# Educational tracking and the polygenic prediction of education

Hannu Lahtinen,[a,b] Pekka Martikainen,[a,b,c] Kaarina Korhonen,[a,b] Tim Morris,[d] Mikko Myrskylä[a,b,c]

a) University of Helsinki; b) Max Planck – University of Helsinki Center for Social Inequalities in Population Health;

c) Max Planck Institute for Demographic Research; d) University College London

**Abstract:** Educational systems that separate students into curriculum tracks later may place less emphasis on socioeconomic family background and allow individuals' personal skills and interests more time to manifest. We tested whether postponing tracking from age 11 to 16 results in stronger genetic prediction of education across a population, exploiting the natural experiment of the Finnish comprehensive school reform between 1972 and 1977. The association between polygenic score of education and achieved education strengthened after the reform by one-third among men and those from low-educated families. We observed no evidence for reform effect among women or those from high-educated families. The first cohort experiencing the new system had the strongest increases. From the perspective of genetic prediction, the school reform promoted equality of opportunity and optimal allocation of human capital. The results also suggest that turbulent circumstances, including puberty or ongoing restructuring of institutional practices, may strengthen genetic associations in education.

**Keywords:** educational tracking; educational attainment; polygenic score; gene–environment interaction; natural experiment

**Replication Package:** Instructions for data access and code to reproduce the analysis can be found at https://github.com/halahti/SocSci23

THE level of educational attainment is one of the most persistent social characteristics between generations. The correlation of years of education between parents and children is typically around 0.4 in Western Europe (Bjöörklund and Jäntti 2020; Hertz et al. 2008). Twin decompositions indicate that the strong family-related effect stems from genetic and shared environmental components (Branigan, McCallum, and Freese 2013; Freese and Jao 2017; Silventoinen et al. 2020), and more recently developed molecular genetic methods support this finding (Lee et al. 2018; Okbay et al. 2022). In behavior genetics, it is frequently argued that the high heritability (within-population variance that is attributable to genetic differences between individuals) of education may act as an indicator of equality of educational opportunity (Ayorech et al. 2017; Conley 2016; Harden 2021; Plomin 2019; Selita and Kovas 2019; Silventoinen et al. 2020). The argument is that in equal environments free of societal obstacles or privileges, the differences between individuals may more greatly reflect individual factors. Thus, identifying the environments in which the genetic prediction of educational attainment is stronger can help us understand the extent to which educational contexts are meritocratic.

Previous research has shown that the amount of tracking in school curricula (i.e., the extent to which students are separated in differing streams, particularly at younger ages) is robustly related to the association between family socioeconomic background and children's educational attainment (Pekkarinen, Uusitalo, and Kerr 2009; Pfeffer 2008; van de Werfhorst 2018; van de Werfhorst and Mijs 2010). In systems in which students are tracked from an early age, family socioeconomic position predicts offspring's educational outcomes more strongly than in later-tracked systems with more universalist curricula. This is possibly because with later tracking, children and adolescents have more time to explore their interests and abilities before being sorted into more specialized tracks. In early-tracked systems, family resources independent of the child's characteristics may play a relatively more pronounced role due to their impact in early education (see also Mollborn et al. 2014).

In this study, we tested whether an analogous pattern is seen in stronger genetic predictions of education when tracking is postponed. We utilized the Finnish comprehensive school reform, conducted between 1972 and 1977, as a natural experiment. The reform abolished early tracking into academically and vocationally oriented schools at the age of 11 and replaced it with a system where students attend the same school until the age of 16. Due to exogenous sorting of individuals between the old and new systems, this design allowed us to circumvent a common limitation of many previous analyses of gene–environment interactions, namely the non-random distribution of genomes across environments (Schmitz and Conley 2017). Moreover, the gradual enrollment of the reform across the country between 1972 and 1977 allowed us to control for secular trends and regional differences potentially confounding the estimates on educational attainment. Specifically, we posed three research questions: (a) How did the school reform modify the association between the known genetic propensity for educational attainment and achieved education; (b) Were effects heterogeneous across individuals from different socioeconomic backgrounds or between men and women; and (c) Did effects vary according to the proximity of the reform?

## The Interplay of Genetics and the Environment in Education

Almost all human traits are to some degree heritable with a measurable genetic component (Polderman et al. 2015; Turkheimer 2000), and education is no exception to this (Branigan, McCallum, and Freese 2013; Silventoinen et al. 2020). Complex behavioral and social traits are "polygenic"; that is, they are associated with a huge number of single nucleotide polymorphisms (SNPs; the most fundamental unit of variation in DNA). Almost 4,000 SNPs have been identified that relate to educational attainment (Okbay et al. 2022). Where multiple SNPs have been identified to associate with a trait, their estimated effect sizes can be summarized into an index variable called a polygenic score (PGS; see Choi, Mak, and O'Reilly 2020; Mills and Tropf 2020; Mills, Barban, and Tropf 2020). A PGS consists of a sum of SNP allele counts, weighted by their effect sizes in predicting the outcome,

obtained from an independent genome-wide association study (GWAS). A PGS therefore provides an estimate of an individual's known genetic propensity for a trait using all SNPs that have been identified to associate with the trait.

The relationship between an individual's genome and their educational attainment is complex and operates through many pathways, such as psychological mechanisms including conscientiousness (Poropat 2009) and cognitive ability (Malanchini et al. 2020). SNPs associated with education have been identified to express predominantly in brain tissue and neurons in particular (Lee et al. 2018). However, genetic effects do not manifest into achieved education in a vacuum but operate in an environmental context of educational institutions shaped by social forces. Thus, the sociological study of the interplay between institutional conditions and individual genotypes presents an exciting opportunity to understand the formation and maintenance of educational inequalities. With PGSs, gene–environment interactions can be modeled straightforwardly in regression frameworks familiar to many sociologists.

However, reliable identification of gene–environment interactions has turned out to be difficult in practice. Analysts must be cautious of endogeneity, as the same reservations that restrict causal inferences from regression main-effect coefficients apply to interactions. A spurious gene–environment interaction can arise for various reasons. First, individuals tend to be distributed in different environments, such as childhood families, in a way that involves selection on genotypes. Second, individuals actively navigate toward and create environments that suit their genetic predispositions. Third, the environmental variable assessed in a gene–environment interaction analysis may not have a causal influence on the outcome, but may merely correlate with the causal factor.[1]

Because the obstacles in drawing causal inferences for interactions, including gene-environment interactions, are similar as for main effects, solutions can also be sought from the same direction. Exploiting natural experiments that exert exogenous effects on individuals, such as policy changes, constitutes one strategy (Schmitz and Conley 2017). If individuals cannot self-select into old and new policy regimes, concerns of non-randomly distributed genotypes across categories of environmental exposure are reduced. Here, we employed the Finnish comprehensive school reform as such a natural experiment.

## The Finnish Comprehensive School Reform

Between 1972 and 1977, Finland conducted a comprehensive school reform where the old selective two-track educational system was replaced by a universalist one-track system.[2] The pre-reform system started with a 4-year primary school ("kansakoulu") for all children between the ages of 6–7. At the age of 10–11, individuals chose either to stay in the primary school or enroll in a general secondary school ("oppikoulu"). Individuals who stayed in the primary school until the age of 13 continued to a 2–3-year civic school ("kansalaiskoulu"). In the late 1960s, roughly half of students took the general secondary track and half stayed in primary school (Pekkala Kerr, Pekkarinen, and Uusitalo 2013; Sysiharju 1969). Follow-up education (if any) for those finishing civic school was usually in vocational schools. Most of

those finishing general secondary school continued to a 3-year academic upper secondary school ("lukio" or "gymnasium"), which opened doors for university education (Sysiharju 1969).

In the post-reform system, by contrast, all students enrolled in a 9-year comprehensive school at the age of 6–7, and everyone followed the same curriculum until the age of 15–16 years. The new curriculum was academically oriented, and among the old-system alternatives, resembled the general secondary track more closely than the practically oriented civic school track. After the comprehensive school reform, students could continue either to academic upper secondary schools or vocational schools. Admission was based solely on comprehensive school grades (Pekkarinen 2008).

Before the school reform, private schools in secondary education (comprising 55 percent of the general secondary track students) charged tuition fees, although most of the expenses were covered by taxes (Pekkarinen, Uusitalo, and Kerr 2009). Tuition fees were abolished with the reform, and free-of-charge tuition at all levels of education was implemented. In addition, the new system introduced greater special education resources for lower-performing students (Kivirauma and Ruoho 2007).

It should be underlined that, unlike some comprehensive education reforms implemented in other countries during the same period, this reform did not increase mandatory schooling or minimum school-leaving age.[3] The possible effects shall thus not be attributable to the amount of mandatory schooling but rather the qualitative differences between the old two-track and new one-track systems.

The reform was implemented gradually between 1972 and 1977 in six different regions, as illustrated in Figure 1. The earliest adoption was in Lapland in Northern Finland and the last in the capital region of Helsinki (for a detailed presentation, see Pekkarinen, Uusitalo, and Kerr 2009). The municipality-specific implementation was planned by the National Board of Education and municipalities could not (at least officially) affect the schedule or curricula in the new system (Aho, Pitkaänen, and Sahlberg 2006). Students in Grades 1–5 in the implementation year were transferred to the new system, as well as all subsequent cohorts.

## Possible Consequences of the Reform on the Genetic Prediction of Education

Previous evidence implies that universalist curricula are less reflective of socioeconomic background in achievement in education (van de Werfhorst 2018; van de Werfhorst and Mijs 2010). Pekkarinen and colleagues (2009) observed that the Finnish comprehensive school reform increased intergenerational income mobility between sons and fathers, and Valkonen et al. (1996; 1998) found suggestive support for a decrease in social class background differences in educational attainment. The mechanism driving the equalizing effect of a universalist curriculum may lie in that students have more time to show their individual talents regardless of their families' resources. Following such reasoning, we may expect that early tracking is more weakly associated with individually embedded factors such as known
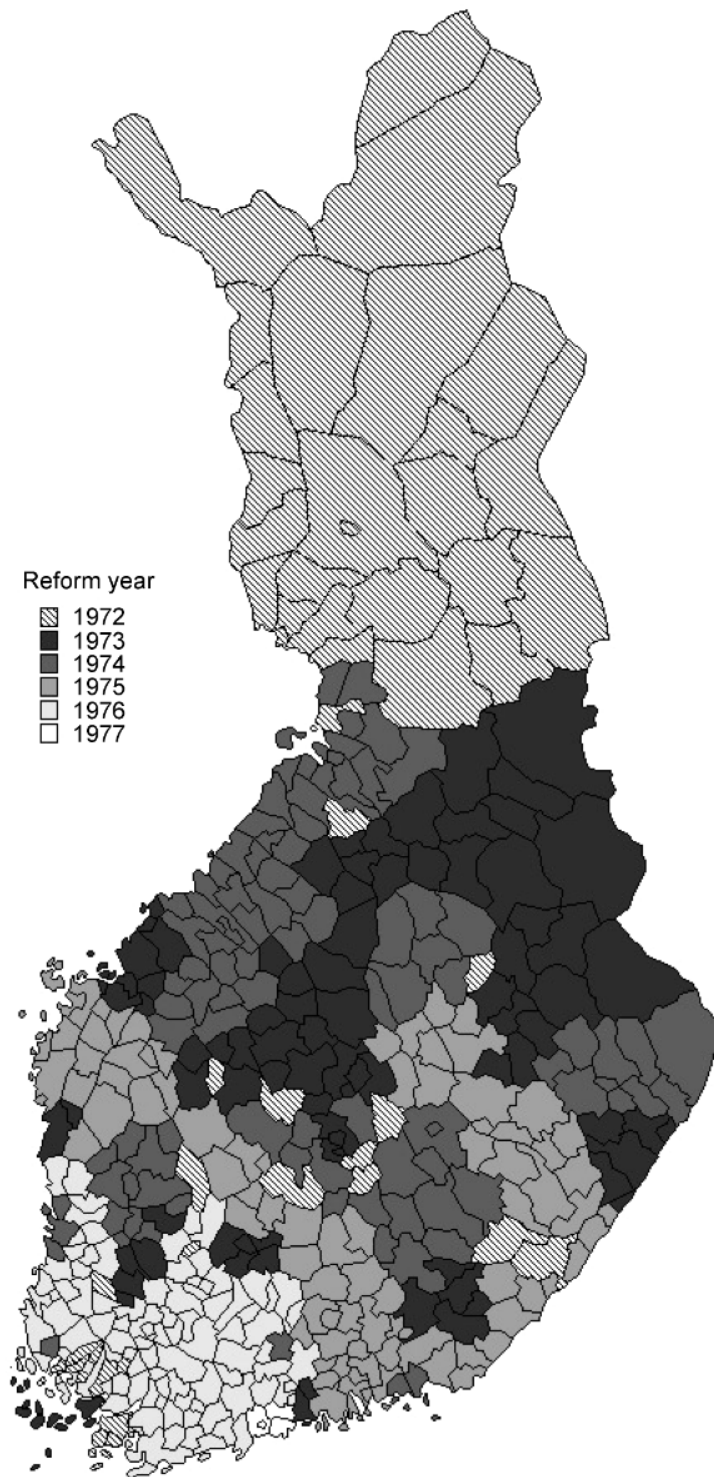
**Figure 1:** School reform implementation years across municipalities in Finland. Reproduced with permission from Pekkarinen (2008)

genetic propensity to education. Some twin studies have found corresponding associations between delayed tracking and higher heritability of education (Baier et al. 2022; Knigge et al. 2022; for a null result, see Mönkediek 2022), as well as higher heritability after some other equality-increasing school reforms (Colodro-Conde et al. 2015; Heath et al. 1985). Evidence from twin studies also indicate that the heritability of many traits (Polderman et al. 2015), including cognitive ability (Briley and Tucker-Drob 2013; Haworth et al. 2010), is higher later in the life-course than during childhood and early adolescence, whereas the shared environmental effect is lower. Thus, if tracks are determined at age 16 instead of 11, we should expect stronger genetic and weaker family environmental effects through psychological mechanisms.

The reform-induced change in genetic prediction of education may have a stronger effect among those from disadvantaged socioeconomic backgrounds. Although the comprehensive school reform affected every student (in contrast to, for example, school reforms that increase mandatory schooling years, which do not directly affect those who would have continued education regardless), the change was arguably larger for those who would have continued in the civic school track in the pre-reform system. These individuals were more likely to come from less socioeconomically advantaged families. By contrast, children of higher-educated parents were more likely to embark on the general track. In the 1960s, 12 percent of the general-track graduates' principal providers had academic upper secondary or university education compared to 4 percent of the overall adult population (Sysiharju 1969). Pekkala Kerr et al. (2013) observed that the Finnish comprehensive school reform increased cognitive test scores especially for individuals with low-educated parents. Other potential mechanisms for heterogeneous genetic prediction by socioeconomic background include decreasing economic costs of education, change in peer groups, and better prospects for future education, which were likely to disproportionately affect individuals from less socioeconomically advantaged origins.

The heterogeneous effect by socioeconomic background may also come from the opposite direction. Opponents of the reform raised concerns that the universal curriculum would also have an undesirable equalizing effect (Okkonen 2017). Following the critics' reasonings, the performance of students who would have enrolled in the general secondary track in the old system may have become hindered by the inclusion of less academically oriented peers in the classroom (Lazear 2001), or other loss of quality in the previously positively selected learning environments. As the individuals from advanced socioeconomic backgrounds were more likely to continue on the general secondary track in the old system, it is possible that the quality of their education provision was reduced and therefore genetic associations within this group were weakened.

Relatedly, there has been a discussion of the Scarr–Rowe interaction hypothesis in education. Such a hypothesis, originating from studies of socioeconomic differences in cognitive ability (Rowe, Jacobson, and Van den Oord 1999; Scarr-Salapatek 1971; Tucker-Drob and Bates 2016), states that individuals from more advantaged socioeconomic backgrounds have better possibilities to follow their motivations and abilities, whereas those with less advantaged backgrounds suffer more from

environmental obstacles. Applied to education, this implies the genetic prediction of education should be relatively higher among those with socioeconomically advantaged backgrounds and the environmental effects higher among individuals with less advantaged backgrounds. The evidence on such a pattern in education has been mixed overall, with possible contextual differences (Baier and Lang 2019; Baier et al. 2022; Erola et al. 2022; Lin 2020). Relatedly, Trejo and colleagues (2018) found that an education PGS was more predictive for educational attainment among socioeconomically advantaged high schools.

However, some studies have found a contrasting pattern, where the genetic associations with education are weaker among those with advantaged socioeconomic backgrounds, supporting the "Saunders"[4] hypothesis (Baier et al. 2022; Lin 2020). This hypothesis states that genetic effects are more pronounced in challenging or uncertain environments (Baier et al. 2022; Saunders 2010). Harden et al. (2020) found that socioeconomically advantaged schools may offer protection from a failure in class among students with lower PGS.[5] The Saunders hypothesis can be analogous to a "diathesis-stress" model of gene–environment interactions in epidemiological and psychopathological literature (Boardman, Daw, and Freese 2013; Manuck and McCaffery 2014). For example, environments with less social control or more limited personal resources can strengthen the genetic effects of addictive substance consumption and related harm (Neale et al. 2021; Pasman, Verweij, and Vink 2019).

In addition to socioeconomic background, gender-specific heterogeneous reform effects may be expected. A slight majority of the students in the old general secondary track were girls; for example, 57 percent in 1964/1965 (Sysiharju 1969). Thus, the reform possibly brought greater change for curricula of boys on average. Furthermore, in the new system, the finishing grade point average, on which the qualification for further education was solely based, was determined between the ages of 15–16. At this age, boys typically experience the most intense period of puberty, whereas girls have passed peak puberty (Euling et al. 2008). Puberty constitutes a candidate for producing an uncertain and challenging environment, which, following the Saunders/diathesis-stress hypotheses, can magnify genetic effects among men in particular.

There are also grounds to expect modification in the opposite direction. Studies from the United States and Finland have observed a strengthening polygenic prediction of education among women across cohorts, which experienced expansion of their educational opportunities (Herd et al. 2019; Lahtinen et al. 2023). It has also been argued that the new curriculum content, with increased emphasis on language, reading, and coursework, favored girls in general. Pekkarinen (2008) found that the Finnish comprehensive school reform increased the achieved educational level of women but not men. It is possible that these factors resonate to genetic prediction in a way that may enhance the genetic associations among women in particular.

Finally, it is possible that the reform's effect was dependent on the time-related proximity of its implementation. The "reform intensity" may have been lower among the first birth cohorts transferred to the new system, as they completed their first school years following the old curriculum (see Ollikainen 2021; Pekkarinen 2008). However, if the mechanism producing change in genetic prediction is the

tracking system and later sorting of students, such reform intensity should be irrelevant because the pioneering cohort already had the two-track system fully abolished in their secondary educational paths. Moreover, the pioneering cohorts also had the unique position of being the first to experience new practices, which had to be established and refined; these practices may have been more defined for later cohorts. Aligned with the Saunders and diathesis-stress hypotheses, such instability may lead genetic associations to manifest more strongly at the inception of the new educational regime.

## Educational Stratification in Finland

The association between parental and offspring education (Hertz et al. 2008; Pfeffer 2008), as well as that between siblings (Graätz et al. 2021; Lahtinen, Martikainen, and Tarkiainen 2022), in Finland is among the smallest in international comparisons. This reflects the objectives of the universalist welfare state, where the reduction of educational differentials by social background has been explicitly pursued (Kalalahti and Varjo 2020). The cohorts born in the early 1960s experienced a period of strong educational expansion; the average level of education increased strongly, particularly among women, whose educational level exceeded that of men among the studied cohorts (Härkönen and Sirniö 2020; Pekkarinen 2012). The genetic prediction of education, in turn, increased in Finland across the cohorts born early in the twentieth century, but was already stabilized among cohorts born at the turn of the 1960s (Lahtinen et al. 2023). The birth cohorts around the reform years may have even historically low intergenerational educational reproduction, with a Spearman's rank-order parent–child correlation of education around 0.25 (Härkönen and Sirniö 2020; Karhunen and Uusitalo 2017; Lahtinen et al. 2023). It is even possible that part of the weak intergenerational association may be attributed to the comprehensive school reform addressed in this study (Pekkarinen, Uusitalo, and Kerr 2009; Valkonen, Pensola, and Jalovaara 1996).

## Data & Methods

### Data and Variables

The sample used in this study consists of population-based FINRISK survey rounds from 1992 (n: 6,024), 1997 (n: 8,387), 2002 (n: 8,775), 2007 (n: 6,216), and 2012 (n: 5,748), collected every 5 years in North Karelia, Northern Savonia, Lapland, Northern Ostrobothnia, Kainuu, Turku and Loimaa, as well as Helsinki and Vantaa. These data are pooled together with data from the Health2000 survey (n: 6,693, with replacement n: 762 during an additional survey round in 2011), which covered the whole of mainland Finland. These data have been evaluated to represent the population of their target age well in the areas studied (Aromaa and Koskinen 2004; Borodulin et al. 2018). These studies included clinical examinations, during which DNA samples were collected. The response rates vary between 65 percent and 93 percent, with an overall average of 73 percent. The genetic data were put through quality control and imputation procedures according to SISu v3 reference panel

protocols (Pärn, Fontarnau, et al. 2018; Pärn, Isokallio, et al. 2018).[6] After quality control, genotyped information was available for 88 percent of all respondents. The genotyped data were linked to population registers including annual information on municipality of residence in 1972–1977 when the reform took place as well as the highest level of education in the household in 1970 and 1975, and the personal educational degrees obtained by the end of 2019. The linkage was conducted by statistics officials based on unique personal identity codes, and the research team analyzed pseudonymized dataset confidentially under supervision in the remote access environment of Statistics Finland.

We restricted the analysis to genotyped participants in cohorts ±4 years around the reform in their municipality of residence, including individuals from cohorts born between 1957 and 1969 and whose municipality of residence in 1972 could be determined in mainland Finland. This gave us an initial sample of 6,159 individuals. We excluded those who moved between municipalities with different school reform implementation years between 1972 and 1977 (n: 78). Finally, to mitigate possible bias between PGS and the outcome arising from a shared environment between related individuals, sometimes titled "cryptic relatedness" bias (Choi, Mak, and O'Reilly 2020; Mills, Barban, and Tropf 2020), we randomly excluded one individual from pairs sharing more than 0.177 of the variation of their genome (n: 135 excluded). These exclusions gave us the final sample size of 5,946 individuals.

We measured individuals' educational achievement as the years of education based on the expected number of years of schooling after basic level to obtain the highest degree an individual has. Following the guidelines of the International Standard Classification of Education 2011 (UNESCO 2012), these were the following: (a) no qualification beyond basic education (ISCED 0–2): 0 years; (b) upper-secondary education (ISCED 3): 3 years; (c) specialized vocational qualification (ISCED 4): 4 years; (d) lowest tertiary degree (ISCED 5): 5 years; (e) bachelor's, or equivalent, including basic polytechnic degree (ISCED 6): 6 years; (f) master's level or equivalent (ISCED 7): 8 years; and (g) doctoral or licentiate degree (ISCED 8): 12 years.

Figure 2 illustrates included birth cohorts with their respective number of observations in each reform region. The exposure to the school reform was an indicator variable defined based on the implementation year of the municipality of residence and year of birth. Students were transferred to the new system if they were in the fifth grade (i.e., they turned 11 in the year) during the year of reform implementation. This means an individual was assigned as exposed to the reform if the difference between the reform implementation in the municipality of residence and the year of birth was at least 11 (e.g., those born in 1961 or later were assigned "exposed" if the reform took place in 1972 in their municipality of residence). We also performed analyses with the exposure variable in three categories: (a) pre-reform cohorts (4 cohorts in each region); (b) the first cohort in the new system (1 cohort in each region); and (c) subsequent reform cohorts (3 cohorts in each region).

The PGS of educational attainment was defined according to the GWAS by Okbay et al. (2022). To avoid overfitting in the genetic data, individuals in our study sample were removed from the GWAS summary statistics. In addition, individuals from 23andme were excluded due to the company's privacy policies. A PGS provides a summary measure of the known genetic propensity for a trait, multiplying

| Reform year region | 1957 | 1958 | 1959 | 1960 | 1961 | 1962 | 1963 | 1964 | 1965 | 1966 | 1967 | 1968 | 1969 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1977 | | | | | | 94 | 58 | 78 | 66 | 86 | 80 | 79 | 68 |
| 1976 | | | | | 130 | 135 | 132 | 139 | 122 | 135 | 139 | 133 | |
| 1975 | | | | 110 | 138 | 135 | 125 | 128 | 116 | 122 | 132 | | |
| 1974 | | | 215 | 195 | 179 | 212 | 200 | 199 | 193 | 194 | | | |
| 1973 | | 152 | 141 | 136 | 127 | 123 | 111 | 111 | 124 | | | | |
| 1972 | 88 | 80 | 80 | 87 | 86 | 74 | 73 | 86 | | | | | |
| | 1957 | 1958 | 1959 | 1960 | 1961 | 1962 | 1963 | 1964 | 1965 | 1966 | 1967 | 1968 | 1969 |
| | **Birth cohort** | | | | | | | | | | | | |

**Figure 2:** Number of observations by birth cohort and reform region (total N: 5,946). Lighter-shaded cells are pre-reform cohorts ("non-exposed"), darker-shaded cells are post reform ("exposed") cohorts

the effect size (obtained from GWAS summary statistics) of a given copy of a SNP by the number of copies an individual has (0, 1 or 2). We employed the SBayesR method (Lloyd-Jones et al. 2019) to adjust the PGS for linkage disequilibrium (LD; correlation between occurrences of SNPs in different loci), using the base GWAS summary results by Okbay et al., as well as a banded LD matrix provided by the authors of the GCTB software (Zeng et al. 2020). These re-weighted scores were then assigned to the individuals using autosomal HapMap3 variants with a minor allele frequency of at least 0.01 in our data. The PGS is standardized to have a mean of 0 and standard deviation (SD) of 1.

The Finnish population is genetically homogenous, save for a difference between the Eastern and Western regions, following approximately the border set in the historical treaty of Pähkinäsaari in 1323 (Kerminen et al. 2017). To account for population stratification by genetic ancestry, we adjusted our models for the first 10 principal components of the genome (Price et al. 2006). The reform regions roughly correspond to this division, with early reform regions mainly consisting of areas in eastern and later reform regions of the western genetic sphere. Our region-specific analyses thus further controlled for such population stratification.

To account for the subtle bias that may have arose from, for example, differing genotyping equipment or differences in participant recruitment between data collections, our models were adjusted for genotyping batch and survey round. We also measured family education based on the highest education in the household in 1970, and if missing, in 1975, in two categories: all members of the household had basic education / at least one member had further education.

## Analysis Strategy

Our main parameter of interest was the difference-in-difference (DiD) estimate on whether the PGS's effect on achieved education was different before and after the reform. For these purposes, we performed estimation in two stages. First, we fit reform region–specific ordinary least squares regressions to estimate the interaction between PGS and reform status through the following model:

$$edu_i = b_0 + b_1 PGS_i + b_2 Reform_i + b_3 (PGS * Reform)_i + \overline{b_4}' \overline{Controls_i} + \epsilon_i \quad (1)$$

where $edu_i$ denotes years of education for individual i; $b_0$ is the intercept; PGS is the polygenic score of education; *Reform* is an indicator of being exposed to the school reform; and $\overline{Controls}$ is a vector of control variables. They include continuous year of birth, a female indicator, the first 10 principal components of the genome, data collection round indicators, and genotyping batch indicators. $\epsilon$ is the residual term. We used Huber–White standard errors to account for potential heteroscedasticity of residuals (Mansournia et al. 2021).

Second, we combined these separate reform region-specific estimates utilizing the fixed-effect inverse variance–weighted meta-analysis method (Borenstein et al. 2009, chap. 11). We conducted meta-analysis on the pre-reform effect parameter ($b_1$ of equation 1), post-reform effect parameter ($b_1+b_3$ of equation 1)[7], as well as the DiD estimate, which is the interaction parameter $b_3$. Calculating a meta-analytic DiD by subtracting the meta-analytic post-reform estimate from pre-reform estimate would create slightly different results. These differences are attributable to the heteroscedasticity correction of the standard errors, as it affects the parameters' given weight. Among these two options, we chose a strategy that was generally more conservative.

The rationale behind resorting to this two-step modeling strategy stems from the fact that the exposure to the reform was defined based on the combination of year of birth and reform region. Thus, including both in the model would create very high (or even full if they are measured categorically and interacted) multi-collinearity with the reform variable. This would also create complex models with hard-to-interpret parameters sensitive to lower-order interactions. Estimating within-region regression also controls for regional differences in education in a flexible manner, including potential interactive patterns.

After estimating the effect for the whole sample, we assessed heterogeneous effects by conducting identical estimation stratified by educational family background (basic education/at least one household member with a further degree during adolescence) and gender (men/women, excluding the gender indicator from regression). Finally, we also performed these analyses with an alternate formulation of the reform status including three categories (unexposed/first cohort of exposure/later cohorts of exposure).

Table 1 presents descriptive statistics of the variables used. The PGS of education and the achieved years of education had a correlation of 0.29. The average PGS was slightly lower in regions with earlier adoption of the reform, and the average level of education was also slightly lower in these regions. This might indicate genetic

**Table 1:** Descriptive statistics

|                                  | Mean | SD  |
|----------------------------------|------|-----|
| PGS of education                 | 0.00 | 1.0 |
| Years of education beyond basic  | 4.29 | 2.4 |
| Correlation: PGS*education        | 0.29 |     |

|                                          | % | PGS Mean | PGS SD | Years of education Mean | Years of education SD |
|------------------------------------------|------|------|------|------|------|
| **Reform exposure**                      |      |      |      |      |      |
| No                                       | 51   | 0.01 | 1.0  | 4.2  | 2.3  |
| Yes                                      | 49   | -0.01| 1.0  | 4.4  | 2.4  |
| **Reform exposure (alternative measurement)** |  |      |      |      |      |
| No                                       | 51   | 0.01 | 1.0  | 4.2  | 2.3  |
| 1st reform cohort                        | 13   | 0.01 | 1.0  | 4.3  | 2.4  |
| Subsequent cohorts                       | 36   | -0.02| 1.0  | 4.5  | 2.4  |
| **Gender**                               |      |      |      |      |      |
| Men                                      | 45   | 0.01 | 1.0  | 3.9  | 2.3  |
| Women                                    | 55   | 0.00 | 1.0  | 4.6  | 2.3  |
| **Year of birth**                        |      |      |      |      |      |
| 1957                                     | 1    | -0.12| 0.9  | 3.7  | 1.8  |
| 1958                                     | 4    | 0.03 | 0.9  | 3.9  | 2.2  |
| 1959                                     | 7    | 0.01 | 1.0  | 4.1  | 2.2  |
| 1960                                     | 9    | -0.06| 1.0  | 4.0  | 2.1  |
| 1961                                     | 11   | -0.05| 1.0  | 4.2  | 2.4  |
| 1962                                     | 13   | 0.11 | 1.0  | 4.3  | 2.4  |
| 1963                                     | 12   | -0.05| 1.0  | 4.2  | 2.4  |
| 1964                                     | 12   | -0.02| 1.0  | 4.4  | 2.4  |
| 1965                                     | 10   | 0.00 | 1.0  | 4.5  | 2.3  |
| 1966                                     | 9    | 0.05 | 1.0  | 4.5  | 2.5  |
| 1967                                     | 6    | 0.00 | 0.9  | 4.5  | 2.7  |
| 1968                                     | 4    | -0.05| 1.0  | 4.3  | 2.4  |
| 1969                                     | 1    | 0.16 | 1.0  | 4.9  | 2.6  |
| **Reform region**                        |      |      |      |      |      |
| 1972                                     | 11   | -0.09| 1.0  | 4.1  | 2.1  |
| 1973                                     | 17   | -0.06| 1.0  | 4.1  | 2.2  |
| 1974                                     | 27   | 0.00 | 1.0  | 4.3  | 2.3  |
| 1975                                     | 17   | 0.07 | 1.0  | 4.6  | 2.5  |
| 1976                                     | 18   | 0.01 | 1.0  | 4.2  | 2.4  |
| 1977                                     | 10   | 0.08 | 1.0  | 4.3  | 2.8  |
| **Highest family education**             |      |      |      |      |      |
| Only basic                               | 54   | -0.13| 1.0  | 3.8  | 2.1  |
| More than basic                          | 46   | 0.15 | 1.0  | 4.9  | 2.5  |
| Total (N)                                | 5,946|      |      |      |      |

**Table 2:** The association between one standard deviation change in education PGS and achieved years of education before and after the comprehensive school reform

|  | Pre-reform | | Post-reform | | Difference-in-difference | | |
|---|---|---|---|---|---|---|---|
|  | b | 95% CI | b | 95% CI | b | 95% CI | N |
| **Whole sample** | 0.59 | (0.51; 0.67) | 0.68 | (0.61; 0.76) | 0.08 | (-0.03; 0.19) | 5,946 |
| **By subgroup** | | | | | | | |
| Men | 0.49 | (0.37; 0.61) | 0.69 | (0.57; 0.80) | 0.17 | (0.01; 0.34) | 2,680 |
| Women | 0.66 | (0.55; 0.76) | 0.67 | (0.57; 0.78) | 0.01 | (-0.14; 0.15) | 3,266 |
| Basic family education | 0.44 | (0.35; 0.54) | 0.60 | (0.49; 0.71) | 0.16 | (0.02; 0.31) | 3,236 |
| More than basic family education | 0.67 | (0.54; 0.79) | 0.64 | (0.53; 0.76) | −0.04 | (-0.21; 0.13) | 2,710 |

Inverse-variance weighted (fixed effects) meta-analytic estimates from reform region -specific regressions including reform indicator, education PGS, and their interactions; adjusted for gender (if not stratified), year of birth, first ten principal components of the genome, study collection round and genotyping batch. These reform region -specific regression results are presented in Tables A1 and A2 in the appendix. Possible inequivalence between difference-in-difference estimates (which is the meta-analyzed interaction coefficient from reform region–specific regressions) and post-reform minus pre-reform estimates stems from the heteroscedasticity correction of the standard errors of reform-specific regressions.

population stratification effects across regions, as the reform roughly moved from north to south and east to west. This further warrants reform-region-specific estimation that implicitly adjusts for such differences. The lower number of observations in early and late birth cohorts stems from the fact that they are only included in the analysis of some regions; for example, the 1957 cohort is only included in the analysis of the region where the reform year was 1972 and cohort 1969 where the reform year was 1977 (see also Figure 2). The code to reproduce the analysis can be found at https://github.com/halahti/SocSci23.

## Results

Table 2 presents the association between the PGS of education and years of education before and after the reform, obtained from meta-analyses of reform-region-specific regressions presented in online supplement Tables A1 and A2. Among the whole study sample, a 1-SD-higher PGS predicted 0.59 (95% confidence interval [CI] 0.51; 0.67, p=$9*10^{-50}$) more achieved years of education before the reform and 0.68 (95% CI 0.61; 0.76, p=$6*10^{-66}$) more years after it. Thus, the reform saw a small increase in coefficients, which did not reach any conventional level of statistical significance in a two-tailed test (DiD estimate = 0.08, 95% CI -0.03; 0.19, p=0.167).

When investigating this change among individuals with differing levels of educational family background or gender, we observed evidence for an increase in association between PGS and education among those whose family members had only basic education (DiD=0.16, 95% CI 0.02; 0.31, p=0.029) and among men (DiD=0.17, 95% CI 0.01; 0.34, p=0.041). We did not observe changes in the coefficient of the PGS among those with higher than a basic-level educational background or among women. We also observed a stronger genetic prediction of education

**Table 3:** The association between one standard deviation change in education PGS and years of education before, during and after the comprehensive school reform

| | Pre-reform cohorts | | 1st Reform cohort | | Subsequent cohorts | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | b | 95% CI | b | 95% CI | b | 95% CI | N |
| **Whole sample** | 0.59 | (0.51; 0.67) | 0.79 | (0.64; 0.94) | 0.65 | (0.56; 0.74) | 5,946 |
| **By subgroup** | | | | | | | |
| Men | 0.49 | (0.37; 0.61) | 0.84 | (0.60; 1.09) | 0.63 | (0.50; 0.76) | 2,680 |
| Women | 0.66 | (0.55; 0.76) | 0.72 | (0.54; 0.90) | 0.65 | (0.52; 0.77) | 3,266 |
| Basic family education | 0.45 | (0.35; 0.54) | 0.79 | (0.57; 1.01) | 0.54 | (0.42; 0.66) | 3,236 |
| More than basic family education | 0.67 | (0.54; 0.79) | 0.70 | (0.50; 0.89) | 0.63 | (0.50; 0.77) | 2,710 |

**Difference-in-difference**

| | 1st vs. pre-reform | | Subseq. vs. pre-reform | |
| --- | --- | --- | --- | --- |
| | b | 95% CI | b | 95% CI |
| **Whole sample** | 0.17 | (0.00; 0.33) | 0.05 | (-0.07; 0.17) |
| **By subgroup** | | | | |
| Men | 0.33 | (0.05; 0.60) | 0.11 | (-0.06; 0.29) |
| Women | 0.04 | (-0.17; 0.25) | -0.01 | (-0.17; 0.15) |
| Basic family education | 0.34 | (0.10; 0.58) | 0.10 | (-0.06; 0.26) |
| More than basic family education | -0.03 | (-0.26; 0.21) | -0.04 | (-0.22; 0.14) |

Inverse-variance weighted (fixed effects) meta-analytic estimates from reform region -specific regressions including status in three categories, education PGS, and their interactions, adjusted for gender (if not stratified), year of birth, first ten principal components of the genome, study collection round and genotyping batch. These reform region -specific regression results are presented in Tables A3, A4 and A5 in the appendix. Possible inequivalence between difference-in-difference estimates (which is the meta-analyzed interaction coefficient from reform region–specific regressions) and post-reform minus pre-reform estimates stems from the heteroscedasticity correction of the standard errors of reform-specific regressions.

among individuals with higher than basic family education compared to those with low family education before the reform (difference in pre-reform estimates between educational backgrounds was 0.22, 95% CI 0.06; 0.38, p=0.007), but not after the reform (b=0.04, 95% CI -0.11; 0.20, p=0.581).[8] Overall, after the reform, PGS predicted education in a more homogenous manner across all the studied subgroups.

Table 3 presents the effects of the comprehensive school reform via an alternative measure of the reform, namely in three categories: cohorts before the reform (pre-reform cohorts), the first cohort that was exposed to the reform (1st reform cohort), and all subsequent cohorts. These are again based on meta-analysis reform-region-specific regressions presented in Tables A3, A4, and A5 in the online supplement. The PGS of education was strongly related to years of education among the cohort that was the first to experience the school reform, among whom a 1-SD-higher PGS was associated with 0.79 (95% CI 0.64; 0.94, $p=5*10^{-26}$) additional years of education. The corresponding figures were lower among cohorts more than 1 year after the

reform (0.65, 95% CI 0.56; 0.74, p=5*10$^{-44}$) as well as cohorts before the reform (0.59, 95% CI 0.51; 0.67, p=10$^{-49}$). The DiD of the coefficient of PGS between the first reform cohort and pre-reform cohorts was 0.17 (95% CI 0.00; 0.33, p=0.051).

Likewise in the previous analyses, the DiD results presented in Table 3 indicate a substantial reform effect for the first cohort among those with only basic parental education (0.34, 95% CI 0.10; 0.58, p=0.006) and among men (0.33, 95% CI 0.05; 0.60, p=0.019), whereas no evidence for reform effect were observed among those whose parents had more than basic education or among women.

To test the robustness of these results, Figure A6 in the online supplement presents an additional analysis where the PGS is interacted with a categorical year-to-reform variable (ranging between -4 and +3, 0 denoting reform year) in predicting the years of education. The results of this analysis were consistent with our earlier results: Among those with only basic parental education and among men, there were larger coefficients after the reform, and the first reform cohort had clearly the strongest association. Of note, two cohorts before the reform were positive outliers among those with more than basic parental education and among women. However, there was substantial uncertainty in these estimates, as evidenced by their wide 95 percent CIs, and these results should therefore be interpreted with caution.

## Discussion

In this study, we investigated whether the comprehensive school reform, implemented regionally between 1972 and 1977 in Finland and replacing the old two-track basic educational system with a universal curriculum for all students, affected the polygenic prediction of educational attainment. Our first research question was how the universalist curriculum reform modified associations for the known genetic propensity for educational attainment. We observed weak evidence for a modest increase in the strength of the association between education PGS and achieved education among the cohorts exposed to the reform. When investigating this effect more closely in population subgroups, we found this modest overall effect was driven by more substantial effects among men and those whose family members had only basic education. In turn, among women and those from higher-educated families, we observed no evidence for such an effect. Thus, our results answer our second research question that heterogeneous effects existed across genders and socioeconomic backgrounds. Our results provide partial support for the results of a twin study from the Netherlands (Knigge et al. 2022), which indicated greater heritability of education when there is less tracking. Our study exploited a design that allows for stronger causal interpretation, and our strategy of using directly measured genetic variation allowed for triangulating the previous evidence from alternative (twin) methods with orthogonal limitations.

We observed that the reform increased genetic prediction among adolescents whose family members had no education beyond a basic level, but the estimate of reform effect was close to nil among individuals from higher-educated families. Correspondingly, we observed a Scarr–Rowe interaction effect—namely stronger genetic prediction among individuals with more advantaged socioeconomic position—before but not after the reform. This suggests that universalism in educational

curricula may be one mechanism that explains the context-specific emergence of the Scarr–Rowe interaction (Baier and Lang 2019; Baier et al. 2022). For example, perhaps the strongest pattern of a Scarr–Rowe interaction in educational attainment has been found in Germany (Baier and Lang 2019), a context known for its high level of curriculum tracking. Baier et al. (2022) observed that Germany had comparatively low heritability of education, attributing it to the strong tracking in the German educational system (however, for mixed results, see Mönkediek 2022). Additionally, our results may provide (a partial) explanation on previous twin results in Finland finding lower heritability of education in earlier than later birth cohorts (Erola et al. 2022; Silventoinen et al. 2004, 2020).

These family background-related observations—an increase in the genetic prediction of education among those with low-educated families as well as the simultaneous disappearance of the Scarr–Rowe interaction—can be interpreted in terms of equality of opportunity. Many behavioral geneticists have advanced a thesis that higher heritability of education is an indicator of equality of opportunity (Ayorech et al. 2017; Conley 2016; Harden 2021; Plomin 2019; Selita and Kovas 2019; Silventoinen et al. 2020). Correspondingly, a typical interpretation of the Scarr–Rowe interaction is that those with advantaged origins have less environmental obstacles in realizing their individual potential. From both of these perspectives, the reform was successful in improving the equality of educational opportunity in Finland, which was also an explicit political goal of the reform (Aho, Pitkänen, and Sahlberg 2006; Kalalahti and Varjo 2020). However, a caveat here is that the greatest increase in genetic prediction may have been short term. The improvement in equality in educational achievement fits previous evidence on the increase in intergenerational income mobility (Pekkarinen, Uusitalo, and Kerr 2009) and cognitive test scores among those with less advantaged socioeconomic backgrounds (Pekkala Kerr, Pekkarinen, and Uusitalo 2013) as a result of this reform.

Because children inherit their genomes from their parents, genetic pathways produce a straightforward mechanism in the intergenerational transmission of traits. Thus, there is some tension between the increase in socioeconomic mobility after the reduction of tracking observed in previous studies (Pekkarinen, Uusitalo, and Kerr 2009; van de Werfhorst 2018; van de Werfhorst and Mijs 2010) and a simultaneous increase in the genetic prediction of education observed here. High heritability and low intergenerational persistence of education have been considered indicators of open societies and equality of opportunity. However, Engzell and Tropf (2019) found an inverse relationship between the heritability of education and the intergenerational educational correlation, suggesting they are not mutually exclusive views of equality of opportunity. Our results analogously suggest educational policies that reduce the intergenerational socioeconomic persistence have the potential to strengthen the genetic prediction of education in a population.

To answer the second research question concerning gender differences, we observed that the reform strengthened the genetic prediction of education among men but not women. Following Pekkarinen (2008), one potential mechanism for these gender-divergent findings lies in puberty. In the new system, the age of grades and decisions affecting future educational paths coincided with psychologically the most intense period of puberty among boys, whereas among adolescent girls,

this phase had started passing (Euling et al. 2008; Koivusilta and Rimpelä 2004; Pekkarinen 2008). Consistent with the Saunders or diathesis-stress models of gene–environment interactions (Baier et al. 2022; Boardman, Daw, and Freese 2013; Saunders 2010), puberty may create unstable or capricious circumstances where the effect of genetic endowments becomes more pronounced. Overall, the psychological regression at puberty is found to be stronger on average among adolescent boys than girls (Aalberg and Siimes 2007, 74–82), and previous evidence indicates that school performance specifically varies more strongly with the puberty timing among boys than girls (Koivusilta and Rimpelä 2004).

Finally, our third research question concerned whether reform effects vary according to the proximity of the reform. We observed that the genetic prediction of education was particularly strong for the first cohort experiencing the reform. Again, this interaction was driven by the population subgroups of men and those with lower-educated families. It is possible that the turbulence around the reform, with lots of practices still seeking their established form, resulted in a stronger differentiation on individuals by their genetic tendencies. Similar to the gender-specific results discussed in this section, this is another pattern roughly analogous to the Saunders or diathesis-stress interaction models.

## Methodological Considerations

Strengths of the study include population-representative data with high response rates for genetically informed datasets and register-based measurements, which can reduce biases stemming from self-reported data, including social desirability, subjectiveness, or faulty recall. Employing exogenous variation provided by a natural experiment helped minimize biases stemming from non-random allocation of genotypes across environments, which is a common limitation in sociogenomic research (Morris et al. 2020). The gradual implementation of the school reform across the country allowed us to control possible secular trends and regional differences. The Finnish data infrastructure and nature of the policy change provided us a rare opportunity to test the effects of limiting school tracking with these benefits, but as always in the case of single-country studies, the generalizability of the results should be tested in other contexts when a possibility arises.

Despite using exogenous environmental variation, the environmental and genetic components in gene–environment interactions may be confounded. This is because PGSs may themselves capture environmental or social factors, which are correlated with genetic variants (Kong et al. 2018; Morris et al. 2020). This could be due to population-level biasing factors (Freedman et al. 2004) or indirect genetic effects (also termed "genetic nurture" or "dynastic effects") whereby the non-inherited parental genotypes also associate with offspring education to inflate estimates of direct genetic effects (Kong et al. 2018; Morris, Davies, Hemani, et al. 2020; Wang et al. 2021). These cases would lead to correlations within the data whereby individuals with higher PGSs also having more educationally beneficial family environments (causally) independent of their inherited DNA variants. Previous work has demonstrated that such environmental biases are larger for educational attainment than many other traits (Howe et al. 2022), and therefore

it is possible that there remains some bias in the genetic component of our study. This limitation could be addressed using within-sibship analysis designs, but these require large samples of siblings that do not exist in the genotyped Finnish cohorts.

Although our main effects were estimated with high statistical precision, robust discovery of an interaction requires a much larger sample than for a corresponding main effect (Gelman, Hill, and Vehtari 2020, chap. 16).[9] Hence, the power of our study to estimate interactions may have been limited. Additionally, despite ever-increasing GWAS sample sizes, the PGSs still contain noise and account for only part of the total heritability of traits. GWAS summary statistics are usually meta-analyses of many cohorts in differing environmental circumstances and are likely to capture SNP effects that manifest uniformly across environments, which can make them conservative when analyzing gene–environment interactions (Conley 2016).

Factors stemming from the reform's implementation can bias our results to the conservative direction. First, as a compromise to the political struggle involved in the reform, ability groups for students were retained in foreign languages and mathematics until 1985 (Pekkala Kerr, Pekkarinen, and Uusitalo 2013). Attending the lowest group closed the doors for further education (Kalalahti and Varjo 2020; Kauko 2019). Second, teachers were only gradually educated focusing specifically on the new system (Aho, Pitkänen, and Sahlberg 2006, 49–51). Third, there were pilot schools that were exposed to the reform earlier than other schools in some municipalities, which we could not identify from our data. Additionally, delaying (or in very rare cases, advancing) the school start year of children or repeating a primary school grade are possible sources of noise in our measurement. Neither, however, were common during the period studied. In the 1960s, an annual 1.3–1.5 percent of the students in Helsinki had to repeat a grade in primary school (Oinonen 1969, 116). In Oinonen's (1969, 106–9) data of a large school district, 13 out of 430 students did not start their school together with their expected cohort in 1963 or 1964.

## Conclusion

Exploiting the natural experiment of the Finnish comprehensive school reform between 1972–1977, we found that a decrease in tracking in the schooling system led to a stronger genetic prediction of educational achievement, particularly among those with less socioeconomically advantaged family backgrounds. Thus, the reform was successful in achieving its central goal of increasing socioeconomic equality in the educational system. This is consistent with results studying the reform with regards of socioeconomic outcomes (Pekkala Kerr, Pekkarinen, and Uusitalo 2013; Pekkarinen 2008; Pekkarinen, Uusitalo, and Kerr 2009; Valkonen, Pensola, and Jalovaara 1996; 1998), expanding it to a novel sociogenetic perspective.

Because educational attainment and equality in it has well-known downstream effects on a wide variety of factors—relating to, for example, later socioeconomic achievement (Hout and DiPrete 2006), crime and deviance (Huttunen et al. 2023), societal integration (Wiborg 2009), political participation (Lindgren, Oskarsson, and Persson 2019), as well as health and mortality[10] (Davies et al. 2018; Lundborg, Lyttkens, and Nystedt 2016)—the reform provided further promise for increasing

equality of opportunity in them. In addition—assuming again that increasing the genetic prediction of education implies better realization of the potential of individuals—the results suggest the reform not only reduced inequality but simultaneously promoted the optimal allocation of human capital, with potentially wider macro-economic benefits.

These results also highlight the debate regarding whether individuals' educational outcomes could be improved by personalized curricula using their genetic data (Asbury and Plomin 2013; Sabatello 2018). Morris, Davies and Davey Smith (2020) demonstrated that PGSs are too coarse to be suitable for tailoring, particularly if predictors such as parental factors or prior achievement are available. Furthermore, our results suggest that more personalized curricula (particularly at early ages) can be counterproductive in optimizing the realization of capacities. Students' and their parents' preferences, test results, or even PGSs cannot be assumed to be independent from social forces, including socioeconomic resources that shape and lay constraints on them.

Finally, the results suggest that conditions marked by instabilities and pending major changes—whether in the form of puberty among adolescents or reform implementation at the institutional level—may be powerful magnifiers of genetic effects (see discussion on "dimmers" in Domingue et al. 2020). Future research could explore and provide a more systematic picture on the potential of turbulent conditions interacting with the genome in affecting later life success, health, and well-being. Although the genome as such is one of the most stable factors of an individual, its consequences need not be insensitive to instability in environmental conditions.

## Notes

[1] A fourth type of confounding is also possible. If we examined only a limited set of alleles in the analysis, the assessed genetic variant may not have necessarily been the causal variant, but only correlated with it. PGSs use the information across the whole genome, which alleviates this concern (Schmitz and Conley 2017), although with the cost of increasing ambiguity on the actual biological mechanisms involved.

[2] The suitability of the reform as a natural experiment has been extensively assessed in the work of Pekkarinen, Uusitalo, and Pekkala Kerr (Pekkarinen 2008; Pekkarinen, Uusitalo, and Kerr 2009; Pekkala Kerr, Pekkarinen, and Uusitalo 2013).

[3] With an exception that a few municipalities still had only a 2-year civic school. However, the reform changed the proportion of those in school at age 15 only negligibly (Pekkala Kerr, Pekkarinen, and Uusitalo 2013, 583).

[4] Originating from Saunders' (2010) observation that the advantaged families were able to protect their offspring from downward mobility even among their lower-ability offspring.

[5] Analogous results for a protective effect of socioeconomically advantaged schools have also been obtained using IQ instead of education PGS (Morris, Dorling, and Davey Smith 2016).

[6] At the quality-control stage, individuals with ambiguous gender information, less than 5 percent genotype missingness, ±4 standard deviations of heterozygosity, and non-Finnish ancestry were removed. Only genetic variants with a missingness of less than 2 percent and Hardy-Weinberg equilibrium p-value greater than $10^{-6}$ were retained.

7 Standard error (se) for the post-reform effect $b_1 + b_3$ was estimated with the delta method as $\sqrt{se_1^2 + se_3^2 + 2*covariance(b_1, b_3)}$ .

8 Standard error (se) for difference in coefficients (b) between educational backgrounds ($b_{highedu} + b_{lowedu}$) was estimated as $\sqrt{se_{highedu}^2 + se_{lowedu}^2}$ (Clogg, Petkova, and Haritou 1995, 1276).

9 Zheng et al. (2023) provided an excellent discussion and simulations regarding statistical power in the context of gene-cohort interactions (which are one type of gene-environment interactions).

10 Mixed evidence regarding tracking on health/mortality and their inequality has been found (Böckerman et al. 2021; Delaruelle, van de Werfhorst, and Bracke 2019; Ravesteijn et al. 2017).

# References

Aho, Erkki, Kari Pitkänen, and Pasi Sahlberg. 2006. "Policy Development and Reform Principles of Basic and Secondary Education in Finland since 1968." The Education Working Paper Series. Washington D.C.: the Education Unit at the World Bank (HDNED). https://policycommons.net/artifacts/1515541/policy-development-and-reform-principles-of-basic-and-secondary-education-in-finland-since-1968/2190989/.

Aromaa, Arpo, and Seppo Koskinen. 2004. *Health and Functional Capacity in Finland: Baseline Results of the Health 2000 Health Examination Survey*. Publications of the National Public Health Institute, KTL B12/2004. Helsinki: National Public Health Institute. URLhere

Asbury, Kathryn, and Robert Plomin. 2013. *G Is for Genes: The Impact of Genetics on Education and Achievement*. Chichester: John Wiley & Sons. https://doi.org/10.1002/9781118482766

Ayorech, Ziada, Eva Krapohl, Robert Plomin, and Sophie von Stumm. 2017. "Genetic Influence on Intergenerational Educational Attainment." *Psychological Science* 28 (9): 1302-10. https://doi.org/10.1177/0956797617707270

Baier, Tina, and Volker Lang. 2019. "The Social Stratification of Environmental and Genetic Influences on Education: New Evidence Using a Register-Based Twin Sample." *Sociological Science* 6: 143-71. https://doi.org/10.15195/v6.a6

Baier, Tina, Volker Lang, Michael Grätz, Kieron J. Barclay, Dalton C. Conley, Christopher T. Dawes, Thomas Laidley, and Torkild H. Lyngstad. 2022. "Genetic Influences on Educational Achievement in Cross-National Perspective." *European Sociological Review* 38 (6): 959-74. https://doi.org/10.1093/esr/jcac014

Björklund, Anders, and Markus Jäntti. 2020. "Intergenerational Mobility, Intergenerational Effects, Sibling Correlations, and Equality of Opportunity: A Comparison of Four Approaches." *Research in Social Stratification and Mobility* 70: 1-11. https://doi.org/10.1016/j.rssm.2019.100455

Boardman, Jason D., Jonathan Daw, and Jeremy Freese. 2013. "Defining the Environment in Gene-Environment Research: Lessons from Social Epidemiology." *American Journal of Public Health* 103 (S1): S64-72. https://doi.org/10.2105/AJPH.2013.301355

Böckerman, Petri, Mika Haapanen, Christopher Jepsen, and Alexandra Roulet. 2021. "School Tracking and Mental Health." *Journal of Human Capital* 15 (2): 291-345. https://doi.org/10.1086/712728

Borenstein, Michael, Larry V. Hedges, Julian PT Higgins, and Hannah R. Rothstein. 2009. *Introduction to Meta-Analysis*. Chichester: Wiley. https://doi.org/10.1002/97804707 43386

Borodulin, Katja, Hanna Tolonen, Pekka Jousilahti, Antti Jula, Anne Juolevi, Seppo Koskinen, Kari Kuulasmaa, Tiina Laatikainen, Satu Männistö, and Markku Peltonen. 2018. "Cohort Profile: The National FINRISK Study." *International Journal of Epidemiology* 47 (3): 696-696i. https://doi.org/10.1093/ije/dyx239

Branigan, Amelia R., Kenneth J. McCallum, and Jeremy Freese. 2013. "Variation in the Heritability of Educational Attainment: An International Meta-Analysis." *Social Forces* 92 (1): 109-40. https://doi.org/10.1093/sf/sot076

Briley, Daniel A., and Elliot M. Tucker-Drob. 2013. "Explaining the Increasing Heritability of Cognitive Ability across Development: A Meta-Analysis of Longitudinal Twin and Adoption Studies." *Psychological Science* 24 (9): 1704-13. https://doi.org/10.1177/09 56797613478618

Choi, Shing Wan, Timothy Shin-Heng Mak, and Paul F. O'Reilly. 2020. "Tutorial: A Guide to Performing Polygenic Risk Score Analyses." *Nature Protocols* 15 (9): 2759-72. https://doi.org/10.1038/s41596-020-0353-1

Clogg, Clifford C., Eva Petkova, and Adamantios Haritou. 1995. "Statistical Methods for Comparing Regression Coefficients between Models." *American Journal of Sociology* 100 (5): 1261-93. https://doi.org/10.1086/230638

Colodro-Conde, Lucía, Frühling Rijsdijk, María J. Tornero-Gómez, Juan F. Sánchez-Romera, and Juan R. Ordoñana. 2015. "Equality in Educational Policy and the Heritability of Educational Attainment." *PLoS One* 10 (11): e0143796. https://doi.org/10.1371/jour nal.pone.0143796

Conley, Dalton. 2016. "Socio-Genomic Research Using Genome-Wide Molecular Data." *Annual Review of Sociology* 42: 275-99. https://doi.org/10.1146/annurev-soc-08171 5-074316

Davies, Neil M., Matt Dickson, George Davey Smith, Gerard J. van den Berg, and Frank Windmeijer. 2018. "The Causal Effects of Education on Health Outcomes in the UK Biobank." *Nature Human Behaviour* 2 (2): 117. https://doi.org/10.1038/s41562-017 -0279-y

Delaruelle, Katrijn, Herman van de Werfhorst, and Piet Bracke. 2019. "Do Comprehensive School Reforms Impact the Health of Early School Leavers? Results of a Comparative Difference-in-Difference Design." *Social Science & Medicine* 239: 112542. https://doi.or g/10.1016/j.socscimed.2019.112542

Domingue, Benjamin W., Sam Trejo, Emma Armstrong-Carter, and Elliot M. Tucker-Drob. 2020. "Interactions between Polygenic Scores and Environments: Methodological and Conceptual Challenges." *Sociological Science* 7: 465-86. https://doi.org/10.15195/v7. a19

Engzell, Per, and Felix C. Tropf. 2019. "Heritability of Education Rises with Intergenerational Mobility." *Proceedings of the National Academy of Sciences* 116 (51): 25386-88. https://doi.org/10.1073/pnas.1912998116

Erola, Jani, Hannu Lehti, Tina Baier, and Aleksi Karhula. 2022. "Socioeconomic Background and Gene-Environment Interplay in Social Stratification across the Early Life Course." *European Sociological Review* 38 (1): 1-17. https://doi.org/10.1093/esr/jcab026

Euling, Susan Y., Marcia E. Herman-Giddens, Peter A. Lee, Sherry G. Selevan, Anders Juul, Thorkild I. A. SØrensen, Leo Dunkel, John H. Himes, Grete Teilmann, and Shanna H. Swan. 2008. "Examination of US Puberty-Timing Data from 1940 to 1994 for Secular

Trends: Panel Findings." *Pediatrics* 121 (Supplement_3): S172-91. https://doi.org/10.1542/peds.2007-1813D

Freedman, Matthew L., David Reich, Kathryn L. Penney, Gavin J. McDonald, Andre A. Mignault, Nick Patterson, Stacey B. Gabriel, et al. 2004. "Assessing the Impact of Population Stratification on Genetic Association Studies." *Nature Genetics* 36 (4): 388-93. https://doi.org/10.1038/ng1333

Freese, Jeremy, and Yu-Han Jao. 2017. "Shared Environment Estimates for Educational Attainment: A Puzzle and Possible Solutions." *Journal of Personality* 85 (1): 79-89. https://doi.org/10.1111/jopy.12226

Gelman, Andrew, Jennifer Hill, and Aki Vehtari. 2020. *Regression and Other Stories*. Cambridge University Press. https://doi.org/10.1017/9781139161879

Grätz, Michael, Kieron J. Barclay, Øyvind N. Wiborg, Torkild H. Lyngstad, Aleksi Karhula, Jani Erola, Patrick Präg, Thomas Laidley, and Dalton Conley. 2021. "Sibling Similarity in Education across and within Societies." *Demography* 58 (3): 1011-37. https://doi.org/10.1215/00703370-9164021

Harden, Kathryn Paige. 2021. *The Genetic Lottery: Why DNA Matters for Social Equality*. Princeton: University Press. https://doi.org/10.1515/9780691226705

Harden, Kathryn Paige, Benjamin W. Domingue, Daniel W. Belsky, Jason D. Boardman, Robert Crosnoe, Margherita Malanchini, Michel Nivard, Elliot M. Tucker-Drob, and Kathleen Mullan Harris. 2020. "Genetic Associations with Mathematics Tracking and Persistence in Secondary School." *NPJ Science of Learning* 5 (1): 1-8. https://doi.org/10.1038/s41539-020-0060-2

Härkönen, Juho, and Outi Sirniö. 2020. "Educational Transitions and Educational Inequality: A Multiple Pathways Sequential Logit Model Analysis of Finnish Birth Cohorts 1960-1985." *European Sociological Review* 36 (5): 700-719. https://doi.org/10.1093/esr/jcaa019

Haworth, Catherine MA, Margaret J. Wright, Michelle Luciano, Nicholas G. Martin, Eco JC de Geus, Catharina EM van Beijsterveldt, Meike Bartels, Danielle Posthuma, Dorret I. Boomsma, and Oliver SP Davis. 2010. "The Heritability of General Cognitive Ability Increases Linearly from Childhood to Young Adulthood." *Molecular Psychiatry* 15 (11): 1112-20. https://doi.org/10.1038/mp.2009.55

Heath, A. C., K. Berg, L. J. Eaves, M. H. Solaas, L. A. Corey, J. Sundet, P. Magnus, and W. E. Nance. 1985. "Education Policy and the Heritability of Educational Attainment." *Nature* 314 (6013): 734-36. https://doi.org/10.1038/314734a0

Herd, Pamela, Jeremy Freese, Kamil Sicinski, Benjamin W. Domingue, Kathleen Mullan Harris, Caiping Wei, and Robert M. Hauser. 2019. "Genes, Gender Inequality, and Educational Attainment." *American Sociological Review* 84 (6): 1069-98. https://doi.org/10.1177/0003122419886550

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina. 2008. "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends." *The B.E. Journal of Economic Analysis & Policy* 7 (2): 1-46. https://doi.org/10.2202/1935-1682.1775

Hout, Michael, and Thomas A. DiPrete. 2006. "What We Have Learned: RC28's Contributions to Knowledge about Social Stratification." *Research in Social Stratification and Mobility* 24 (1): 1-20. https://doi.org/10.1016/j.rssm.2005.10.001

Howe, Laurence J., Michel G. Nivard, Tim T. Morris, Ailin F. Hansen, Humaira Rasheed, Yoonsu Cho, Geetha Chittoor, Rafael Ahlskog, Penelope A. Lind, and Teemu Palviainen. 2022. "Within-Sibship Genome-Wide Association Analyses Decrease Bias in Estimates of

Direct Genetic Effects." *Nature Genetics* 54 (5): 581-92. https://doi.org/10.1038/s415 88-022-01062-7

Huttunen, Kristiina, Tuomas Pekkarinen, Roope Uusitalo, and Hanna Virtanen. 2023. "Lost Boys? Secondary Education and Crime." *Journal of Public Economics* 218: 104804. https: //doi.org/10.1016/j.jpubeco.2022.104804

Kalalahti, Mira, and Janne Varjo. 2020. "Revisiting Universalism in the Finnish Education System." *Research on Finnish Society* 12 (1): 1-16.

Karhunen, Hannu, and Roope Uusitalo. 2017. "50 vuotta koulutusmahdollisuuksien tasa-arvoa." *Yhteiskuntapolitiikka* 82 (3): 296-303.

Kauko, Jaakko. 2019. "The Finnish Comprehensive School: Conflicts, Compromises, and Institutional Robustness." In *Great Policy Successes: Or, A Tale About Why It's Amazing That Governments Get So Little Credit for Their Many Everyday and Extraordinary Achievements as Told by Sympathetic Observers Who Seek to Create Space for a Less Relentlessly Negative View of Our Pivotal Public Institutions.*, by Mallory Compton and Paul Hart, 122-42. Oxford: University Press.

Kerminen, Sini, Aki S. Havulinna, Garrett Hellenthal, Alicia R. Martin, Antti-Pekka Sarin, Markus Perola, Aarno Palotie, Veikko Salomaa, Mark J. Daly, and Samuli Ripatti. 2017. "Fine-Scale Genetic Structure in Finland." *G3: Genes, Genomes, Genetics* 7 (10): 3459-68. https://doi.org/10.1534/g3.117.300217

Kivirauma, Joel, and Kari Ruoho. 2007. "Excellence through Special Education? Lessons from the Finnish School Reform." *International Review of Education* 53 (3): 283-302. https: //doi.org/10.1007/s11159-007-9044-1

Knigge, Antonie, Ineke Maas, Kim Stienstra, Eveline L. De Zeeuw, and Dorret I. Boomsma. 2022. "Delayed Tracking and Inequality of Opportunity: Gene-Environment Interactions in Educational Attainment." *NPJ Science of Learning* 7 (1): 1-13. https://doi.org/10.103 8/s41539-022-00122-1

Koivusilta, Leena, and A. Rimpelä. 2004. "Pubertal Timing and Educational Careers: A Longitudinal Study." *Annals of Human Biology* 31 (4): 446-65. https://doi.org/10.1080/ 03014460412331281719

Kong, Augustine, Gudmar Thorleifsson, Michael L. Frigge, Bjarni J. Vilhjalmsson, Alexander I. Young, Thorgeir E. Thorgeirsson, Stefania Benonisdottir, Asmundur Oddsson, Bjarni V. Halldorsson, and Gisli Masson. 2018. "The Nature of Nurture: Effects of Parental Genotypes." *Science* 359 (6374): 424-28. https://doi.org/10.1126/science.aan6877

Lahtinen, Hannu, Kaarina Korhonen, Pekka Martikainen, and Tim Morris. 2023. "Polygenic Prediction of Education and Its Role in the Intergenerational Transmission of Education: Cohort Changes Among Finnish Men and Women Born in 1925-1989." *Demography* 60 (5): 1523-47. https://doi.org/10.1215/00703370-10963788

Lahtinen, Hannu, Pekka Martikainen, and Lasse Tarkiainen. 2022. "Changes in Sibling Similarity in Education among Finnish Cohorts Born in 1950-89: The Contribution of Paternal and Maternal Education." *Longitudinal and Life Course Studies* 13 (4): 496-526. https://doi.org/10.1332/175795921X16600597647058

Lazear, Edward P. 2001. "Educational Production." *The Quarterly Journal of Economics* 116 (3): 777-803. https://doi.org/10.1162/00335530152466232

Lee, James J., Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghzian, Meghan Zacher, Tuan Anh Nguyen-Viet, et al. 2018. "Gene Discovery and Polygenic Prediction from a Genome-Wide Association Study of Educational Attainment in 1.1 Million Individuals." *Nature Genetics* 50 (8): 1112-21. https://doi.org/10.1038/s41588-018-0147-3

Lin, Meng-Jung. 2020. "The Social and Genetic Inheritance of Educational Attainment: Genes, Parental Education, and Educational Expansion." *Social Science Research* 86: 102387. https://doi.org/10.1016/j.ssresearch.2019.102387

Lindgren, Karl-Oskar, Sven Oskarsson, and Mikael Persson. 2019. "Enhancing Electoral Equality: Can Education Compensate for Family Background Differences in Voting Participation?" *American Political Science Review* 113 (1): 108-22. https://doi.org/10.1017/S0003055418000746

Lloyd-Jones, Luke R., Jian Zeng, Julia Sidorenko, Loïc Yengo, Gerhard Moser, Kathryn E. Kemper, Huanwei Wang, Zhili Zheng, Reedik Magi, and Tonu Esko. 2019. "Improved Polygenic Prediction by Bayesian Multiple Regression on Summary Statistics." *Nature Communications* 10 (1): 1-11. https://doi.org/10.1038/s41467-019-12653-0

Lundborg, Petter, Carl Hampus Lyttkens, and Paul Nystedt. 2016. "The Effect of Schooling on Mortality: New Evidence from 50,000 Swedish Twins." *Demography* 53 (4): 1135-68. https://doi.org/10.1007/s13524-016-0489-3

Malanchini, Margherita, Kaili Rimfeld, Andrea G. Allegrini, Stuart J. Ritchie, and Robert Plomin. 2020. "Cognitive Ability and Education: How Behavioural Genetic Research Has Advanced Our Knowledge and Understanding of Their Association." *Neuroscience & Biobehavioral Reviews* 111: 229-45. https://doi.org/10.1016/j.neubiorev.2020.01.016

Mansournia, Mohammad Ali, Maryam Nazemipour, Ashley I Naimi, Gary S Collins, and Michael J Campbell. 2021. "Reflection on Modern Methods: Demystifying Robust Standard Errors for Epidemiologists." *International Journal of Epidemiology* 50 (1): 346-51. https://doi.org/10.1093/ije/dyaa260

Manuck, Stephen B., and Jeanne M. McCaffery. 2014. "Gene-Environment Interaction." *Annual Review of Psychology* 65: 41-70. https://doi.org/10.1146/annurev-psych-010213-115100

Mills, Melinda C., Nicola Barban, and Felix C. Tropf. 2020. *An Introduction to Statistical Genetic Data Analysis*. Cambridge: MIT Press.

Mills, Melinda C., and Felix C. Tropf. 2020. "Sociology, Genetics, and the Coming of Age of Sociogenomics." *Annual Review of Sociology* 46 (1): 553-81. https://doi.org/10.1146/annurev-soc-121919-054756

Mollborn, Stefanie, Elizabeth Lawrence, Laurie James-Hawkins, and Paula Fomby. 2014. "When Do Socioeconomic Resources Matter Most in Early Childhood?" *Advances in Life Course Research* 20: 56-69. https://doi.org/10.1016/j.alcr.2014.03.001

Mönkediek, Bastian. 2022. "How Variants of Tracking Affect the Role of Genes and Environment in Explaining Child Attendance at Upper Secondary School." *Research in Social Stratification and Mobility* 81: 100714. https://doi.org/10.1016/j.rssm.2022.100714

Morris, Tim T., Neil M. Davies, and George Davey Smith. 2020. "Can Education Be Personalised Using Pupils' Genetic Data?" *Elife* 9: e49962. https://doi.org/10.7554/eLife.49962

Morris, Tim T., Neil M. Davies, Gibran Hemani, and George Davey Smith. 2020. "Population Phenomena Inflate Genetic Associations of Complex Social Traits." *Science Advances* 6 (16): eaay0328. https://doi.org/10.1126/sciadv.aay0328

Morris, Tim T., Danny Dorling, and George Davey Smith. 2016. "How Well Can We Predict Educational Outcomes? Examining the Roles of Cognitive Ability and Social Position in Educational Attainment." *Contemporary Social Science* 11 (2-3): 154-68. https://doi.org/10.1080/21582041.2016.1138502

Neale, Zoe E., I. Sally, Chun Kuo, and Danielle M. Dick. 2021. "A Systematic Review of Gene-by-Intervention Studies of Alcohol and Other Substance Use." *Development and Psychopathology* 33 (4): 1410-27. https://doi.org/10.1017/S0954579420000590

Oinonen, Päivi. 1969. *Kouluvalmiuden Ongelma: Milloin Lapsi on Kypsä Kouluun*. Porvoo: WSOY.

Okbay, Aysu, Yeda Wu, Nancy Wang, Hariharan Jayashankar, Michael Bennett, Seyed Moeen Nehzati, Julia Sidorenko, Hyeokmoon Kweon, Grant Goldman, and Tamara Gjorgjieva. 2022. "Polygenic Prediction of Educational Attainment within and between Families from Genome-Wide Association Analyses in 3 Million Individuals." *Nature Genetics* 54 (4): 437-49. https://doi.org/10.1038/s41588-022-01016-z

Okkonen, V. 2017. "Peruskoulua Vastaan: Yksityisoppikoulut Ja Yhteiskuntajärjestys 1966-1975 [Against the Comprehensive School: Private Grammar Schools and Societal Order]." *Scripta Lingua Fennica Edita C* 444.

Ollikainen, Jani-Petteri. 2021. "Comprehensive School Reform and Labor Market Outcomes over the Lifecycle: Evidence from Finland." *Labour Economics* 68 (January): 101952. https://doi.org/10.1016/j.labeco.2020.101952

Pärn, Kalle, Javier Nunez Fontarnau, Marita A. Isokallio, Timo Sipilä, Elina Kilpelainen, Aarno Palotie, Samuli Ripatti, and Priit Palta. 2018. "Genotyping Chip Data Lift-over to Reference Genome Build GRCh38/Hg38." Protocols.io. https://doi.org/10.17504/protocols.io.nqtddwn

Pärn, Kalle, Marita A. Isokallio, Javier Nunez-Fontarnau, Aarno Palotie, Samuli Ripatti, and Priit Palta. 2018. "Genotype Imputation Workflow v3.0." Protocols.Io. https://doi.org/10.17504/protocols.io.nmndc5e

Pasman, Joëlle A., Karin JH Verweij, and Jacqueline M. Vink. 2019. "Systematic Review of Polygenic Gene-Environment Interaction in Tobacco, Alcohol, and Cannabis Use." *Behavior Genetics* 49 (4): 349-65. https://doi.org/10.1007/s10519-019-09958-7

Pekkala Kerr, Sari, Tuomas Pekkarinen, and Roope Uusitalo. 2013. "School Tracking and Development of Cognitive Skills." *Journal of Labor Economics* 31 (3): 577-602. https://doi.org/10.1086/669493

Pekkarinen, Tuomas. 2008. "Gender Differences in Educational Attainment: Evidence on the Role of Tracking from a Finnish Quasi-Experiment." *Scandinavian Journal of Economics* 110 (4): 807-25. https://doi.org/10.1111/j.1467-9442.2008.00562.x

Pekkarinen, Tuomas. 2012. "Gender Differences in Education." *Nordic Economic Policy Review* 1 (1): 165-94. https://doi.org/10.2139/ssrn.2019442

Pekkarinen, Tuomas, Roope Uusitalo, and Sari Kerr. 2009. "School Tracking and Intergenerational Income Mobility: Evidence from the Finnish Comprehensive School Reform." *Journal of Public Economics* 93 (7-8): 965-73. https://doi.org/10.1016/j.jpubeco.2009.04.006

Pfeffer, Fabian T. 2008. "Persistent Inequality in Educational Attainment and Its Institutional Context." *European Sociological Review* 24 (5): 543-65. https://doi.org/10.1093/esr/jcn026

Plomin, Robert. 2019. *Blueprint: How DNA Makes Us Who We Are*. Cambridge: MIT Press.

Polderman, Tinca JC, Beben Benyamin, Christiaan A. De Leeuw, Patrick F. Sullivan, Arjen Van Bochoven, Peter M. Visscher, and Danielle Posthuma. 2015. "Meta-Analysis of the Heritability of Human Traits Based on Fifty Years of Twin Studies." *Nature Genetics* 47 (7): 702-9. https://doi.org/10.1038/ng.3285

Poropat, Arthur E. 2009. "A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance." *Psychological Bulletin* 135 (2): 322. https://doi.org/10.1037/a0014996

Price, Alkes L., Nick J. Patterson, Robert M. Plenge, Michael E. Weinblatt, Nancy A. Shadick, and David Reich. 2006. "Principal Components Analysis Corrects for Stratification in Genome-Wide Association Studies." *Nature Genetics* 38 (8): 904-9. https://doi.org/10.1038/ng1847

Ravesteijn, Bastian, Hans van Kippersluis, Mauricio Avendano, Pekka Martikainen, Hannu Vessari, and Eddy Van Doorslaer. 2017. "The Impact of Later Tracking on Mortality by Parental Income in Finland," Tinbergen Institute Discussion Paper. https://doi.org/10.2139/ssrn.2924537

Rowe, David C., Kristen C. Jacobson, and Edwin JCG Van den Oord. 1999. "Genetic and Environmental Influences on Vocabulary IQ: Parental Education Level as Moderator." *Child Development* 70 (5): 1151-62. https://doi.org/10.1111/1467-8624.00084

Sabatello, Maya. 2018. "A Genomically Informed Education System? Challenges for Behavioral Genetics." *Journal of Law, Medicine & Ethics* 46 (1): 130-44. https://doi.org/10.1177/1073110518766027

Saunders, Peter. 2010. *Social Mobility Myths*. London: Civitas.

Scarr-Salapatek, Sandra. 1971. "Race, Social Class, and IQ: Population Differences in Heritability of IQ Scores Were Found for Racial and Social Class Groups." *Science* 174 (4016): 1285-95. https://doi.org/10.1126/science.174.4016.1285

Schmitz, Lauren, and Dalton Conley. 2017. "Modeling Gene-Environment Interactions with Quasi-Natural Experiments." *Journal of Personality* 85 (1): 10-21. https://doi.org/10.1111/jopy.12227

Selita, Fatos, and Yulia Kovas. 2019. "Genes and Gini: What Inequality Means for Heritability." *Journal of Biosocial Science* 51 (1): 18-47. https://doi.org/10.1017/S0021932017000645

Silventoinen, Karri, Aline Jelenkovic, Reijo Sund, Antti Latvala, Chika Honda, Fujio Inui, Rie Tomizawa, Mikio Watanabe, Norio Sakai, and Esther Rebato. 2020. "Genetic and Environmental Variation in Educational Attainment: An Individual-Based Analysis of 28 Twin Cohorts." *Scientific Reports* 10 (1): 1-11. https://doi.org/10.1038/s41598-020-69526-6

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 (5): 544-55. https://doi.org/10.1002/ajhb.20060. https://doi.org/10.1002/ajhb.20060

Sysiharju, Anna-Liisa. 1969. "Social Factors and Recruitment to the Gymnasium in Finland." *Pedagogisk Forskning* 13 (1): 79-102. https://doi.org/10.1080/0031383690130104

Trejo, Sam, Daniel W. Belsky, Jason D. Boardman, Jeremy Freese, Kathleen Mullan Harris, Pam Herd, Kamil Sicinski, and Benjamin W. Domingue. 2018. "Schools as Moderators of Genetic Associations with Life Course Attainments: Evidence from the WLS and Add Health." *Sociological Science* 5 (August): 513-40. https://doi.org/10.15195/v5.a22

Tucker-Drob, Elliot M., and Timothy C. Bates. 2016. "Large Cross-National Differences in GeneX Socioeconomic Status Interaction on Intelligence." *Psychological Science* 27 (2): 138-49. https://doi.org/10.1177/0956797615612727

Turkheimer, Eric. 2000. "Three Laws of Behavior Genetics and What They Mean." *Current Directions in Psychological Science* 9 (5): 160-64. https://doi.org/10.1111/1467-8721.00084

UNESCO. 2012. *International Standard Classification of Education: ISCED 2011.* UNESCO Institute for Statistics.

Valkonen, Tapani, Tiina Pensola, and Marika Jalovaara. 1996. "Peruskoulu-Uudistuksen Vaikutus Sukupuolten, Alueiden Ja Sosiaaliryhmien Välisiin Eroihin Perusasteen Jälkeisessä Koulutuksessa." In *Toteuttaako Peruskoulu Tasa-Arvoa?*, by Ritva Jakku-Sihvonen, Aslak Lindström, and Sinikka Lipsanen, 316-35. 1/96. Helsinki: Opetushallitus.

Valkonen, Tapani, Tiina Pensola, and Marika Jalovaara. 1998. "Peruskoulu-Uudistuksen Vaikutus Sosiaaliryhmien Välisiin Eroihin Ylioppilastutkinnon Suorittamisessa." In *Rekisteriaineistot Yhteiskunta- Ja Sosiaalitutkimuksessa.*, by Tapani Valkonen, Seppo Koskinen, and Tuija Martelin, 229-40. Helsinki: Gaudeamus.

Wang, Biyao, Jessie R. Baldwin, Tabea Schoeler, Rosa Cheesman, Wikus Barkhuizen, Frank Dudbridge, David Bann, Tim T. Morris, and Jean-Baptiste Pingault. 2021. "Robust Genetic Nurture Effects on Education: A Systematic Review and Meta-Analysis Based on 38,654 Families across 8 Cohorts." *The American Journal of Human Genetics* 108 (9): 1780-91. https://doi.org/10.1016/j.ajhg.2021.07.010

Werfhorst, Herman G. van de. 2018. "Early Tracking and Socioeconomic Inequality in Academic Achievement: Studying Reforms in Nine Countries." *Research in Social Stratification and Mobility* 58 (December): 22-32. https://doi.org/10.1016/j.rssm.2018.09.002

Werfhorst, Herman G. van de, and Jonathan JB Mijs. 2010. "Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective." *Annual Review of Sociology* 36: 407-28. https://doi.org/10.1146/annurev.soc.012809.102538

Wiborg, Susanne. 2009. *Education and Social Integration: Comprehensive Schooling in Europe.* New York: Palgrave Macmillan.

Zeng, Jian, Jian Yang, Futao Zhang, Zhili Zheng, Luke Lloyd-Jones, and Mike Goddard. 2020. "GCTB: A Tool for Genome-Wide Complex Trait Bayesian Analysis." https://cnsgenomics.com/software/gctb/.

Zheng, Boyan, Jason Fletcher, Jie Song, and Qiongshi Lu. 2023. "Analysis of Sex-Specific Gene-by-Cohort and Genetic Correlation-by-Cohort Interaction in Educational and Reproductive Outcomes Using the UK Biobank Data." *Journal of Health and Social Behavior*, August, 00221465231188166. https://doi.org/10.1177/00221465231188166

**Hannu Lahtinen:** Population Research Unit, University of Helsinki; Max Planck – University of Helsinki Center for Social Inequalities in Population Health, Helsinki, Finland. E-mail: hannu.lahtinen@helsinki.fi.

**Pekka Martikainen:** Population Research Unit, University of Helsinki; Max Planck Institute for Demographic Research, Rostock, Germany; Max Planck – University of Helsinki Center for Social Inequalities in Population Health, Helsinki, Finland. E-mail: pekka.martikainen@helsinki.fi.

**Kaarina Korhonen:** Population Research Unit, University of Helsinki; Max Planck – University of Helsinki Center for Social Inequalities in Population Health, Helsinki, Finland. E-mail: kaarina.korhonen@helsinki.fi.

**Tim Morris:** Centre for Longitudinal Studies, Social Research Institute, University College London. E-mail: t.t.morris@ucl.ac.uk.

**Mikko Myrskylä:** Max Planck Institute for Demographic Research, Rostock, Germany; University of Helsinki, Helsinki, Finland; Max Planck – University of Helsinki Center for Social Inequalities in Population Health, Rostock, Germany and Helsinki, Finland. E-mail: myrskyla@demogr.mpg.de.