Freese, *Northwestern University*

To assess heterogeneity in the influence of genetic variation on educational attainment across environmental contexts, we present a meta-analysis of heritability estimates in fifteen samples and thirty-four subgroups differing by nationality, sex, and birth cohort. We find that heritability, shared environment, and unshared environment each explain a substantial percentage of the variance in attainment across all countries, with between-sample heterogeneity in all three variance components. Although we observe only meager differences in the total family effect by cohort or sex, we observe large cohort and sex differences in the composition of the family effect, consistent with a history of higher heritability of educational attainment for males and for individuals born in the latter half of the twentieth century. Heritability also varies significantly by nation, with the direction of variation specific by sample.

The authors would like to thank three anonymous reviewers; Larry V. Hedges; Andrew C. Heath; Eric Turkheimer; Jason Schnittker; Sonia Brescianini for the Italian Twin Registry; Juan Ordoñana for the Registro de Gemelos de Murcia (Murcia Twin Registry); Rainer Riemann for the Bielefeld Twin Study; and Gary Swan and Ruth Krasnow for the Twin Research Registry at SRI International. The MIDUS I study (Midlife in the US) was supported by the John D. and Catherine T. MacArthur Foundation Research Network on Successful Midlife Development. The MIDUS II research was supported by a grant from the National Institute on Aging (P01-AG020166) to conduct a longitudinal follow-up of the MIDUS I investigation. The TwinsUK study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE project grant agreement (HEALTH-F4-2007-201413). The study also receives support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St. Thomas's NHS Foundation Trust in partnership with King's College London. TDS is an NIHR senior investigator and is holder of an ERC Advanced Principal Investigator award. Genotyping was performed by the Wellcome Trust Sanger Institute, support of the National Eye Institute via an NIH/CIDR genotyping project. This research also uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from twenty-three other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (<http://www.cpc.unc.edu/addhealth>). No direct support was received from grant P01-HD31921 for this analysis. The research reported here was supported by the Institute of Education Sciences, US Department of Education, through grant R305B080027 to Northwestern University. The opinions expressed are those of the authors and do not represent views of the institute or the US Department of Education.

We find a markedly larger impact of shared environment on attainment than has been found for other social outcomes, with the percent of variation in attainment attributable to shared environment exceeding the percent attributable to heritability in one-third of the studies in our sample. Our findings demonstrate the heritability of educational attainment to be environmentally contingent, affirm the widespread and enduring role of shared environment in determining ultimate socioeconomic attainment, and emphasize the importance of considering behavioral genetics techniques in models of social mobility.

Although it is now taken for granted that many physiological outcomes are influenced by genetic differences, a number of studies of heritability over the past three decades have considered behaviors and attitudes once thought to be primarily social, including income and likelihood of divorce (Behrman, Taubman, and Wales 1977; McGue and Lykken 1992). In sociology, research has focused mainly on identifying social variables that alter the effect of genetics on social outcomes, contradicting notions of a “nature-versus-nurture” dichotomy in which genes and environment represent distinct causal explanations (e.g., Boardman et al. 2008; Guo and Stearns 2002). In this work, authors have suggested that estimates of the percentage of variation in social outcomes explained by genetic and environmental differences are likely to be context specific, varying systematically across different social conditions, policy environments, or subgroups of the population (Jencks 1980; Boardman 2009). For scholars of social inequality, these findings are particularly meaningful, as they suggest that models of variation in social outcomes that take into account only social predictors are missing an integral part of the story (Eckland 1967; Freese 2008).

Education was an early domain of interest among researchers assessing the heritability of social outcomes, primarily using test-score data to contribute to long-standing debates on the genetic basis of intelligence or achievement (e.g., Bouchard and McGue 1981; Devlin, Daniels, and Roeder 1997; Guo and Stearns 2002; Nielsen 2006).¹ Studies of the heritability of years of education attained, on the other hand, are relatively few (Heath, Berg, et al. 1985; Behrman and Taubman 1989; Silventoinen et al. 2004), despite attainment being a common operationalization of educational performance across many fields and a long-standing outcome of interest in research on social mobility. While test scores are a socially contingent outcome, attainment is arguably even more so, under the presumption that the number of years of education one completes reflects not just cumulative achievement, but a vast range of social factors likely to influence school experiences and continuation decisions through pathways other than academic performance. In this study, we address this omission with a meta-analysis of the largest set of heritability estimates of educational attainment yet assembled, using published and unpublished data ranging across ten countries and with birth years spanning more than a century.

Prior studies of the heritability of educational outcomes using micro-level data have identified a number of variables by which the estimated influence of genetic variation on performance varies, most commonly indicators of socioeconomic status (Guo and Stearns 2002; Rowe, Jacobson, and van den Oord 1999) or

egalitarian policy initiatives (Heath, Berg, et al. 1985). With few exceptions, these studies have been based on comparisons between subgroups in a single sample drawn from within a single country. Although still informative, this approach potentially complicates generalization beyond the environmental contexts of the original studies, as the social settings from which the individual samples were drawn will themselves encompass a vast array of relevant factors such as governmental structure, religious or racial heterogeneity, and levels of inequality and mobility on many social dimensions. A meta-analysis trades explanatory depth for historic or geographic breadth, allowing us to explore variation in average heritability of attainment across a broad range of samples as opposed to between subgroups in a single sample. While micro-level data are useful for capturing within-population variation in heritability by sample-specific characteristics such as individuals' relative poverty level, here we are uniquely able to evaluate the effect of larger-scale social categories such as nationality, and to assess the significance of sex and birth cohort across a vastly larger and more diverse respondent population than would be otherwise attainable.

Beyond identifying new social dimensions by which the heritability of attainment varies, the ability to compare estimates of heritability across sex, nation, and birth cohort makes this analysis particularly relevant in a social mobility context (e.g., Björklund, Jäntti, and Solon 2003; Diewald 2010). In the total family effect frequently calculated in studies of social mobility, the influence of genetic variation and family-level social influence are captured as a combined estimate (Hauser and Mossel 1985; Hauser and Wong 1989; Sieben and de Graaf 2003). Although identifying differences in the family effect on attainment across varying social groups (Kuo and Hauser 1995, 1997) and birth cohorts (Dronkers 1993a, 1993b; de Graaf and Huinink 1992; Smith and Cheung 1986; Hauser and Featherman 1974) has long been a topic of interest in the social-mobility literature, few attempts have been made to identify group differences in the effect of family environment net of the effect of genetic relatedness. By comparing separate estimates of the impact of heritability and familial environment as calculated in our sample, we are able not only to observe nation, sex, and cohort differences in the total family effect on educational attainment, but also to detect differences in how that effect is decomposed into genetic and environmental components. Our findings emphasize the importance of calculating genetic effects separately from social family effects when estimating models of social mobility, as changes in the importance of home environment relative to genes appear to reflect group differences in mobility trends even while the sum total family effect remains stable.

Genetic and Environmental Influences on Educational Performance

The most conventional approach to assessing the percentage of variation in an outcome accounted for by the effects of genetics is the twin study, in which the difference in within-pair correlations between identical (monozygotic, MZ) and fraternal (dizygotic, DZ) twins are used to generate an estimate of heritability

(Winerman 2004). This approach is predicated on the assumption that higher correlations among MZ versus DZ twins can be attributed to MZ twins being genetically (nearly) identical, whereas DZ twins share only approximately half their genes by descent. While heritability can also now be assessed using molecular genome-wide data, the only published study to do so for educational attainment with data that also include twins found similar estimates using conventional twin and genome-wide estimation techniques (Martin et al. 2011).

The term *heritability* refers to the percentage of outcome variation accounted for by additive genetic variation,² while remaining variation in the outcome is attributed to two categories of environmental influence: the *shared environment*, those factors experienced similarly by twin pairs, and the *unshared environment*, those factors that individual twins experience uniquely. In the case of educational outcomes, heritability estimates have been hypothesized to be lower for individuals in environments that may work to stunt achievement, such as being raised in poverty, and higher for individuals not exposed to such constraints (Rowe, Jacobson, and van den Oord 1999; Guo and Stearns 2002; Turkheimer et al. 2003). In principle, at least, environmental constraint could also depress heritability by compelling individuals to attain at higher levels than they otherwise would (see Boardman, Daw, and Freese [forthcoming] on “social control” and heritability-by-environment interaction). One possible example of such a constraint would be the legal requirement to remain in school until a minimum age, whereby individuals otherwise inclined to attain fewer years of education may continue in school due to legal strictures.

The “shared” or “common” environment component reflects the proportion of variance in educational outcomes explained by all non-genetic sources of twin similarity in a phenotype, including characteristics that twins experience similarly due to having been reared together. As we illustrate in table 1, across a wide range of social outcomes the percent of variation explained by this component has been consistently small, and nearly always smaller than the percent of variation explained by genes. The finding is so ubiquitous that Turkheimer (2004) coined it the “second law of behavioral genetics”: that “the environmental effect of being raised in the same family is substantially smaller than the genetic effect and is often close to zero” (161; see also Turkheimer [2000]). This assumption poses a challenge for sociologists, as the effects of many social forces are ostensibly captured in the common environment estimate: characteristics of the family such as parental education, occupation, wealth, income, parenting style, and sibship size; within-population variation in race and ethnicity; and characteristics of the neighborhood such as poverty level, access to schools and other institutional resources, crime, and so on.³

Variation not explained by genetic variation or by family environment is attributed to the effect of “unshared” environment, which refers to individual-specific factors that influence twins to differ on an outcome of interest. With estimates of shared environment so consistently low in prior studies (table 1), unshared environment is typically expected to explain the vast majority of outcome variation not accounted for by genetic differences (Turkheimer 2000, 2004). An illness or accident that affects only one twin would be captured here, as would differences

Table 1. Selected Studies of the Heritability of Social Outcomes, by Percent Heritability and Percent Common Environment

	High Common Environment (>33%)	Low Common Environment (≤33%)
High Heritability (>33%)	<ul style="list-style-type: none"> - Smoking initiation in men (Li et al. 2003) 	<ul style="list-style-type: none"> - Alcoholism (Heath, Bucholz, et al. 1997) - Altruism and aggression (Rushton et al. 1986) - Age at first intercourse for respondents born ≥1952 (Dunne et al. 1997) - Attitudes on school prayer, property tax, moral majority, capitalism, astrology, draft laws, pacifism, unions, Republicans, socialism, foreign aid, X-rated movies, immigration, women's liberation (Alford, Funk, and Hibbing 2005) - Attitudes on the death penalty, open-door immigration, doing athletic activities, voluntary euthanasia, exercising, organized religion, reading books, roller coasters (Olson et al. 2001) - College plans (Nielsen 2006) - Depression in women (Kendler et al. 2006) - Electoral participation (Fowler, Baker, and Dawes 2008)¹⁴ - Extraversion (Pedersen et al. 1988) - Impulsivity (Pedersen et al. 1988) - Income (Taubman 1976) - Inhibition (Robinson et al. 1992) - IQ (e.g., Rowe, Jacobson, and van den Oord 1999)¹⁵ - Loneliness (Boomsma et al. 2005) - Low birth weight (Clausson, Lichtenstein, and Cnattingius 2000) - Perceived self-confidence (McGuire et al. 1994) - Perfectionism (Tozzi et al. 2004) - Sexual orientation in women (Bailey et al. 1993) - Smoking initiation in women (Li et al. 2003) - Smoking persistence (Li et al. 2003) - Social cognitive skills (Scourfield et al. 1999)
Low Heritability (≤33%)	<ul style="list-style-type: none"> - Age at first intercourse for men born <1952 (Dunne et al. 1997) - Attitudes on abortion and on living together (Alford, Funk, and Hibbing 2005) - Preference for loud music (Olson et al. 2001) 	<ul style="list-style-type: none"> - Attitudes on the death penalty, censorship, military drill, gay rights, segregation, busing, nuclear power, Democrats, divorce, modern art, federal housing, liberals (Alford, Funk, and Hibbing 2005) - Preference for sweets, organized sports, looking one's best at all times, legalizing racial discrimination, public speaking, wearing clothes that draw attention, being the leader of groups (Olson et al. 2001) - Depression in men (Kendler et al. 2006) - Neuroticism (Pedersen et al. 1988) - Stress coping styles (Kato and Pedersen 2005)

in birth weight, assigned teachers, or peers. Although the sources of variance captured in the estimate of unshared environment are expected to be individual specific and random, population differences may reflect broad social trends that lead family members in a given society to vary more or less from one another on average. In a society where the philosophy regarding individuation of same-aged siblings leads to twins being deliberately placed in different classrooms, for example, one might expect a higher estimated influence of unshared environment than in societies in which twins are deliberately kept together in school.

In interpreting the shared and unshared environment components, an important distinction is made between the *objective* environment, meaning the environment as measured by an observer, and the *effective* environment, meaning the environment as defined by its actual influence upon individuals (Turkheimer and Waldron 2000). An objective environmental condition experienced similarly by twins may still affect the two individuals differently, while objectively dissimilar environmental conditions may nonetheless produce identical effects. Turkheimer and Waldron (2000) pose the example of divorce: although twins reared together would both be objectively exposed to the divorce of their parents, it need not affect the two individuals in the same way. Our estimates are thus capturing *effective* shared and unshared environment, as we observe only actual outcomes, and not the objective conditions that shaped those outcomes.

The Present Study

In this study we consider all three variance components, assessing the extent to which the percentages of outcome variation in educational attainment attributable to genetic differences, shared environment, and unshared environment vary across samples differing by sex, birth cohort, and nation. We begin by estimating the degree of between-study heterogeneity in the percent of variance explained by each component among our full set of samples, which we expect to be high given the varying environments from which our samples are drawn. Where we find significant between-study heterogeneity, we introduce our measures of sex, cohort, and nationality in order to assess the percent of variation accounted for by each of these sample characteristics.

Our expectation that sex and birth cohort will affect heritability of educational attainment is based on observed historical attainment trends. Although gender inequality in attainment declined over much of the twentieth century, men consistently achieved higher average rates of attainment in the majority of the nations in our sample, suggesting that men and women may have been subjected to different levels of social constraints with respect to educational continuation.^{4,5} The attainment differential alone does not indicate which gender was subjected to greater constraints—whether men were pushed to progress in school farther than they otherwise might, or females were held back from progressing as far in school. However, limitations on women's occupations through the early twentieth century are well documented, and bounded the return to education for women (Durbin and Kent 1989). Heath, Berg, et al. (1985) found that the heritability of educational attainment was higher in cohorts of Norwegian men

born after 1940, but had remained relatively stable for women, for which one interpretation is that constraints were greater on women than on men. Heath, Berg, et al. (1985) attribute the cohort shift observed in their data to the “more liberal social and educational policies introduced in Norway after the Second World War” (734), a period that also saw rapid change in the educational systems in many of the other countries in our sample.⁶

If this pattern of greater social constraints on achievement for women and for individuals born earlier in the twentieth century can be extrapolated across the countries and years from which our samples were drawn, we expect to see significant differences by sex and birth cohort in the influence of genetic variation on attainment, with the effect being larger for men and for later-cohort respondents. Sex and cohort differences in the percentage of variance in educational attainment explained by genetic versus shared environmental factors would make these estimates a useful tool for assessing social mobility, suggesting that a decrease in the importance of common environment relative to genes may reflect an increase in the opportunity for intergenerational mobility even if the total family effect remains stable.

Net of sex and cohort, we expect that the vast number of differences between nations will also affect the level of social constraints that individuals experience with respect to educational attainment. Systems of government, extent of religious and ethnic diversity, and structure of public education are but a few examples of such differences, all of which are implicitly held constant in a single-nation study of variation in heritability. We anticipate that these differences will influence both magnitude and variation of our heritability estimates, leading to significant heterogeneity across the ten countries in our full sample. The geographic diversity of our samples leaves open the possibility that an observed effect by nation might reflect broader regional factors for which the nation may serve as a proxy. We test for this possibility among a subgroup of only those samples drawn from Scandinavian countries, assessing whether heterogeneity of heritability estimates is significantly reduced within a set of nationally distinct but geographically proximate countries with a similar type of welfare state (Esping-Andersen 1990).

Data and Methods

Despite the few published studies explicitly focused on the heritability of educational attainment (Heath, Berg, et al. 1985; Silventoinen et al. 2004), twin correlations for level of education attained are often available in studies that consider education as relevant for other outcomes, such as schooling (Taubman 1976; Ashenfelter and Kruger 1994; Behrman, Taubman, and Wales 1977; Miller, Mulvey, and Martin 1997; Isacsson 1999; Bingley, Christensen, and Walker 2005). Even when not published, data on educational attainment are also routinely collected by twin studies. We were able to locate fifteen twin samples in which correlations of educational attainment were either published or obtainable from the researchers managing the data. Attainment was assessed as an ordinal scale, with values defined either by degree or by the number of years of

education equivalent to a given degree. The number of categories differed by study, and the relevant degrees differed by nation, reflecting differences in educational systems. The TwinsUK study (Moayyeri et al. 2012) stands as an exception since there was no categorical measure of attainment, but rather a measure of the age at which education was terminated.⁷ Many of the samples are subdivided by sex or by age cohort, amounting to a total of thirty-four subgroups; we define subgroups by sex, birth cohort, and nationality. The data obtained are described in table 2, with sex, sample type, source, and birth years noted.

To conduct our search, we began with electronic databases, using search terms such as “twin study,” “twin registry,” “educational attainment,” “educational level,” and “heritability.” We reviewed the references of the relevant studies found, and pursued any additional studies that might have contained relevant data. We also conducted a similar search in the reverse direction: we sought out lists of twin registries internationally, and reviewed the citation lists on available registry websites to determine whether any studies had been published that included results for education. Finally, we expanded our list to include unpublished estimates from available sources, some of which we computed ourselves, and some of which we obtained through direct correspondence with study investigators.

We considered a sample to be “population based” if respondents were randomly selected from a defined population (e.g., MIDUS; AddHealth), or if all twins from a population census were automatically impaneled (e.g., the Italian Twin Registry; the Minnesota Twin Registry). The remaining samples consisted of individuals from within a particular geographic region who were solicited for survey participation on the basis of twinship (e.g., the SRI and TwinUK samples). We found one study based on a convenience sample of twins attending a twin convention (Ashenfelter and Krueger 1994); while the heritability estimate in this study did not differ substantively from others in our sample, the study was outside our exclusion criteria due to potential bias from the sampling method.

Although relevant twin registries exist in a number of additional non-English-speaking countries (e.g., Korea, Sri Lanka, China, Japan), we were not able to obtain data or estimates from these sources. As a result, our analysis is limited in international scope, spanning only the United States, Western Europe, and Australia. The published studies we found were also entirely from developed nations, although this is perhaps unsurprising, as Sung et al. (2006) find that only 2.7 percent of the papers published using twin-study designs since 1950 have used data from low- or middle-income countries.

For each sample, we coded an indicator variable for sex and birth cohort, with our cohort cutoff set at 1950. Male twins and twins born in the early cohort serve as the reference categories in all models. While nearly all of our samples are single sex, the Swedish sample is mixed sex;⁸ likewise, while twenty-two of our samples are drawn from a relatively narrow range of birth years and can be classified reasonably into either pre- or post-1950 birth cohorts, the remaining samples were drawn from too wide a span of birth years to be readily classified. The variables for sex and cohort range between zero and one, with the values

Table 2. List of Studies

Nationality	Sex	Birth Years	Source	Sample Title	Population-Based
Australia (1)	Male & female	1893–1950	Baker et al. (1996)	Australian Twin Register	No
Australia (2)	Male & female	1951–1965	Baker et al. (1996)	Australian Twin Register	No
Australia (3)	Male & female	1964–1971	Miller, Mulvey, and Martin (1997)	Australian Twin Register	No
Denmark	Male only	1954–2004	Bingley, Christensen, and Walker (2005)	Danish Twins Registry	No
Finland	Male & Female	1936–1955	Silventoinen, Krueger, et al. (2004)	Finnish Twin Cohort Study	Yes
Germany	Male & female	1912–1985	Obtained from investigator	Bielefeld Twin Study	No
Italy	Male & female	1913–1977	Obtained from investigator	Italian Twin Registry	Yes
Norway (1)	Male & female	1915–1939	Heath, Berg, et al. (1985)	Norwegian Twin Panel	No
Norway (2)	Male & female	1940–1949	Heath, Berg, et al. (1985)	Norwegian Twin Panel	No
Norway (3)	Male & female	1950–1960	Heath, Berg, et al. (1985)	Norwegian Twin Panel	No
Spain	Male & female	1940–1965	Obtained from investigator	Murcia Twin Registry	No
Sweden	Mixed	1926–1958	Isacsson (1999)	Swedish Twin Registry	No
United States					
AddHealth	Male & female	1976–1984	Calculated from sample	National Longitudinal Study Of Adolescent Health	Yes
MIDUS	Male & female	1921–1969	Calculated from sample	Midlife in the United States Study	Yes
Vietnam Veterans	Male only	1939–1957	Taubman (1976)	Vietnam-Era Twin (VET) Registry (NAS-NRC)	No
Minnesota	Male & female	1936–1955	Lykken et al. (1990)	Minnesota Twins Registry	Yes

(Continued)

Table 2. continued

Nationality	Sex	Birth Years	Source	Sample Title	Population-Based
WW2 Veterans	Male only	1917–1927	Behrman, Taubman, and Wales (1977)	World War II Veterans Twin Registry (NAS-NRC)	No
SRI	Male & female	1911–1985	Obtained from investigator	Twin Research Registry at SRI International	No
United Kingdom (1)	Mixed ¹⁶	1917–1950	Calculated from sample	TwinsUK	No
United Kingdom (2)	Mixed	1951–1985	Calculated from sample	TwinsUK	No

zero and one denoting homogeneous samples. A mixed sample is coded as the proportion of twin pairs in group one, denoted p . The expectation maximization (EM) algorithm (Dempster, Laird, and Rubin 1977) is used to calculate maximum likelihood estimates when values of p are missing for some samples.⁹ As sex data are missing only for the Swedish sample, we expect that our results for sex should not be strongly affected, although the larger percentage of samples for which birth cohort was missing is important to note when interpreting findings. All computations were done in the statistical software package R (<http://cran.r-project.org/>) using functions written for this analysis.

Assessing Heritability: The Twin Study

While there are many methods of assessing heritability (Hatemi et al. 2010; Neale and Maes 2004), our estimates of heritability are calculated using the simple, conventional equation $2(r_{MZ} - r_{DZ})$, where r_{MZ} , r_{DZ} are the intra-pair correlations of educational attainment for MZ and DZ twins, respectively (Plomin et al. 2008, 382). The simple estimator typically closely approximates estimates derived through more complex means and permits us to maximize cases from our meta-analysis by being calculable simply from twin correlations rather than requiring microdata.

As noted, twin studies partition variance in an outcome into that which is resolved by additive genetic variance (i.e., heritability), by the environment shared by members of a twin pair, and by the unshared environment. The first two of these are assumed to induce similarity between MZ twins, while only the last induces divergence, and so the unshared environment component (e^2) is calculated as

$$e^2 = 1 - r_{MZ}.$$

The within-pair correlations on a given outcome—in our case, educational attainment—

can therefore be expressed in terms of genetic (h^2) and shared environmental (c^2) components.

The correlation for MZ twins reflects full similarity of both genotype and shared environment,

$$r_{MZ} = h^2 + c^2,$$

whereas the correlation for same-sex DZ twins reflects half genetic similarity, but still full similarity of shared environment,

$$r_{DZ} = \frac{1}{2} h^2 + c^2.$$

Solving for h^2 and c^2 , we find that our heritability estimate is equivalent to twice the difference between the MZ and DZ correlations,

$$h^2 = 2(r_{MZ} - r_{DZ}),$$

while our estimate of shared environment is equivalent to the remainder of the MZ correlation not accounted for by heritability,

$$c^2 = 2 r_{DZ} - r_{MZ} = r_{MZ} - h^2.$$

For a given sample, the heritability, shared environment, and unshared environment components sum to one. Correlations, sample sizes, and the resulting estimates of h^2 , c^2 , and e^2 for our full set of studies are presented in table 3.

We do not here employ more complicated strategies for estimating heritability, including those that distinguish additive and non-additive sources of heritability or those that attempt to differentiate the main effect of genetic relatedness from the interaction between genetics and features of the environment that vary within populations. Although it is preferable to differentiate these effects in studies using micro-level data, in a meta-analysis one is largely constrained to the information available in published sources. As noted, by limiting our data requirements to only quantities derivative from r_{MZ} and r_{DZ} , we are able to include a sufficient number of samples so as to make the comparative aims of this analysis feasible.

Meta-Analysis

Meta-analysis is a statistical technique in which results from two or more studies of the same concept are mathematically combined to produce a single overall estimate of the effect size (Cooper, Hedges, and Valentine 1994). Although in its simplest form, a meta-analysis may simply be a weighted average, more advanced models are commonly used to take into account characteristics of

Table 3. Twin Correlations and Variance Component Estimates

Nationality ($k = 34$)	Sex	Cohort	r_{MZ}	N_{MZ}	r_{DZ}	N_{DZ}	b^2		c^2	
							$2(r_{MZ} - r_{DZ})$	$r_{MZ} - b^2$	$1 - r_{MZ}$	e^2
Australia (1)	Male	1	0.70	216	0.53	94	0.34	0.36	0.30	0.30
	Female	1	0.77	520	0.55	299	0.44	0.33	0.23	0.23
Australia (2)	Male	2	0.74	226	0.47	161	0.54	0.20	0.26	0.26
	Female	2	0.75	479	0.49	290	0.52	0.23	0.25	0.25
Australia (3)	Male	2	0.674	282	0.532	164	0.284	0.39	0.326	0.326
	Female	2	0.705	320	0.319	158	0.772	-0.067	0.295	0.295
Denmark	Male	2	0.62	4370	0.444	7068	0.352	0.268	0.38	0.38
Finland	Male	1	0.83	1506	0.58	3504	0.50	0.33	0.17	0.17
	Female	1	0.86	2028	0.62	3870	0.48	0.38	0.14	0.14
Germany	Male	Mixed	0.680	133	0.306	47	0.748	-0.068	0.320	0.320
	Female	Mixed	0.717	421	0.479	172	0.476	0.241	0.283	0.283
Italy	Male	Mixed	0.71	752	0.61	406	0.20	0.51	0.29	0.29
	Female	Mixed	0.79	1342	0.7	712	0.18	0.61	0.21	0.21
Norway (1)	Male	1	0.86	259	0.77	313	0.18	0.68	0.14	0.14
	Female	1	0.89	405	0.75	425	0.28	0.61	0.11	0.11
Norway (2)	Male	1	0.82	253	0.48	284	0.68	0.14	0.18	0.18
	Female	1	0.85	342	0.68	400	0.34	0.51	0.15	0.15
Norway (3)	Male	2	0.85	370	0.47	463	0.76	0.09	0.15	0.15
	Female	2	0.89	518	0.66	576	0.46	0.43	0.11	0.11

Spain	Male	Mixed	128	0.758	0.519	155	0.478	0.28	0.242
	Female	Mixed	228	0.821	0.562	231	0.518	0.303	0.179
	Mixed	1	2492	0.76	0.55	3368	0.42	0.34	0.24
United States									
AddHealth	Male	2	100	0.611	0.477	94	0.268	0.343	0.389
	Female	2	117	0.623	0.650	93	-0.054	0.677	0.377
Vietnam Veterans	Male	1	1019	0.76	0.54	907	0.44	0.32	0.24
Minnesota	Male	Mixed	512	0.65	0.42	772	0.46	0.19	0.35
	Female	Mixed	758	0.72	0.57	1154	0.30	0.42	0.28
WW2 Veterans (NAS-NRC)	Male	1	1234	0.764	0.545	1167	0.438	0.326	0.236
MIDUS	Male	Mixed	164	0.668	0.538	124	0.293	0.375	0.332
	Female	Mixed	186	0.707	0.561	198	0.260	0.447	0.293
SRI	Male	Mixed	170	0.65	0.48	28	0.34	0.31	0.35
	Female	Mixed	390	0.68	0.50	123	0.36	0.32	0.32
United Kingdom (1)	Mixed	1	457	0.717	0.521	393	0.391	0.326	0.283
United Kingdom (2)	Mixed	2	388	0.593	0.474	247	0.238	0.355	0.407
Total			23,085			28,460			

studies that affect reliability of the effect sizes generated. Even in a set of studies testing the same hypothesis in which one believes that there is a single “true” underlying effect size, one may still expect some variation in the results due to study-specific differences.

Alternatively, one may suppose the possibility that there is no “true” effect size, but rather, that the true effect may differ across studies, even in the absence of study-specific error. In other words, studies may differ in their estimated effect sizes because of real differences in the effects being estimated, as opposed to differences due to sampling variation or study characteristics. In our case, we expect that environmental variation between our samples leads to differential exposure to factors such as national educational policies and cultural values, which in turn may alter the extent to which genes and environment explain variation in attainment. We thus consider a two-level hierarchical linear model with random effects to predict the true values of the h^2 , c^2 , and e^2 variance components in each study, as environmental differences may lead to sample-specific random variation (Cooper, Hedges, and Valentine 1994).

At the micro level, we assume that for study i , each observed variance component T_i is an estimate of a true study-specific variance component θ_i plus estimation error ϵ_i .¹⁰ As the sample sizes in all of the studies in our data are reasonably large, we assume that T_i is normally distributed around θ_i with variance v_i :

$$T_i = \theta_i + \epsilon_i, \quad \epsilon_i \sim N(0, v_i). \quad (1)$$

At the macro level, we expect that study-specific differences may lead to between-study variation in θ_i , and so we model the true value of the variance components as

$$\theta_i = \beta_0 + \beta_1 N_{i1} + \beta_2 S_{i2} + \beta_3 C_{i3} + u_i, \quad u_i \sim N(0, \sigma_u^2), \quad (2)$$

where β_0 is the intercept; N_i is the nationality of respondents in study i ; S_i is the sex of respondents in study i ; C_i is the birth cohort of respondents in study i ; and u_i is the random effect of study i , reflecting the deviation of the true value of the variance component from the value predicted by the study characteristics. Random effect u_i is normally distributed with mean zero and variance σ_u^2 , the between-studies variance. Combining the two models,

$$T_i = \beta_0 + \beta_1 N_{i1} + \beta_2 S_{i2} + \beta_3 C_{i3} + u_i + \epsilon_i, \quad u_i + \epsilon_i \sim N(0, v_i^*), \quad (3)$$

where the total variance observed in T_i is $v_i^* = \sigma_u^2 + v_i$.

This model implies two sources of variation in true effect sizes: first, studies may differ by observed characteristics (nationality, sex, and birth cohort); second, studies may differ by a study-specific random effect u . If neither of these

sources contributes significant variation, the true effect size for all studies will be equal to the intercept β_0 plus any sampling error ε . In this case, the between-studies variance σ_θ^2 is zero, and all variation is attributed to ε . To generate the true effect size β_0 for all studies k under this assumption of homogeneity, we weight the results from each study i by the reciprocal of the sampling variances,

$$w_i = 1/v_i,$$

so that studies with more precise estimates of the variance components will be given more weight.

Since each T_i is only an unbiased estimate of β_0 when effect sizes are homogeneous, for each model presented in tables 4 and 5 we first evaluate homogeneity in true effect sizes by calculating the Q test statistic to detect presence or absence of heterogeneity between samples. As per Cochran (1954), Q is computed as the sum of squares of deviations from the overall estimate of effect size, with each study weighted by its inverse variance. The null hypothesis of homogeneity among effect sizes ($H_0: \sigma_\theta^2 = 0$) follows a chi-square distribution with $k - 1$ degrees of freedom, in which k denotes the total number of studies. While the Q -test is standard in the literature, it informs only about the presence or absence of heterogeneity, rather than about the magnitude of true heterogeneity between studies (Huedo-Medina et al. 2006). We thus also report the estimate of the magnitude of the between-study variance, σ_θ^2 , and the I^2 indicator of the proportion of the total variability in effect sizes that results from heterogeneity between studies rather than from random sampling error (Higgins and Thompson 2002; Higgins et al. 2003). The I^2 indicator has the additional benefit of being easy to interpret: a value of 0% ($I^2 = 0$) would indicate that all variability among estimates is due to sampling error, whereas a value of 100 percent ($I^2 = 100$) would indicate that all variability is due to heterogeneity between studies (Higgins and Thompson 2002). The confidence interval around the I^2 indicator is calculated using the formula suggested by Higgins and Thompson (2002). In the instances where we are able to reject the null of homogeneity, we proceed to evaluate possible sources of heterogeneity (model 3), recalculating our weights as

$$w_i = 1/v_i^*$$

to reflect the contribution of between-study variability to overall variability in the estimates.

To calculate v_i^* , we assume that since the estimates of r_{MZ} and r_{DZ} are calculated from two separate groups of respondents, they are statistically independent. The variances of the estimators h^2 and c^2 are therefore

$$Var(h^2) = Var(2(r_{MZ} - r_{DZ})) = 4(Var(r_{MZ}) + Var(r_{DZ}))$$

Table 4. Correlations and Heritability Variance Components, Full Sample (reference category: American males born pre-1950)

	r_{MZ}		r_{DZ}		b^2		c^2		e^2	
	1	2	1	2	1	2	1	2	1	2
Grand Mean	0.747*** (0.015)		0.551*** (0.018)		0.400*** (0.024)		0.361*** (0.026)		0.253*** (0.015)	
Nationality										
Australia	0.022 (0.018)		-0.023 (0.033)		0.097 (0.076)		-0.074 (0.069)		-0.022 (0.018)	
Denmark	-0.055 + (0.030)		-0.009 (0.046)		-0.095 (0.114)		0.038 (0.098)		0.055 + (0.030)	
Finland	0.116*** (0.016)		0.036 (0.029)		0.155* (0.067)		-0.042 (0.061)		-0.116*** (0.016)	
Germany	-0.012 (0.023)		-0.096 (0.066)		0.177 (0.146)		-0.187 (0.136)		0.012 (0.023)	
Italy	0.044* (0.020)		0.134*** (0.031)		-0.181* (0.075)		0.222*** (0.066)		-0.044* (0.020)	
Norway	0.147*** (0.013)		0.119*** (0.023)		0.055 (0.054)		0.092 + (0.049)		-0.147*** (0.013)	
Spain	0.084*** (0.025)		0.015 (0.047)		0.137 (0.106)		-0.056 (0.097)		-0.084*** (0.025)	
Sweden	0.032 (0.024)		-0.014 (0.039)		0.085 (0.091)		-0.057 (0.081)		-0.032 (0.024)	

United Kingdom	-0.056* (0.020)	-0.053 (0.043)	-0.014 (0.103)	-0.040 (0.091)	0.056* (0.020)
Female	0.035*** (0.007)	0.065*** (0.018)	-0.061 (0.041)	0.096* (0.038)	-0.035*** (0.007)
Post-1950 Birth Cohort	0.036* (0.014)	0.082* (0.027)	0.079 (0.061)	0.118* (0.056)	0.036* (0.014)
Intercept	0.711*** (0.012)	0.532*** (0.014)	0.366*** (0.047)	0.349*** (0.042)	0.289*** (0.012)
\bar{Q}	1010.08***	514.497***	143.114***	152.918***	1010.08***
σ^2_{θ}	0.007	0.008	0.012	0.013	0.007
I^2	96.73	93.59	76.94	78.42	96.73
95% CI	96.08–	91.97–	68.1–	70.32–	96.08–
	97.28	94.87	83.34	84.32	97.28

Note: Standard errors in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Table 5. Correlations and Heritability Variance Components, Comparison by Region

	r_{MZ}		r_{DZ}		b^2		c^2		e^2	
	1	2	1	2	1	2	1	2	1	2
Scandinavia ($k = 10$)										
Grand Mean	0.823*** (0.027)		0.601*** (0.033)		0.432*** (0.036)		0.385*** (0.042)		0.177*** (0.027)	
Female	0.072 (0.053)		0.110 (0.066)		-0.058 (0.122)		0.133 (0.117)		-0.072 (0.053)	
Post-1950 Birth Cohort	-0.041 (0.054)		-0.195 (0.068)		0.083 (0.118)		-0.136 (0.113)		0.041 (0.054)	
Intercept	0.803*** (0.039)		0.581*** (0.050)		0.434** (0.090)		0.367** (0.086)		0.197** (0.039)	
Q	679.442*** 0.007	387.970*** 0.000	402.630*** 0.008	165.609*** 0.001	74.492*** 0.012	75.335*** 0.004	94.409*** 0.013	88.626*** 0.004	679.442*** 0.007	387.970*** 0.000
σ^2_θ										
I^2	98.675	98.196	97.765	95.773	87.918	90.708	90.467	92.102	98.675	98.196
95% CI	98.284–	97.504–	96.961–	93.535–	79.830–	84.100–	84.610–	86.814–	98.284–	97.504–
	98.978	98.696	98.356	97.236	92.763	94.570	94.095	95.269	98.481	98.696

Note: Standard errors in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, * $p < 0.1$

and

$$\text{Var}(c^2) = \text{Var}(2r_{DZ} - r_{MZ}) = 4\text{Var}(r_{DZ}) + \text{Var}(r_{MZ}).$$

The variance of the estimated correlation coefficients can be accurately approximated by

$$\text{Var}(r) \approx \frac{(1 - r^2)^2}{n - 1},$$

where n is the sample size. The variances of the h^2 and c^2 estimators are then

$$\text{Var}(h^2) \approx 4 \left(\frac{(1 - r_{MZ}^2)^2}{n_{MZ} - 1} + \frac{(1 - r_{DZ}^2)^2}{n_{DZ} - 1} \right)$$

and

$$\text{Var}(c^2) \approx 4 \left(\frac{(1 - r_{DZ}^2)^2}{n_{DZ} - 1} \right) + \frac{(1 - r_{MZ}^2)^2}{n_{MZ} - 1}.$$

Since $e^2 = 1 - r_{MZ}$, the variance of e^2 is equal to the variance of r_{MZ} . The variance of e^2 will therefore be smaller than the variances of h^2 and c^2 , as can be seen in the forest plots presented in figures 1 through 3.

Our weighting system has the drawback that the estimates of h^2 , c^2 , and e^2 are functions of r_{MZ} and r_{DZ} , with the variances being monotonically decreasing functions of r_{MZ} and r_{DZ} . This could potentially bias estimates by over-weighting some effect sizes while underweighting others. For a meta-analysis involving simple estimates of correlation, this problem is usually addressed by using a variance stabilizing transformation; however, the transformations used for correlations are not suitable for linear functions of correlations. The potential for bias was investigated using simulations based on the empirical distribution of r_{MZ} and r_{DZ} values, and based on the results of these simulations, we concluded that any bias that exists in the estimates is small enough to be considered inconsequential.

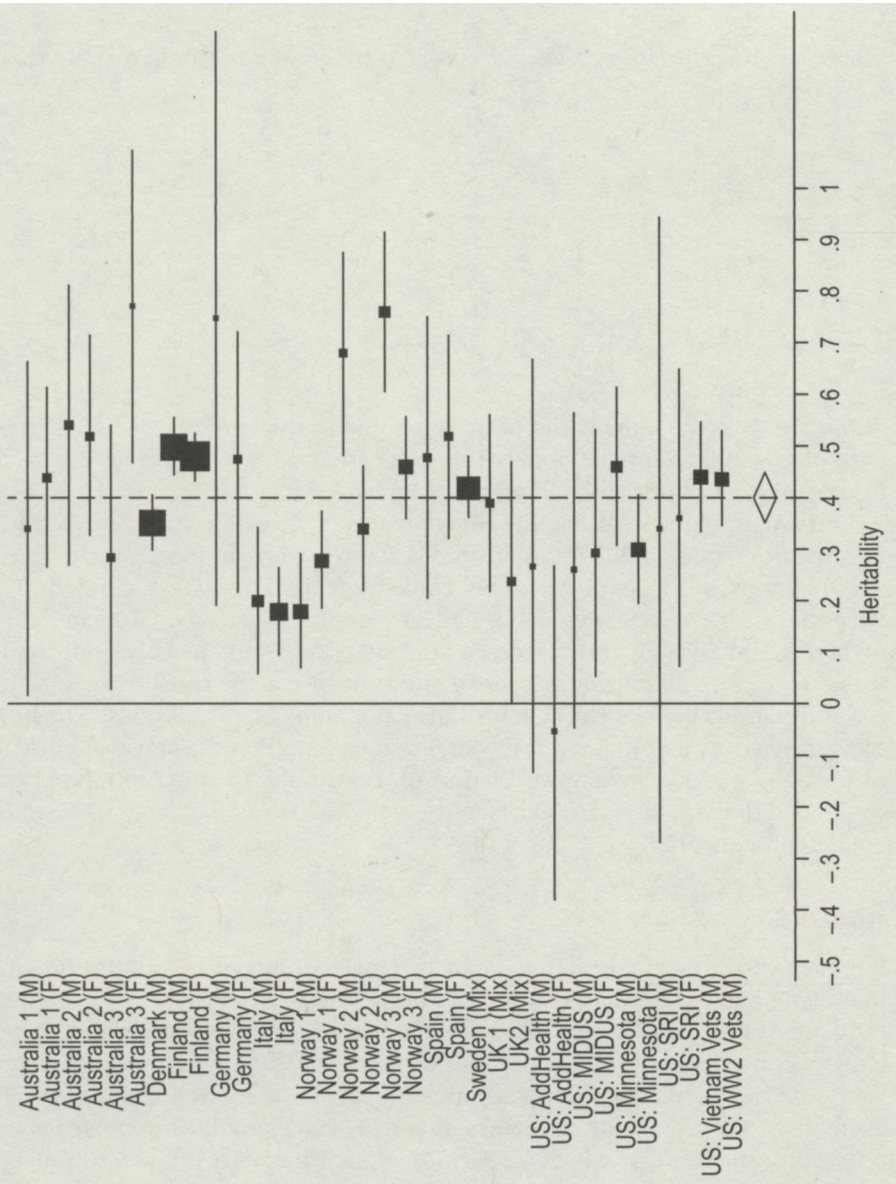
Results

For each of the twin correlations and variance component estimates presented in tables 4 and 5, our analysis proceeds in two steps. In step one, we generate an estimate of the grand mean (β_0), and calculate Q to test for sample homogeneity. If we are unable to reject the null of homogeneity, suggesting that the true values do not vary across populations, we interpret the grand mean as our best estimate of the true value. If we are able to reject the null of homogeneity, we proceed to step two, in which we seek to explain the observed heterogeneity by

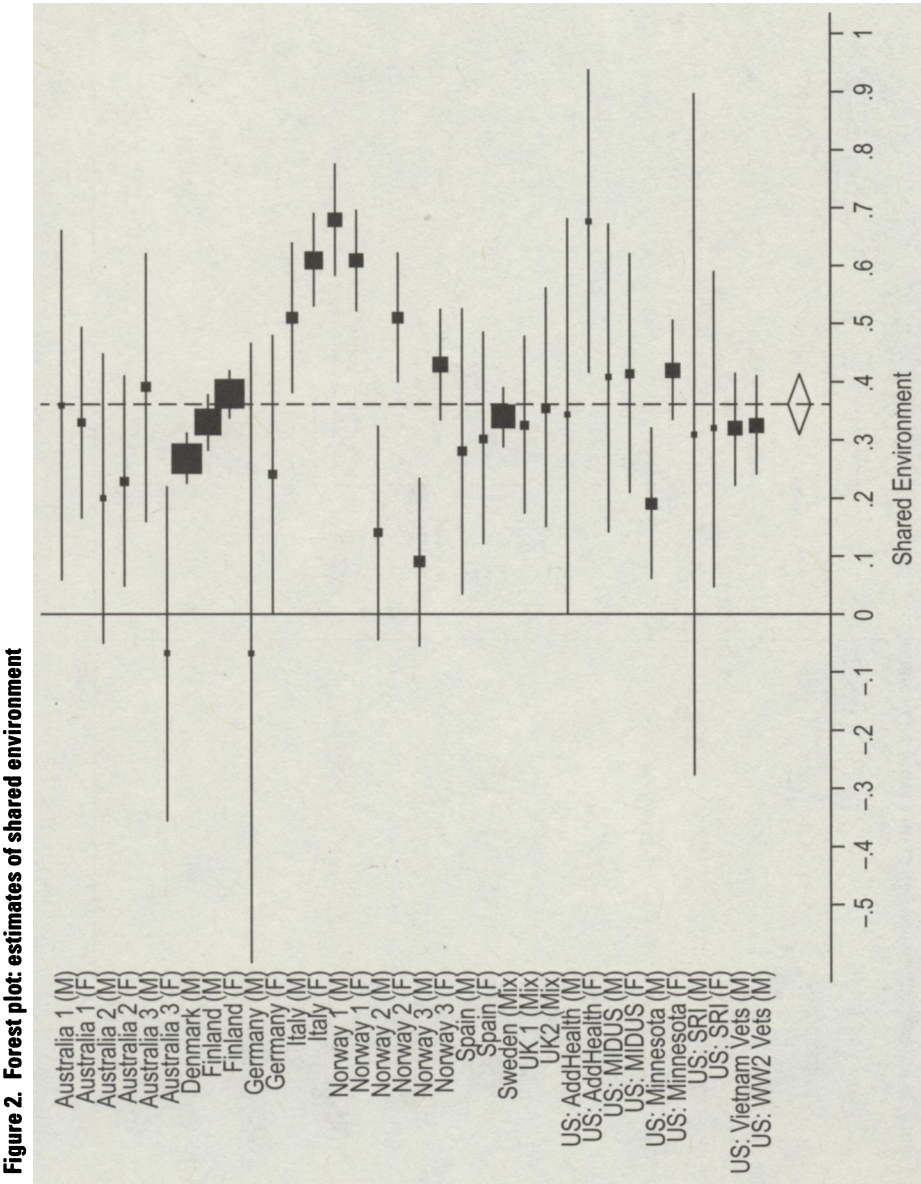
introducing covariates (nation, sex, and birth cohort) and study-specific random effects. Because $e^2 = 1 - r_{MZ}$, patterns of significance are identical across e^2 and r_{MZ} in all models.

In table 4, we present the results of our meta-analysis of the MZ and DZ correlations and the estimates of h^2 , c^2 , and e^2 across our full set of samples. In all models, heritability, shared environment, and unshared environment each

Figure 1. Forest plot: estimates of heritability



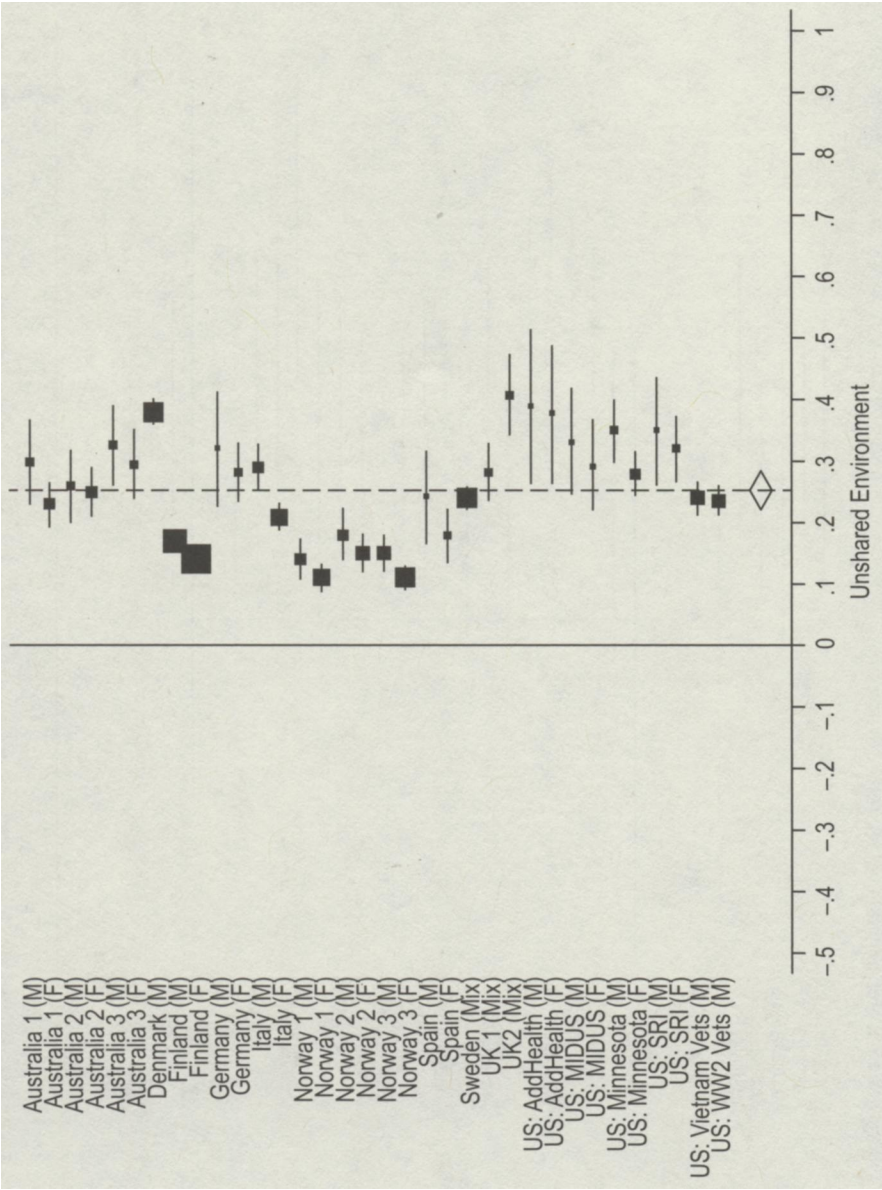
explain a percentage of the variance in attainment that is significantly greater than zero. As hypothesized, the contribution of each component varies by sample: we are able to reject the null of homogeneity in the grand mean across both sets of correlations and all three variance components, with a large percentage of the variation attributable to between-study variance (I^2 values range from 77 percent for h^2 to 97 percent for r_{MZ} and e^2). Our estimates of the magnitude of between-study variance (σ^2_θ) are also relatively large—for h^2 and c^2 , the



between-studies standard deviation is 0.11, equivalent to about one-fourth of the grand mean.

Upon introduction of random effects and covariates, we see that estimates of e^2 (and r_{MZ}) vary significantly by nation, sex, and birth cohort, with unshared environment explaining four percentage points more of the variation in attainment among men than women, and four percentage points less of the variation among twins born after 1950 than among those born earlier. These significant

Figure 3. Forest plot: estimates of unshared environment



group differences imply equivalent shifts in the sum of the remaining two variance components, which together constitute the total family effect. Microdata studies evaluating the influence of family on educational attainment across a range of Western nations have found family effects to either decline or remain relatively stable over the past century;¹¹ our non-zero but substantively meager decline in total family effect between cohorts in our international sample is consistent with these findings. Estimates of r_{DZ} also vary significantly by nation, sex, and birth cohort, with DZ correlations seven percentage points higher for women than men, and eight percentage points higher for those born in the later cohort than those born earlier.

Despite the small shift in the total family effect by sex and cohort, we see a far larger difference by sex and cohort in the composition of the family effect, with variation in genes explaining more of the variation in educational attainment for men and those born later in the century, and variation in shared environment explaining more of the variation for women and for respondents born earlier. Shared environment explains ten percentage points more of the variance in attainment for women than for men, and twelve percentage points less of the variance in attainment for those born after 1950 than for those born earlier. Conversely, genetic variation explains six percentage points less of the variance in attainment for women than men, and eight percentage points more of the variance in attainment for those born after 1950 than for those born earlier.¹² These differences are in the direction one would expect if women and respondents in the earlier cohort were exposed to greater social constraints with respect to educational attainment than men and respondents born later in the century.

As hypothesized, nation also appears to affect the extent to which genetic versus environmental differences account for variation in educational attainment. With the United States as our reference category, we observe significant differences ($p < .05$) in at least one variance component for five out of nine countries: Finland, Italy, Norway, Spain, and the United Kingdom. The direction of change among variance components varies by country. Our finding of high levels of heterogeneity across all models in our subgroup of Scandinavian samples (table 5) suggests that nation is indeed the salient factor driving the heterogeneity observed in table 4, rather than nation serving as a proxy for broader regional characteristics. While neither birth cohort nor sex are significant among the Scandinavian samples, the coefficients on sex and cohort for all three variance components are again in the direction one would expect if women and respondents in the earlier birth cohort were exposed to greater social constraints, similar to the full-sample results in table 4.¹³

Although we observed heterogeneity in the grand means for all three variance components in our full-sample analysis (table 4), the relative magnitudes of our grand mean estimates of heritability and common environment constitute an unanticipated yet notable finding. Counter to the well-established expectation that the estimate of common environment will be far below the estimate of heritability and often close to zero (Turkheimer 2004), here we observe grand means of heritability and common environment that are both large and substantively quite similar ($h^2 = 0.400$, while $c^2 = 0.361$). As our inability to reject

the null of homogeneity suggests that these estimates do vary by sample, the descriptive statistics presented in table 3 lend additional insight here: in eleven of our thirty-four samples, our estimate of common environment actually exceeds our estimate of heritability. Far from having a near-zero effect, in only five of our samples does the estimate of common environment fall below 20 percent, and in two it approaches 70 percent.

Discussion

As Hauser and Wong (1989) note, early efforts to separately estimate the genetic and environmental components of variation in social outcomes were often framed as attempts to “resolve the old debate about nature versus nurture as sources of social inequality” (151). Our results reflect a far more complicated relationship between nature and nurture than a simple dichotomy, suggesting that variables such as nation, sex, and birth cohort influence the extent to which genetic and environmental factors explain variation in educational attainment. While heritability, shared environment, and unshared environment each explain a significant percentage of the variation in attainment across our samples, we find that the relative contributions of the three components vary between studies. National differences explain a portion of this observed heterogeneity, while we do not observe a similar effect for regional differences more broadly. This finding is important to consider when attempting to generalize results from single-nation microdata studies of heritability to other nations, given the numerous unique environmental characteristics implicit in any national policy environment.

With respect to variation in the influence of genetic differences on educational attainment by sex and birth cohort, we find our coefficients to be consistent with twentieth-century attainment trends. For men and individuals born in the latter half of the twentieth century, more of the variance in attainment can be explained by genetic variation, whereas shared environment explains a greater percentage of the variance in attainment for women and those born in the earlier half of the century. In a mobility analysis where genetic and family social factors are assessed cumulatively, these shifts would be masked by the relatively small sex and cohort differences in the total family effect. We thus suggest calculating separate estimates of heritability and common environment as a method of distinguishing the impact of family characteristics that are transmitted strictly socially from those family characteristics that appear to follow from the greater genetic similarity within biological families (e.g., Björklund, Jännti, and Solon 2003; Jencks and Tach 2006; Nielsen 2008; Diewald 2010). Shifts in the balance between these two types of family effects may reflect shifts in patterns of mobility even when the sum of the two remains stable.

Although we observed significant heterogeneity among all three variance components in our full-sample analysis (table 4), the sheer magnitude of the common environment estimates we observe, consistently across many samples, may surprise those who are used to seeing very low common environment estimates from twin-study models. While such clear evidence of non-genetic, within-family influence on educational attainment may be less surprising to sociologists, the

unusual robustness of this result relative to other outcomes studied in behavioral genetics cannot be understated. Table 1 suggests that children from the same household are likely to differ more in their sexual orientation, athletic participation, personality traits such as perfectionism and extroversion, alcohol consumption, smoking habits, and even IQ than they do in their educational attainment. Indeed, for educational attainment, not only is our grand mean estimate of common environment of similar magnitude to our grand mean estimate of heritability, but in 32 percent of the studies in our sample, the estimated effect of common environment actually *exceeds* that of genetic differences (table 3).

We suggest that this divergence from prior literature in the effect of shared environment relative to genes may be in part due to differences between the outcome of educational attainment and the types of outcomes for which heritability estimates have been most commonly calculated. First, unlike the attitudes, beliefs, personality traits, and habitual behaviors listed in table 1, educational attainment has no intrinsic volatility. Attainment is determined with a clear cut-off, as either one has completed a year of education or one has not, and once attained, education can never be subsequently unattained. This greatly reduces concern about test-retest variation, as any characteristic that fluctuates within an individual can also be expected to fluctuate between twins, inflating the estimate of unshared environment. Second, even if ultimate attainment is often reached after a child has left the home, most educational attainment takes place during the years in which twins are likely to be living in the same household in the same neighborhood, going to the same schools, and so on. The divergent life trajectories that twins may experience after leaving the home can be expected to be less relevant for attainment than for later outcomes, even more traditional sociological outcomes such as income (Taubman 1976, 867). Future behavior genetic models of social outcomes may want to consider how these aspects of the outcome itself—intrinsic volatility and extent of common environment during the time the outcome occurred—may be expected to influence results, particularly when comparing estimates between outcomes that may differ in these respects.

Along with the benefits of meta-analysis, we also recognize its shortcomings. While we have attempted to address these issues, coding variation between studies, imperfect information regarding sex mix and cohort composition, and an inability to engage more advanced modeling techniques when calculating heritability estimates are among the trade-offs of obtaining a sample as large and diverse as the one we present here. Further analysis by subgroup would be a valuable extension of this research, and should be possible in the foreseeable future given the number of international twin studies in which respondents are simply too young at present for ultimate educational attainment to be a reasonable outcome.

In addition, advances in the availability of molecular genetic data may very soon allow significantly more fine-grained analyses than what we present here. To date, social science studies of gene-environment interaction using molecular genetic data have largely focused on small numbers of specific genetic variants, whose contribution to the overall population heritability of an outcome is unknown but almost certainly very small (Duncan and Keller 2011). Far more

promising are approaches based on genome-wide data, which can be used to identify causally implicated regions of the genome with far more breadth of coverage and fewer false positives than candidate gene approaches (Visscher et al. 2012). Genome-wide approaches can also be used to estimate the heritability of outcomes in samples of individuals who are not related, thereby addressing any concerns about particular assumptions of twin models (Benjamin et al. 2012); as noted earlier, the only published example to do so thus far for educational attainment found results similar to what had earlier been reported for twins in the same sample (Martin et al. 2011). As genome-wide data become more broadly available, these data can be used to probe variation in the relationship between genetic differences and educational attainment across populations, offering a more direct and comprehensive examination than is possible using the more basic twin models presented here.

Notes

1. See Feldman and Lewontin (1975) and Goldberger (1979), two early critiques of heritability estimates.
2. Our heritability estimate captures narrow-sense heritability, or additive genetic effects only (Plomin et al. 2008, 382). Although we are thus omitting dominant and epistatic variance from our estimate, research suggests that additive variance accounts for the majority, often close to 100 percent, of the total genetic variance for complex traits (Hill, Goddard, and Visscher 2008).
3. Some have suggested interpreting estimates of common environment as an upper bound on the extent to which policy can ameliorate inequality on a given social outcome (Rowe 1994). Among others, Jencks and Tach (2006) and Manski (2011) provide strong arguments for why the genetic or social etiology of a trait does not necessarily determine whether that trait can be modified at either the individual or population level.
4. See Breen et al. (2010) for twentieth-century attainment trends in Italy, Great Britain, Sweden, and Germany; Núñez (2003) on Spain; Alexander and Eckland (1974) and Everett et al. (2007) on the United States; and Booth and Kee (2009) on Australia.
5. Gender differences will vary by the definition of attainment used. Since our outcome is ultimate attainment, it is the mean attainment rates that most interest us here.
6. See Lowe (2012) on how World War II affected the structure of education in Australia, Italy, the United States, the United Kingdom, and Germany.
7. While differences in heritability by nation may also capture measurement differences across studies, structural variation between national education systems complicates efforts to distinguish scaling choices made by study investigators from actual between-country variation in the number of socially meaningful education categories. In all studies in our sample, both factors are likely relevant. As such, we chose not to standardize across scales, since by collapsing categories one loses more information from societies that make a greater number of educational distinctions than from societies that make fewer distinctions. For one robustness check of the effect of measurement variation on the heterogeneity of heritability of attainment within a single country, see footnote 13.
8. We include the UK samples as mixed-sex estimates due to an insufficient number of men to run sex-by-cohort correlations separately, but the precise ratio of males to females in these estimates is known.

9. The EM algorithm is a method for maximum likelihood estimation in the presence of latent variables. This approach does acknowledge the uncertainty in the value of missing cohort and sex data, which is reflected in the estimated sampling variance. We also ran our models using listwise deletion, and while significance drops off (as would be expected from the decrease in sample size), the magnitude of the coefficients on both sex and cohort remains unchanged from the model in which all samples were retained.
10. For more on random effects models in meta-analysis, see Chapter 16 in Cooper, Hedges, and Valentine (1994), from which our models were adapted.
11. A decline in the effect of family factors on educational attainment was found in populations drawn from the United States (Kuo and Hauser 1995), the Netherlands (Dronkers 1993a), and West Germany (de Graaf and Huinink 1992); no cohort change in the effect of family factors on attainment was observed in samples drawn from Australia (Borgers et al. 1995), Hungary (Toka and Dronkers 1996), or a second sample from the Netherlands (Dronkers 1993b).
12. To address the possibility that sex may be serving as a proxy for unmeasured characteristics—for example, veteran status, as our veteran samples consist of men only—we ran all models presented in table 3 on a subgroup consisting of only the fourteen samples for which we have data on both males and females separately. We find no notable differences in the significance or direction of the coefficient on sex across any of the five models.
13. As one robustness check regarding the impact of differences in measurement scales on heterogeneity in our variance components, we also ran our models on a subgroup consisting of only our ten US samples, in which coding scales ranged from six attainment categories (Minnesota) to thirteen attainment categories (AddHealth). After controlling for sex and birth cohort, we observed no residual heterogeneity in either MZ or DZ correlations or any of our three variance components, suggesting that variation in the number of coding categories alone was not sufficient to generate the extent of heterogeneity we observed by nation in our full sample.
14. In a sample from Los Angeles, estimates of common environment and heritability were both > 33 percent; in a nationally representative sample, the common environment estimate was substantially lower than 33 percent.
15. The heritability estimates for both IQ and achievement have been found to vary significantly by population subgroup (e.g., Guo and Stearns 2002; Rowe, Jacobson, and van den Oord 1999). Full population estimates have generally found heritability to be high and common environment to be low, with some exceptions (e.g., Thompson, Determan, and Plomin 1993).
16. We include the UK samples as mixed-sex estimates due to an insufficient number of males to calculate sex-by-cohort correlations. However, the precise ratio of males to females is known.

References

- Alexander, Karl L., and Bruce K. Eckland. 1974. "Sex Differences in the Educational Attainment Process." *American Sociological Review* 39:668–82.
- Alford, John R., Carolyn L. Funk, and John R. Hibbing. 2005. "Are Political Orientations Genetically Transmitted?" *American Political Science Review* 99:153–67.
- Ashenfelter, Orley, and Alan Krueger. 1994. "Estimates of the Economic Return to Schooling from a New Sample of Twins." *American Economic Review* 84:1157–73.

- Bailey, J. Michael, Richard C. Pillard, Michael C. Neale, and Yvonne Agyei. 1993. "Heritable Factors Influence Sexual Orientation in Women." *Archives of General Psychiatry* 50:217–23.
- Baker, Laura A., Susan A. Treloar, Chandra A. Reynolds, Andrew C. Heath, and Nicholas G. Martin. 1996. "Genetics of Educational Attainment in Australian Twins: Sex Differences and Secular Changes." *Behavior Genetics* 26(2):13.
- Behrman, Jere R., and Paul Taubman. 1989. "Is Schooling 'Mostly in the Genes'? Nature-Nurture Decomposition Using Data on Relatives." *Journal of Political Economy* 97:1425.
- Behrman, Jere R., Paul Taubman, and Terence Wales. 1977. "Controlling for and Measuring the Effects of Genetics and Family Environment in Equations for Schooling and Labor Market Success." In *Kinometrics: Determinants of Socioeconomic Success Within and Between Families*. New York: North-Holland Publishing Company.
- Benjamin, Daniel J., David Cesarini, Matthijs J. H. M. van der Loos, Christopher T. Dawes, Philipp D. Koellinger, Patrik K. E. Magnusson, Christopher F. Chabris, Dalton Conley, David Laibson, Magnus Johannesson, and Peter M. Visscher. 2012. "The Genetic Architecture of Economic and Political Preferences." *Proceedings of the National Academy of Sciences*.
- Bingley, Paul, Kaare Christensen, and Ian Walker. 2005. "Twin-Based Estimates of the Returns to Education: Evidence from the Population of Danish Twins." Aarhus, Denmark: Aarhus Business School.
- Björklund, Anders, Markus Jäntti, and Gary Solon. 2003. "Influences of Nature and Nurture on Earnings Variation: A Report on a Study of Various Sibling Types in Sweden." In *Unequal Chances: Family Background and Economic Success*, edited by S. Bowles, H. Gintis, and M. Osborne. New York: Russell Sage Foundation.
- Boardman, Jason D. 2009. "State-Level Moderation of Genetic Tendencies to Smoke." *American Journal of Public Health* 99:480–86.
- Boardman, Jason D., Jonathan Daw, and Jeremy Freese. Forthcoming. "Defining the Environment in Gene-Environment Research: Lessons from Social Epidemiology." *American Journal of Public Health*.
- Boardman, Jason D., Jarron M. Saint Onge, Brett C. Haberstick, David S. Timberlake, and John K. Hewitt. 2008. "Do Schools Moderate the Genetic Determinants of Smoking?" *Behavior Genetics* 38:13.
- Boomsma, Dorret, Gonneke Willemsen, Conor V. Dolan, Louise C. Hawkey, and John T. Cacioppo. 2005. "Genetic and Environmental Contributions to Loneliness in Adults: The Netherlands Twin Register Study." *Behavior Genetics* 35:745–52.
- Booth, Alison, and Hiau Joo Kee. 2009. "The University Gender Gap in Australia: A Long-run Perspective." Paper presented at the Conference Australian Economic History in the Long Run, Canberra, Australia, March 26–27.
- Borgers, N., J. Dronkers, L. Rollenberg, M. D. R. Evans, and J. Kelley. 1995. "The Decline of the Family and Educational Resemblance Between Australian Siblings: Birth Cohorts, Migration, Divorce, and Working Mothers." Paper presented at the Annual Meeting of the American Sociological Association, Washington, D.C.
- Bouchard, Thomas J., and Matt McGue. 1981. "Familial Studies of Intelligence: A Review." *Science* 212:1055–59.
- Breen, Richard, Ruud Luijkx, Walter Müller, and Reinhard Pollak. 2010. "Long-term Trends in Educational Inequality in Europe: Class Inequalities and Gender Differences." *European Sociological Review* 26:31–48.
- Clausson, Britt, Paul Lichtenstein, and Sven Cnattingius. 2000. "Genetic Influence on Birthweight and Gestational Length Determined by Studies in Offspring of Twins." *BJOG: An International Journal of Obstetrics & Gynaecology* 107:375–81.
- Cochran, William G. 1954. "The Combination of Estimates from Different Experiments." *Biometrics* 10:101–29.
- Cooper, Harris, Larry V. Hedges, and Jeffrey C. Valentine. 1994. *The Handbook of Research Synthesis and Meta-Analysis*. New York: Russell Sage Foundation.

- De Graaf, Paul M., and Johannes J. Huinink. 1992. "Trends in Measured and Unmeasured Effects of Family Background on Educational Attainment and Occupational Status in the Federal Republic of Germany." *Social Science Research* 21:28.
- Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society, Series B (Methodological)* 39:1–38.
- Devlin, B., Michael Daniels, and Kathryn Roeder. 1997. "The Heritability of IQ." *Nature* 388:468–71.
- Diebold, Martin. 2010. "On Genetic Variation and Social Inequality." *Zeitschrift für Soziologie* 39:4–21.
- Dronkers, J. 1993a. "Is the Importance of Family Decreasing? Evidence Regarding Dutch Sibling Data and Educational Attainment." Pp. 267–78 in *Adolescence, Careers, and Cultures*, edited by W. Meeus, M. de Goede, W. Kox, and K. Hurrelman. Berlin/New York: DeGruyter.
- . 1993b. "Educational Reform in the Netherlands: Did It Change the Impact of Parental Occupation and Education?" *Sociology of Education* 6(4):262–77.
- Duncan, L. E., and M. C. Keller. 2011. "A Critical Review of the First 10 Years of Candidate Gene-by-Environment Interaction Research in Psychiatry." *American Journal of Psychiatry* 168(10):1041–49.
- Dunne, M. P., N. G. Martin, D. J. Statham, W. S. Slutske, S. H. Dinwiddie, K. K. Bucholz, P. A. F. Madden, and A. C. Heath. 1997. "Genetic and Environmental Contributions to Variance in Age at First Sexual Intercourse." *Psychological Science* 8:211–16.
- Durbin, Nancy E., and Lori Kent. 1989. "Postsecondary Education of White Women in 1900." *Sociology of Education* 62(1):1–13.
- Eckland, Bruce K. 1967. "Genetics and Sociology: A Reconsideration." *American Sociological Review* 32:173–94.
- Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton University Press.
- Everett, Bethany G., Richard G. Rogers, Patrick M. Krueger, and Robert A. Hummer. 2007. "Trends in Educational Attainment by Sex, Race/Ethnicity, and Nativity in the United States, 1989–2005." Working paper, Population Program, Institute of Behavioral Science, University of Colorado at Boulder, Boulder, CO. <http://www.colorado.edu/ibs/pubs/pop/pop2007-0007.pdf>.
- Feldman, M. W., and R. C. Lewontin. 1975. "The Heritability Hang-Up." *Science* 190:1163–68.
- Fowler, James H., Laura A. Baker, and Christopher T. Dawes. 2008. "Genetic Variation in Political Participation." *American Political Science Review* 102:233–48.
- Freese, Jeremy. 2008. "Genetics and the Social Science Explanation of Individual Outcomes." *American Journal of Sociology* 114:35.
- Goldberger, Arthur S. 1979. "Heritability." *Economica* 46:327–47.
- Guo, Guang, Michael E. Roettger, and Tianji Cai. 2008. "The Integration of Genetic Propensities into Social-Control Models of Delinquency and Violence among Male Youths." *American Sociological Review* 73:543–68.
- Guo, Guang, and Elizabeth Stearns. 2002. "The Social Influences on the Realization of Genetic Potential for Intellectual Development." *Social Forces* 80:881–910.
- Hatemi, Peter K., John R. Hibbing, Sarah E. Medland, Matthew C. Keller, John R. Alford, Kevin B. Smith, Nicholas G. Martin, and Lindon J. Eaves. 2010. "Not by Twins Alone: Using the Extended Family Design to Investigate Genetic Influence on Political Beliefs." *American Journal of Political Science* 54:798–814.
- Hauser, Robert M., and David L. Featherman. 1974. "Socioeconomic Achievements of U.S. Men, 1962 to 1972." *Science* 185:325–31.
- Hauser, Robert M., and Peter A. Mossel. 1985. "Fraternal Resemblance in Educational Attainment and Occupational Status." *American Journal of Sociology* 91:650–73.
- Hauser, Robert M., and Raymond Sin-Kwok Wong. 1989. "Sibling Resemblance and Intersibling Effects in Educational Attainment." *Sociology of Education* 62:149–71.

- Heath, A. C., K. Berg, L. J. Eaves, M. H. Salaas, L. A. Corey, J. Sundet, P. Magnus, and W. E. Nance. 1985. "Education Policy and the Heritability of Educational Attainment." *Nature* 314:734–6.
- Heath, A. C., K. K. Bucholz, P. A. F. Madden, S. H. Dinwiddie, W. S. Slutske, L. J. Bierut, D. J. Statham, M. P. Dunne, J. B. Whitfield, and N. G. Martin. 1997. "Genetic and Environmental Contributions to Alcohol Dependence Risk in a National Twin Sample: Consistency of Findings in Women and Men." *Psychological Medicine* 27:1381–96.
- Higgins, Julian P. T., and Simon G. Thompson. 2002. "Quantifying Heterogeneity in a Meta-Analysis." *Statistics in Medicine* 21:1539–58.
- Higgins, Julian P. T., Simon G. Thompson, Jonathan J. Deeks, and Douglas G. Altman. 2003. "Measuring Inconsistency in Meta-Analyses." *BMJ* 327:557–60.
- Hill, William G., Michael E. Goddard, and Peter M. Visscher. 2008. "Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits." *PloS Genetics* 4:10.
- Huedo-Medina, Tania B., Julio Sánchez-Meca, Juan Botella, and Fulgencio Marín-Martínez. 2006. "Assessing Heterogeneity in Meta-Analysis: Q Statistic or I² Index?" *Psychological Methods* 11:14.
- Isacsson, Gunnar. 1999. "Estimates of the return to schooling in Sweden from a large sample of twins." *Labour Economics* 6(4):471–89.
- Jencks, C. 1980. "Heredity, Environment, and Public Policy Reconsidered." *American Sociological Review* 45:723–36.
- Jencks, Christopher, and Laura Tach. 2006. "Would Equal Opportunity Mean More Mobility?" Pp. 23–58 in S. Morgan, D. Grusky, and G. Fields, eds., *Mobility and Inequality: Frontiers of Research from Sociology and Economics*. Stanford, CA: Stanford University Press.
- Kato, Kenji, and Nancy Pedersen. 2005. "Personality and Coping: A Study of Twins Reared Apart and Twins Reared Together." *Behavior Genetics* 35:147–58.
- Kendler, Kenneth S., Margaret Gatz, Charles O. Gardner, and Nancy L. Pedersen. 2006. "A Swedish National Twin Study of Lifetime Major Depression." *American Journal of Psychiatry* 163:6.
- Kuo, Hsiang-Hui Daphne, and Robert M. Hauser. 1995. "Trends in Family Effects on the Education of Black and White Brothers." *Sociology of Education* 68:136–60.
- _____. 1997. "How Does Size of Sibship Matter? Family Configuration and Family Effects on Educational Attainment." *Social Science Research* 26:25.
- Li, Ming D., Rong Cheng, Jennie Z. Ma, and Gary E. Swan. 2003. "A Meta-Analysis of Estimated Genetic and Environmental Effects on Smoking Behavior in Male and Female Adult Twins." *Addiction* 98:23–31.
- Lowe, Roy, ed. 2012. *Education & the Second World War: Studies in Schooling & Social Change*. New York: Routledge.
- Lykken, D. T., T. J. Bouchard, M. McGue, and A. Tellegen. 1990. "The Minnesota Twin Family Registry: Some Initial Findings." *Acta geneticae medicae et gemellologiae* 39:35.
- Manski, Charles F. 2011. "Genes, Eyeglasses, and Social Policy." *Journal of Economic Perspectives* 25(4):9.
- Martin, Nicolas W., Sarah E. Medland, Karin J. H. Verweij, S. Hong Lee, Dale R. Nyholt, Pamela A. Madden, Andrew C. Heath, Grant W. Montgomery, Margaret J. Wright, and Nicholas G. Martin. 2011. "Educational Attainment: A Genome-Wide Association Study in 9538 Australians." *PLoS ONE* 6(6):8.
- McGue, Matt, and David T. Lykken. 1992. "Genetic Influence on Risk of Divorce." *Psychological Science* 3:6.
- McGuire, Shirley, Jenae M. Neiderhiser, David Reiss, E. Mavis Hetherington, and Robert Plomin. 1994. "Genetic and Environmental Influences on Perceptions of Self-Worth and Competence in Adolescence: A Study of Twins, Full Siblings, and Step-Siblings." *Child Development* 65:785–99.
- Miller, Paul, Charles Mulvey, and Nick Martin. 1997. "Family Characteristics and the Returns to Schooling: Evidence on Gender Differences from a Sample of Australian Twins." *Economica* 64:137–54.
- Moayyeri A., C. J. Hammond, A. M. Valdes, T. D. Spector. 2012. "Cohort Profile: TwinsUK and Healthy Ageing Twin Study." *International Journal of Epidemiology*.

- Murray, Sheila E., William N. Evans, and Robert M. Schwab. 1998. "Education-Finance Reform and the Distribution of Education Resources." *American Economic Review* 88:789–812.
- Neale, Michael C., and Hermine H. M. Maes. 2004. *Methodology for Genetic Studies of Twins and Families*. Dordrecht, The Netherlands: Kluwer Academic Publishers B.V.
- Nielsen, François. 2006. "Achievement and Ascription in Educational Attainment: Genetic and Environmental Influences on Adolescent Schooling." *Social Forces* 85:193–216.
- _____. 2008. "The Nature of Social Reproduction: Two Paradigms of Social Mobility." *Sociologica* 3. doi: 10.2383/28771.
- Núñez, Clara Eugenia. 2003. "Within the European Periphery: Education and Labor Mobility in Twentieth-Century Spain." *Paedagogica Historica* 39:621–49.
- Olson, James M., Philip A. Vernon, Julie Aitken Harris, and Kerry L. Jang. 2001. "The Heritability of Attitudes: A Study of Twins." *Journal of Personality and Social Psychology* 80:845–60.
- Pedersen, Nancy L., Robert Plomin, Gerald E. McClearn, and Lars Friberg. 1988. "Neuroticism, Extraversion, and Related Traits In Adult Twins Reared Apart and Reared Together." *Journal of Personality and Social Psychology* 55:950–57.
- Plomin, Robert, John C. DeFries, Gerald E. McClearn, and Peter McGuffin. 2008. *Behavioral Genetics, 5th Edition*. New York: Worth Publishers.
- Robinson, JoAnn L., Jerome Kagan, J. Steven Reznick, and Robin Corley. 1992. "The Heritability of Inhibited and Uninhibited Behavior: A Twin Study." *Developmental Psychology* 28:1030–37.
- Rowe, David C. 1994. "The Limits of Family Influence: Genes, Experience, and Behavior." New York: Guilford Press.
- Rowe, David C., Kristen C. Jacobson, and Edwin J. C. G. van den Oord. 1999. "Genetic and Environmental Influences on Vocabulary IQ: Parental Education Level as Moderator." *Child Development* 70:1151–62.
- Rushton, J. Philippe, David W. Fulker, Michael C. Neale, David K. B. Nias, and Hans J. Eysenck. 1986. "Altruism and Aggression: The Heritability of Individual Differences." *Journal of Personality and Social Psychology* 50:1192–98.
- Scourfield, J., N. Martin, G. Lewis, and P. McGuffin. 1999. "Heritability of Social Cognitive Skills in Children and Adolescents." *British Journal of Psychiatry* 175:559–64.
- Sieben, Inge, and Paul M. de Graaf. 2003. "The Total Impact of the Family on Educational Attainment." *European Societies* 5:35.
- Silventoinen, Karri, Robert F. Krueger, Thomas J. Bouchard, Jaakko Kaprio, and Matt McGue. 2004. "Heritability of Body Height and Educational Attainment in an International Context: Comparison of Adult Twins in Minnesota and Finland." *American Journal of Human Biology* 16:544–55.
- Smith, Herbert L., and Paul P. L. Cheung. 1986. "Trends in the Effects of Family Background on Educational Attainment in the Philippines." *American Journal of Sociology* 91:1387–408.
- Sung, Joohon, Sung-Il Cho, Yun-Mi Song, Kayoung Lee, Eun-Young Choi, Mina Ha, Jihae Kim, H. O. Kim, Yeonju Kim, Eun-Kyung Shin, Yoon-Hee Kim, Keun-Young Yoo, Chan Park, and Kuchan Kimm. 2006. "Do We Need More Twin Studies? The Healthy Twin Study, Korea." *International Journal of Epidemiology* 35:488–90.
- Taubman, Paul. 1976. "The Determinants of Earnings: Genetics, Family, and Other Environments: A Study of White Male Twins." *American Economic Review* 66:858–70.
- Thompson, Lee, Douglas Detterman, and Robert Plomin. 1993. "Differences in Heritability across Groups Differing in Ability, Revisited." *Behavior Genetics* 23:331–36.
- Toka, Gabor, and Jaap Dronkers. 1996. "Sibling Resemblance in Educational Attainment, Occupational Prestige, and Wealth in Hungary during the Communist Regime." *European Sociological Review* 12:251–69.
- Tozzi, Federica, Steven H. Aggen, Benjamin M. Neale, Charles B. Anderson, Suzanne E. Mazzeo, Michael C. Neale, and Cynthia M. Bulik. 2004. "The Structure of Perfectionism: A Twin Study." *Behavior Genetics* 34:483–94.

- Turkheimer, Eric. 2000. "Three Laws of Behavior Genetics and What They Mean." *Current Directions in Psychological Science* 9:5.
- . 2004. "Spinach and Ice Cream: Why Social Science Is So Difficult." Pp. 161–89 in *Behavior Genetics Principles: Perspectives in Development, Personality, and Psychopathology*, edited by L. DiLalla. Washington, DC: American Psychological Association.
- Turkheimer, Eric, Andreana Haley, Mary Waldron, Brian D'Onofrio, and Irving I. Gottesman. 2003. "Socioeconomic Status Modifies Heritability of IQ in Young Children." *Psychological Science* 14:623–8.
- Turkheimer, Eric, and Mary Waldron. 2000. "Nonshared Environment: A Theoretical, Methodological, and Quantitative Review." *Psychological Bulletin* 126:31.
- Visscher, Peter M., Matthew A. Brown, Mark I. McCarthy, and Jian Yang. 2012. "Five Years of GWAS Discovery." *American Journal of Human Genetics* 90:7–24.
- Winerman, Lea. 2004. "A Second Look at Twin Studies." *Monitor on Psychology* 35:46.