Does parental education influence child educational outcomes? A developmental analysis in a genetically-informative full population sample and adoptee design

Steven G. Ludeke*[1]
Miriam Gensowski*[2]
Sarah Y. Junge[3]
Oliver P. John[4]
Simon C. Andersen[3]

Abstract:

Children's educational outcomes are strongly correlated with their parents' educational attainment. This finding is often attributed to the family environment (e.g., differential parental behavior and resources driving child attainment). However, inferring a causal role of the family environment depends on the untested assumption that intergenerational similarities do *not* reflect genes shared between parent and child. We provide data on this assumption with an adoption design in full-population cohorts from Danish administrative data. We test whether parental education predicts children's educational outcomes in both biological and adopted children, at four developmental stages: (a) the child's Conscientiousness during compulsory schooling (grades 4-9), (b) academic performance in those same years, measured by objective achievement tests, (c) enrollment in academically challenging high schools, and (d) graduation success. At all four stages, parental education was a substantial predictor of child outcomes in the full population, but has a much smaller role (if any) in the adoptee subsamples. This suggests that the intergenerational correlation in education is strongly dependent on shared genes between parent and child. Further analysis of the different stages shows that while adoptive parents' education has a significant effect on later stages, such as educational attainment and enrollment, these effects are purely due to effects in earlier stages. When controlling for academic performance, for example, adoptive parents' education becomes insignificant in predicting enrollment in higher education. These findings highlight that interventions to improve educational outcomes for children from disadvantaged families may have the most profound impact when they reach families early in the child's educational development.

Keywords:

Educational outcomes; full-population studies; behavior genetics; adoptees; personality; intergenerational association of education

*Denotes equally contributing co-authors
[1] Department of Psychology, University of Southern Denmark, Odense, Denmark
[2] Department of Economics, University of Copenhagen, Copenhagen, Denmark
[3] Department of Political Science, Aarhus University, Aarhus, Denmark
[4] Department of Psychology, University of California, Berkeley, USA

Corresponding author:
Steven Ludeke, Campusvej 55, 5230, Odense M, Denmark, Lud@health.sdu.dk , tel. +45 2031 5064

- please turn over - page intentionally left blank -

The transmission of social and economic advantage across generations is a perennial topic of interest for both the public and the scientific community. Research on the topic often focuses specifically on education as a pathway through which families reproduce their advantage. The associations between parental and child levels of education are well documented, as are the effects of education on class.

Precisely *how* education is transmitted across generations is less clear. Many studies address the question by showing how children's educational outcomes are correlated with a long list of behaviors that are typical of more educated parents (e.g., using a rich vocabulary, providing assistance with school work, or encouraging the pursuit of advanced education). However, drawing causal inferences from such studies is problematic, most significantly because correlations between the characteristics of parents and of children is so often attributable not to the causal influence of the former on the latter but because of genes common to both (1).

Behavioral genetic study designs represent a useful alternative approach. They rely on natural experiments, such as twinning and adoption (2), in which the degrees of genetic relatedness between family members deviate systematically from that in the typical family. This setting allows for potent tests of parental influence on child characteristics. For example, if the association between parent and child characteristics is present primarily or even exclusively in parents and children who share genes, such studies suggest that genes shared between the parents and child represent the more parsimonious explanation for the association, a conclusion that can be further buffered when sibling or twin analyses also indicate little or no role for the family rearing environment.

Even richer behavioral genetic studies are possible when information on multiple stages of a hypothesized developmental pathway are available. Consider a particularly noteworthy recent study by McGue and colleagues (3): building off of previous work which found that educational attainment in adoptive parents predict the adoptive child's IQ (4, 5) as well as separate research that found a similar effect on the adoptive child's level of educational attainment (6), the McGue study combined these two elements to find that similarity in educational attainment between parents and their adoptive children persisted even after controlling for the earlier effects of parental attainment on adoptive child's IQ and personality. While the authors could not identify a specific mechanism, they advanced hypotheses, such as "academic expectations, social network effects, and the economic benefits of having wealthy parents," for how parental education could affect child education other than through IQ. Yet between early parental influences on the child's IQ, and the

child's eventual educational attainment, there lies a great deal of time and important educational development. Ideally, research can break down this period and pinpoint specific points at which parental education influences the different stages of the pathway from early learning to educational attainment. Behavioral genetic methods can only identify periods of plausibly causal effects if the available data is rich enough and covers different stages of the pathway from early influences to attainment. In the present study, we analyze the linkage between parental attainment and four separate stages in their children's educational development using an unparalleled data resource: administrative registers on a Danish national sample, allowing a comparison of results in the full population against those in adoptees reared by nonfamily members. We use this to critically examine not only the existence, but also the timing, of any causal effects of parental attainment on child attainment.

Pathways of influence

Figure 1 illustrates a model for how parental educational attainment might causally influence child educational attainment. The figure shows that at least four aspects of the child's educational trajectories can be differentiated where parents could influence their children educational achievement: (I) the child's educationally-relevant psychological characteristics (e.g. personality, intelligence); (II) the child's academic performance throughout primary and compulsory schooling; (III) the child's enrollment in advanced education; and, (IV) the child's completion of advanced education, or attainment. Each of these characteristics plausibly affects those listed later, for example if the child's personality (part of category I) impacts their learning and thus their performance on exams (II). Further, each of the characteristics is plausibly affected by the parent's level of education through any of a number of mechanisms.

As noted above, previous research has convincingly demonstrated a non-genetically confounded relationship between parental attainment and both the first and last of these stages. However, we are not aware of any published work that controls for genetic confounding when exploring linkages between parental attainment and child academic performance (II) or child enrollment in advanced education (III). Each of these characteristics could depend on parental attainment either through effects on earlier stages in the child's educational development (e.g., affecting grades by first affecting intelligence or personality), or through direct associations. For example, high attainment parents might improve the

academic performance of their children through the hiring of private tutors (7) or by shaping the child's behavior through parental expectations for education (8, 9).

More importantly still, no previous work has integrated either the child's academic performance or enrollment in advanced education into a genetically-informative study in order to determine the *timing* of when any causal effect of parental attainment on child educational characteristics might occur. Understanding the timing of causal effects will provide a gauge of the plausibility of various hypothesized mechanisms by which parental attainment is suggested to influence child attainment, due to the fact that the mechanisms often correspond to specific stages of the process, and their effects, if present, should thus show up even after conditioning on prior stages in the child's educational development. For example, some mechanisms plausibly imply that highly educated parents may increase their child's *attaining* advanced education even after conditioning on their child's *enrollment* in that education – for example, by using their own personal familiarity with advanced education to help the child succeed in that educational context, or by providing a financial buffer that would allow the child to continue their education despite adverse financial events. Other mechanisms might involve parental attainment causing the child to enroll in advanced education over and above any effect on the child's prior academic performance, such as by using their personal familiarity with advanced education or their social network to facilitate their child's admission to advanced education, or even simply by convincing the child that advanced education is so important that it should be pursued even when one's prior academic performance makes success in that education appear less than certain.

These mechanisms each imply that the specific effect of parental attainment on child attainment *persists* after not only eliminating genetic confounding but also controlling for effects on the child's education in earlier stages. By comparing the relationships at different stages before and after adequate controls are introduced, we can test whether specific mechanisms are plausible pathways through which parental attainment causally influenced child attainment.

The present study

We use data collected by the Danish government for administrative purposes to explore the associations between parental levels of education and the four child educational characteristics described above. To assess academically relevant psychological characteristics (I), we use the child's level of Conscientiousness, a personality trait highly relevant for academic performance (10). Academic performance (II) is assessed via standardized national tests and exit exams. Enrollment in advanced education (III) is assessed using enrollment in

the most academically-rigorous post-elementary educational track (high school, also referred to as gymnasium), which is pursued by the most gifted two-thirds of Danish students and typically begin in the year they turn 17, as described below. Completion (IV) involves completion of high school, which is itself a rigorous and selective accomplishment: in our high school completion sample described below, only 85% of students enrolling in this three-year degree in 2010 will have completed it by 2016 (notes in Appendix A3). For reasons discussed below, the analyses using Conscientiousness are limited to public school students (80% of the population in the relevant cohort), but all other analyses use all Danish residents within the relevant age ranges.

This data provides a number of advantages over previous studies. First, the use of administrative data provides advantages with respect to (a) statistical power (we have far more participants than any previous study on the topic), (b) data accuracy (for example, parental education is not vulnerable to inaccurate self- or child-reporting), and (c) generalizability (as no analyses are affected by things such as volunteer bias, and most analyses use the entire population of a given age cohort). A second class of advantages pertain to our Conscientiousness data: although the measure is very brief, it (d) captures precisely the domain of greatest relevance for the topic in question, and (e) does so at the ages that are not only most relevant for the topic (i.e. before one has completed one's education) but are also when parental characteristics have a greater opportunity to affect the child – that is, when the child is comparatively young and still living in the family home (11). The final strengths are perhaps the most important: not only do we (g) remedy the previous lack of genetically-informative studies of parental attainment on child academic performance and enrollment in advanced education through our use of adoptee sub-samples, but (h) in simultaneously analyzing such a range of components of childhood education, we provide a greater ability to highlight any stages at which a specific causal role for parental attainment on a given child characteristic are less plausible.

Materials and Methods

Because the administrative data we analyze was collected by the Danish government for purposes unrelated to the present study, many elements of the data are available only for a specific subset of the population. As data availability determines the sample used for a given analysis, we begin by describing the measures used for each characteristic before describing the samples. More extensive discussions of measures and samples, both with details to

facilitate replication and an extended discussion of how data availability shaped the parameters for a given sample, are provided in Appendix A1.

Measures

Conscientiousness

Students enrolled in Danish public schools completed an annual assessment on well-being beginning in 2015, with data currently available through 2017. Three items completed by 4th to 9th graders were judged by the authors to reflect Conscientiousness as reflected in the Big Five - particularly the agentic/industrious rather than orderly components of the trait. Results from a supplementary study described in Appendix A1 indicate that this measure not only correlates highly with a general measure Big Five Conscientiousness (12) (r = .65), but that it is particularly characterized by the Conscientiousness facets of Self-Discipline, Self-Efficacy, and Achievement. Previous research suggests that a Conscientiousness measure with this facet profile should be a comparatively effective predictor of academic performance (13).

For each year's assessment, we generate a sum score for each participant and then standardize the score (mean 0 and SD 1). We then average across the participant's score from each year. (For some participants, e.g. those too young to complete the first or second year's assessment, the average will involve scores from only one or two years.) This average score is then standardized again. Internal consistency was high at all grade levels (alphas ranged from .68 to .70), and test-rest correlations are substantial (e.g. r = .55 between 2016 and 2017).

Academic performance

Danish national tests

Since 2010, Danish public school students in 2nd, 4th, 6th, and 8th grade are required to take a test of their reading ability. Exams take place near the end of the school year and are computerized adaptive tests in which questions are determined by the student's performance earlier in the test. The test is scored electronically without teacher input, such that the system automatically calculates scores in three performance areas: language comprehension, decoding, and reading comprehension. Following (14) and (17), we standardize these three individual scores, take the simple average, and re-standardize them within year. Then, we form the average of the available scores between 2014 and 2017 for each individual.

Exit exams

A comprehensive set of exit exams are completed by all Danish students (not just those in public school) at the end of 9th grade. Since 2007 these exams have included a stable set of exam topics (written Danish, oral Danish, reading, spelling, problem-solving (math), skills (math), oral science (physics, chemistry, biology and geography), and oral English). Following our procedure for the Danish national tests, each of these exams scores are standardized and a simple mean is calculated and re-standardized.

Higher educational enrollment and attainment

In the Danish educational system, the first opportunity an individual has to select into rigorous advanced education comes after the completion of 9th grade, at which point one can enter high school. Only 71% of those who were in 9th grade (either public or private) in the 2009/10 academic year will have enrolled in this advanced educational track by 2016.

Because of the young age of the children (discussed below), we use high school for both our advanced educational enrollment and attainment measures. Of course, a continuous measure of educational attainment is generally preferable, and the older age of the parent generation allows us to use such a measure. Specifically, for parents we use registry-based information about the highest educational level achieved and assign the standard duration of each educational program as the participant's "years of education." We take the mean score of both legal parents to represent the average parental education in years. When data is missing for one parent (as it is for 1.7% of children enrolled in Danish public elementary school in 2015), the mean parental education variable is simply the score for the one parent with data. Dichotomizing parental educational attainment does not alter our conclusions.

Participants

The timing of the introduction of the various assessments discussed above are such that no individual has meaningful data on all four categories of child educational characteristics discussed above. Accordingly, we conduct our analyses not on a single sample of participants providing all relevant data but on several often-overlapping samples defined as described below. All samples shared one criterion, namely that parental education data must be available for at least one legal parent. We always contrast the full population (designated Sample *N*a) with the subsample of non-family adoptees (Sample *N*b) – more information on adoptees below.

Samples 1a and 1b – High school completion

Members of samples 1a and 1b were all those born between 1988 and 1993, reflecting limitations in the dates of available information on (a) adoptive status (adoption data is not

available prior to this year) and (b) high school completion, as described in Appendix A1. Table A1 presents a breakdown of data availability (for this and all other samples), which highlights how many participants are eliminated by a given restriction. For the present sample, the table shows that of the 477,384 individuals recorded in the Danish registry with birthdays in the years 1988-1993, we lack parental education data for 14% (who are overwhelmingly children of immigrants), leaving us with a final sample of 412,295 for sample 1a. Sample 1b consists of the 3,297 Sample 1a members who are "non-family" adoptees (described more below).

Samples 2a and 2b – Academic performance (exit exams) and high school enrollment

Samples 2a and 2b include those taking the exit exams between 2007 (the first year the procedure for this period was adopted) and 2014. Table A1 specifies where attrition occurs. Of the 545,792 people enrolled in 9th grade (public or private) between August 2007 and June 2014, 1% are lost due to missing parental education data and 14% are missing ninth-grade exit exam data, leaving us with 465,358 individuals in Sample 2a. Sample 2b consists of the 3,505 Sample 2a members who are non-family adoptees.

Samples 3a and 3b – Conscientiousness and performance on Danish national tests

Participants in these include all students who completed one or more of the annual well-being assessments performed in 2015-2017, and (b) completed a Danish national test in reading between 2014 and 2017. A total of 536,593 children were enrolled in grades 4-9 in any Danish school during the years 2014 to 2017, with approximately 80% enrolled in public schools (for which these assessments were mandatory) at any one point in time. Less than 2% of public school students are eliminated due to missing data on parental education, and 11% are eliminated due to lack of a Conscientiousness score. Less than 5% of the remaining children lack data on the Danish national tests. The final *N* for Sample 3a is 392,163. Sample 3b consists of the 2,799 members of Sample 3a who are non-family adoptees.

Adoptees

The Danish Civil Registration System records not only whether an individual is legally adopted by another person but also whether that adoption was performed by a non-family member – i.e. those not performed by a relative or a step-parent. Because our adoptee analyses are intended to eliminate all potential genetic confounding, we analyze only these "non-family" adoptions. The concern with genetic confounding is also an issue for the sibling analyses we conduct. Because information on the adoptees' birth parents is not available in the Danish registries, we seek to reduce the likelihood that two adoptees share birth parents

by eliminating pairs of adoptees who were adopted on the same day from all analyses of adoptive sibling correlations.

Of the 8,282 individuals distributed throughout our three adoptee samples (1b, 2b, and 3b), 56% are female, and represent a total of 50 countries of birth. The median age at adoption in these samples is 15.3 months, with 75% completed by 28.3 months. Adoption procedures vary based on country of child origin, with some (e.g. South Korea) exhibiting a low (10.5 months) average age of adoption, whereas others (e.g. Thailand) were considerably later (52 months). Adoptive parent attainment was minimally associated with salient characteristics of the adoptee. More educated parents were trivially but statistically significantly more likely to adopt children from less developed countries as scored using the 2007 Human Development Index (HDI) scores (r = -0.02, [-0.04, -0.01]), and adopted children modestly younger at age of adoption (r = -0.07, [-0.10, -0.05]). Consistent with other research (6), parents of children in our three adoptive samples were somewhat more educated (mean years of education = 14.94; SD = 2.03] than were parents of children in our three main non-adoptive samples (1a, 2a, and 3a) [M = 14.00; SD = 2.30].

Empirical strategy

The non-experimental nature of the present work means it is not well-suited to positively support a claimed causal role for parental education, but it can provide critical tests that would cast doubt on such causal claims. One such relatively indirect test involves quantifying sibling resemblance for each child educational characteristic, both in the full population and among adoptees, to evaluate the degree to which these characteristics are influenced by genetics and by the shared environment.[1] Because levels of education in rearing parents are shared by the children in that home, any causal effects of parental education might reasonably be expected to lead to phenotypic similarity between co-resident siblings, even in the absence of genetic similarity between such siblings. Translated into the terminology of standard behavior genetic models, this is to say that parental education's

[1] Because the full sample sibling correlations are based on biological siblings of diverse varieties (MZ and DZ twin pairs, full siblings, half-siblings, not all of which are readily differentiable from each other using registry data) as well as adoptive siblings, one should not simply compare the full sample and adoptee subsample coefficients and attempt to obtain a precise estimate of the role of genetics for the characteristic, even if the magnitude of the genetic effect can be loosely approximated by such a comparison.

causal effects on children are likely to reveal themselves as shared environmental effects.[2] Therefore, we analyze sibling correlations for each characteristic, both in the population at large and in a subset of that population consisting of adoptees. (In households where more than two children meet the requirements for inclusion in a given sample, we use only the oldest two children meeting those requirements for our sibling correlation analyses.)

If a given educational characteristic appears influenced by features of the shared environment – indicated by a nonzero correlation between adoptive siblings for the characteristic – the next step is to explore how parental education predicts that characteristic. For this, we regress child educational characteristics on parental attainment both in the full sample and among adoptees. Stronger coefficients for the full rather than the adoptive sample are typically interpreted as evidence for the presence of genetic influences which affect both educational attainment in the parents and the child feature in question. An absence of *any* degree of relationship among adoptees would provide a challenge to any claims for a causal role of parental education on the child characteristic, as it would indicate that no relationship exists aside from that more parsimoniously attributed to common genetics. By contrast, a positive nonzero relationship would be consistent with causal claims.

If such a nonzero relationship is observed, one further step can then be taken to facilitate assessing alternative accounts of parental education's influence, here by exploring whether parental education can plausibly be said to have a relatively direct causal effect on the specific characteristic in question, or whether instead its impact on that characteristic is mediated by effects on earlier stages of the child's educational development. For this analysis, we add as a predictor the child's score on the previous level of educational development (e.g., academic performance as a predictor of enrollment in advanced education). If parental education does not retain predictive power of the later-stage child educational characteristic, this would speak against any causal accounts which focus on effects that are specific to that educational characteristic – as opposed to having effects on that characteristic by first having impacted earlier stages in the child's educational development. For example, if parental and child attainment correlate partially because of educated parents being able to better help their kids navigate the particular challenges in enrolling in advanced education, that should reveal itself in an association between parent attainment and child enrollment in advanced education among adoptees that persists even

[2] This, of course, depends on the assumption that particular "objective" feature of the shared environment – here, parental levels of education – is also "effectively" shared for the children, in that it affects all children in a given home in generally similar ways.

after controlling for the child's academic performance. Or, if the association between parental and child attainment reflects parents using their wealth to financially support their children's studies, this too should reveal itself in the same kind of association (to the extent that such wealth impacts the child's decision to go to school) or in an effect of parental attainment on adoptive child attainment after controlling for child enrollment in advanced education (to the extent that such wealth impacts the child's ability to stay in school during financial hardship).

Results

We begin our analyses with the outcome of greatest interest – child attainment – and work backwards through the child's educational development (see Figure 1). Figures 2 and 3 show a visual representation of the study's core results, which are also presented in Table A2. These represent regression results with parental attainment predicting various child characteristics, separately for the full sample and for adoptees, with child gender as a control.

Educational attainment

A first indication of the importance of familial influences – whether genetic or shared environmental – on attainment is provided by the significant similarity between siblings in high school completion. In the relevant full sample (1a), siblings correlated .56 ([95% CI: .55, .57], 70,074 sibling pairs) for completion. In the corresponding adoptee sample (1b) sibling similarity is reduced but still substantial (.25; [.11, .39]; 449 sibling pairs) – 45% of the correlation in the full sample, indicating that not only genetics but also features of the shared environment are notable contributors to the outcome.

Panel (a) of Figure 2 allows us to explore whether parental attainment might plausibly account for some of these shared environmental effects. Although parent-child associations in the population at large do not allow us to separate genetic from shared environmental influences, the large size of these associations are still noteworthy: In the full sample (1a), a 1 SD increase in parental years of education predicted a 16.6 percentage point [16.4, 16.7] increased likelihood of high school completion. Thus, whereas at mean parental education the children completed high school at 55%, the predicted probability of completing HS at 1 SD above the mean in parental education was closer to 72%.

Of even greater interest is the corresponding result in the adoptee subsample, where the effect is 2.9 percentage points [1.17, 4.63]. The fact that this value is nonzero is consistent with some causal role for parental attainment on child attainment, even if the modest size of this result (only 17% of the result in the full sample) indicates that genes shared between the parent and child account for the majority of the covariation between these traits. That is,

although the adoptive sibling correlations highlighted a significant role for the shared environment as an influence for child attainment, the regression results suggest that whatever aspects of the shared environment contributed to sibling similarity in attainment do not vary markedly as a function of parental attainment. If parental attainment were a highly effective cause of child attainment, or even simply a highly effective indicator of the presence of some other (nongenetic) causal factor, its association with child attainment in the adoptee sample should be more pronounced.

Nevertheless, because some association between parental and child attainment is evident even among the adoptees, an evaluation of the timing of these effects is merited. Panel (a) of Figure 2 represents these results, showing the association between parental attainment and child high school completion after controlling for the immediately prior stage in educational development, high school enrollment. Introducing this control eliminates 72% of the effect of parental attainment in the full sample, and completely eliminates the effect in the adoptee sample. This suggests that any causal effect of parental attainment on high school completion is less due to parents' contributions during the high school years themselves, but instead reflects the effects of parental attainment on earlier stages of the child's educational development. Therefore, we need to seek the initiation of these causal effects in earlier developmental stages.

Enrollment in advanced education

Results for advanced educational enrollment are highly parallel to those for attainment. First, high school enrollment is substantially familial: the sibling correlation in the corresponding full sample (2a) is identical to that for completion (r = .56, [.55, .57], 79,974 sibling pairs), though the characteristic may be comparatively more influenced by the shared environment and less influenced by genetics: The correlation in the corresponding adoptee sample (2b) is larger than that for educational attainment (r = .33, [.18, .48], 449 sibling pairs, 59% of the effect in the full sample).

Similarly, parental attainment substantially predicts enrollment in advanced education: as shown in Panel (b) of Figure 2, students with 1 SD higher parental years of education were 11.0 percentage points [10.9, 11.1] more likely to enroll in high school in the corresponding full sample. As with child attainment, however, this effect was much smaller among the adoptees, among whom the corresponding value was only 2.2 percentage points [0.8, 3.6]. As with child attainment, this is consistent with some causal role for parental attainment on child enrollment in advanced education, though the modest size of this result (only 20% of the result in the full sample) indicates that genes shared between the parent and

child account for substantially more of the covariation between these characteristics. As before, then, although the adoptive sibling correlations highlighted a significant role for the shared environment as an influence for child educational enrollment, our results suggest that whatever aspects of the home environment contributed to sibling similarity in enrollment were not very strongly related to parental attainment.

A final parallel between the results for attainment and those for enrollment is the most important: Just as conditioning on enrollment eliminated any connection between parental attainment and child high school completion among adoptees, conditioning on academic performance (as assessed with 9th grade exit exams) eliminates roughly 70% of the link between parental attainment and child enrollment in both the full and adoptee samples, rendering the adoptee result non-significant (adoptee conditional result = 0.8 percentage points, [-0.34, 1.92]) and the full sample result greatly diminished (3.1 percentage points, [3.0, 3.2]). This reflects the strong effect of exit exam scores on high school enrollment: in both the full and adoptive samples, a 1 SD increase in exit exam score predicted a roughly 20 percentage point increased likelihood of high school enrollment.

As with child attainment, then, these results are not consistent with a causal role for parental education influencing child educational enrollment as a specific outcome. The covariation between these characteristics predominantly reflected genes shared between the parent and child, though it may also secondarily reflect causal effects of parental education on child educational characteristics preceding enrollment in high school. We thus must progress further back in the child's educational development to search for evidence compatible with a causal effect of parental educational attainment on a specific stage in the child's educational development.

Academic performance

As with educational enrollment and attainment, academic performance was substantially familial. For the most comprehensive measure of academic performance (the 9th grade exit exams), we observed sibling correlations of .51 ([.50, .51], 106,972 sibling pairs) in the relevant full sample (2a). The sibling correlation of .24 ([.16; .32], 534 sibling pairs) in the corresponding adoptee sample (2b) indicates the familial influences are roughly evenly split between genetic and shared environmental sources.

Panels (a) and (b) of Figure 3 show that, as with the previous characteristics, there is a substantial link between parental attainment and child academic performance in the relevant full sample, whether that is assessed using the comprehensive exit exams (beta = .44, [.44, .44]; Sample 2a) or the bi-annual language test (beta = .34, [.34, .34]; Sample 3a). It also

shows that these links are substantially attenuated in the adoptee samples: the corresponding values are betas of .05 [.02, .09] (Sample 2b) and .06 [.02, .10] (Sample 3b).

Academic performance thus resembles the previously discussed educational characteristics (attainment and enrollment in advanced education): Performance is substantially influenced by both genes and environment, and it associates with parental attainment primarily, but not exclusively, through genetic mechanisms. Unfortunately, we are unable to complete a final comparison – whether the causal effects of parental attainment on the characteristic is likely entirely attributable to effects earlier in the child's educational development – due to limitations in available data. Specifically, our assessment of educationally-relevant psychological characteristics (which in Figure 1 is indicated to precede academic performance) is obviously not comprehensive, given that we lack any information concerning the child's IQ and instead have only personality data available. Using the bi-annual language tests to indicate academic performance, so as to allow a comparison with Conscientiousness scores among a large number of children, regression results (Table A2) show that Conscientiousness is, in fact, a reasonably potent predictor of academic performance (betas of .26 [.26, .277] and .32 [.28, .35] in the full and adoptive samples, respectively). However, controlling for Conscientiousness in panel (b) only reduces the effect of parental attainment on child academic performance by roughly 16% in both the full (3a) and adoptee (3b) samples, compared to *much* larger reductions for the previously discussed comparable analyses on attainment and enrollment in advanced education. (Note that panels (a) and (c) of Figure 3 do not show conditional estimates.) Parental attainment thus remains a significant predictor of child academic performance for both the full (beta = .29 [.28, .29]) and adoptee (beta = .05, [.01, .09]) samples.

Academically-relevant psychological characteristics

Sibling correlations for Conscientiousness are, relative to the other traits discussed, comparatively modest, consistent with previous research (16). In the relevant full sample (3a), sibling correlations were only .23 ([.22,.23], N = 96,993), with the adoptive sibling correlations (sample 3b) smaller still (r = .11, [.01,.20], N = 516). However, as shown in Panel (b) of Figure 3, while parental attainment significantly predicts child Conscientiousness (beta = .20, [.20, .21]) in the full sample (3a), this association is markedly reduced among adoptees (beta = .03, [-.01, .07]), this time to statistical insignificance. Accordingly, while there appears to be both a genetic and shared environmental component for Conscientiousness in the present sample, there is no evidence for an environmental effect of parental education on this child characteristic.

Supplementary analyses

Because of the small but statistically significant (negative) association between parental attainment and the HDI score for the child's country of origin, we performed supplementary versions of all analyses described above. In each analysis, an interaction between parental attainment and HDI was not significant, indicating that parental attainment has consistent relationships with child educational characteristics across different countries of origin. The small link between parental attainment and (young) child age at adoption also required investigating, given the possibility that the elevated educational outcomes exhibited by children adopted by educated parents reflected this extra time in the adoptive home. This was not the case, as indicated by supplementary versions of all analyses that included an interaction term for child age at adoption and parental education.

Discussion

The present research points to two primary trends. The first is that although in the full population we found that more educated parents have children who thrive at every stage of educational development, this relationship is overwhelmingly dependent on the nature of the family: When the parents and children do not share genes, the children reared by educated parents look little different from those reared by uneducated parents.

The second finding of particular interest is that any truly causal effects that are behind the relationship of parental attainment with child educational characteristics appear to happen comparatively early in the child's educational development. Rather than specifically influencing the child's enrollment in or completion of advanced education, our results are more consistent with parental attainment instead influencing these characteristics by first affecting child academic performance.

Evaluating alternative explanations

There are well-known assumptions and limitations to the use of adoptee samples (17), but we can address or examine many of these in the present study. From these examinations, we saw no cause for concern. For example, it is true that adoptive parents were moderately more educated than was the full population, but because the variance in years of education completed was only modestly (12%) smaller among adoptive parents than among the full population of parents, there seems little reason to attribute the failure of adoptive parental attainment to predict educational characteristics of the child to insufficient variation in that education.

We also see no reason to attribute these results to the age at which the children were adopted. First, supplementary moderation analyses did not suggest effects were stronger

among those adopted earlier. Second, if adoptive parental education failed to predict child characteristics simply because the adoptive children had been insufficiently exposed to the environment provided by their adoptive parents, then effects of adoptive parental education might be most expected at *later* ages, by which time the children have spent more time in their adoptive environment. That contrasts with our observation that the shared environmental effects of parental attainment were "baked in" to the earlier stages. However, our study can of course not rule out the possibility that parent-child similarity in educational characteristics among the full population partially represents the effects of behaviors of the parents very early in the child's life (such as in the first months or even in utero, before the time when meaningful numbers of our adoptee samples was living in their adoptive homes). Similarly, the registry-based nature of our study means we have no information on whether adoptive children have post-adoption contact with their birth parents. However, because there were no meaningful differences in our results when we limited our analyses to international adoptees (who typically would have little chance for such contact), we see no reason to attribute our results to any such contact that might occur.

To evaluate the generalizability of our adoptee findings, it is instructive to compare them to those from other cultural contexts as well as those derived from other approaches (e.g. twin studies). Previous research has found that genetic influences on educationally-relevant features such as intelligence are suppressed (and shared environmental features enhanced) under deprived socioeconomic conditions (18). The wealth and redistributive policies of northern Europe are such that comparatively fewer families are likely to find themselves in such deprivation, with the expected effect being that in such societies a comparatively larger role will be played by genetics and a smaller role by the shared environment. Comparing twin studies from different contexts (see 21) supports this expectation: when compared to results from other countries, there is substantially higher heritability and (with the exception of Norway) lower shared environmentality for educational attainment in Denmark (20), Finland, Sweden, and Germany.[3]

This context helps to show that the estimated role for the shared environment in the present study – .25 for educational attainment, as indicated by the adoptive sibling correlation

[3] This trend was not identified in the meta-analysis (19), reflecting that study's inclusion of a working paper on a segment of the Danish Twin Registry with many young participants. As can be expected, the young age of the participants appears to have markedly suppressed the familiality of educational attainment. A published study on an appropriately-aged segment of the Danish Twin Registry (20) that was not included in the meta-analysis instead shows results that are very much in line with those of other northern European nations.

– is a plausible result for a Danish study, even if it is smaller than the result derived from all twin studies from around the world observed (.36, see 21). In fact, .25 is precisely the shared environmental estimate also observed by the Danish twin study (20). By contrast, the shared environmental estimate of .37 derived from McGue's American adoptive sample (3) is a reasonable match to the estimate expected from American twin samples (19). We thus have no clear indication that adoptee samples, especially of the size and completeness of that used here, are particularly prone to understate the importance of the shared environment for these features. At the same time, there is clearly a reason to think that studies in other nations may point to a greater role for the shared environment. Potentially, this might also indicate other countries would find a greater causal role for parental education in other environments, although a comparison of the results of McGue (who observed an odds ratio for college completion among adoptees of 1.3) and the educational attainment results of our own high school completion result among adoptees (odds ratio 1.13) gives us no reason to expect this, as we cannot reject equality of the estimates (z-value for test of equality was 1.3). Nevertheless, further study in contexts where a greater share of the population faces economic deprivations is clearly merited before the present results are unquestioningly extrapolated to such contexts.

Timing of parental influence

Two pieces of our study require consideration of timing. The first is the issue of which educational characteristics are most important. The second concerns when parents matter most.

Most literature on educational attainment focuses on completion of college degrees, whereas the present study uses high school attainment, which as noted above was necessitated by the young age of those whose adoptee status is recorded in public Danish Registries. Importantly, in the Danish context high school ("gymnasium") plays a very different role than in educational systems such as the U.S. Enrolling in high school is not open to all students, instead requiring students to demonstrate substantial academic competence. Students with a less academic orientation more often pursue one of many alternatives with a more vocational angle. The significance of attaining a high school degree is perhaps best indicated by considering its associations with life outcomes. In terms of further education, although only 31% of the Danish population obtains a college degree, fully 62% of those who completed high school will complete such a tertiary degree (notes about this result in Appendix A3). Economically, completion of both high school and subsequently college seems to be associated with meaningful rewards: A comparison of earnings published

by Statistics Denmark (21) notes that men and women with a high school degree earn 21% and 7% more on average than those with compulsory schooling only, and those with a bachelor earn another 18% or 19% more (men/women), and those with a master's/long university degree 48% or 55%. Accordingly, many of the dynamics that influence college enrollment and completion – whether they be considerations of future earning potential, academic interests, or academic competencies – are highly relevant for high school enrollment and completion in Denmark.

At least one major difference requires further attention, however, and that is the age typical for enrollment and attainment for high school versus for college. Previous research has found that the influence of parents and the shared environment changes substantially over the life course, diminishing as the child ages (11). Accordingly, our estimates for the importance of the shared environment and the correlation in educational characteristics between adoptive parents and children might be expected to be overestimated in the present work compared to analyses of college enrollment and completion. However, the identical estimates for the role of the shared environment in our adoptee study of high school completion and the previously noted Danish twin study of college completion does speak against this possibility.

Implications for interventions

The pronounced degree of intergenerational transmission of educational characteristics has long served as a source of inspiration for those looking to provide greater opportunities to children from less advantaged families. In particular, identifying the kinds of behaviors common to more educated parents can serve to inspire attempts to facilitate those with less education to copy those behaviors, in the hopes of improving the educational outcomes for children in these families. Results from the present study highlight that such interventions might be reasonably expected to be more potent if they affect characteristics relevant to earlier rather than later stages in educational development. This is because our results were consistent with any causal effects of parental attainment on child attainment being already transmitted to the child by the time of the 9th grade exit exams. Interventions aiming to help less educated parents behave in ways that facilitated their child's eventual educational attainment might be best targeted at interventions that contribute to the child's academic performance prior to this age, rather than, for example, focusing specifically on influencing the child's decision to enroll in advanced education; or the child's ability to gain entry to that advanced education; or the child's ability to complete that education (22). At the same time, it is noteworthy that one potential pathway that meets these criteria was not

supported. We did find that in the full population, educated parents had children who scored meaningfully higher on Conscientiousness. We also observed that scores on this trait were effective predictors of academic performance. However, we found no evidence that the relationship between parental education and child Conscientiousness was causal, as the relationship was not present among the adoptive sample. Despite the utility of high Conscientiousness, then, our results do not support exploring how the environment provided by highly educated parents facilitates Conscientiousness in their children, as the relationship may simply reflect shared genes.

Conclusion

Using a national, genetically-informative sample, our study provides the most thorough exploration of how parental educational attainment associates with their children's educational development ever performed. Both genes and the shared environment influenced each stage of the child's educational development, beginning with the child's personality, through the child's academic performance, and into their advanced educational enrollment and attainment. However, our findings were not consistent with parental attainment being of pronounced *causal* importance for these educational outcomes, and what causal effects were supported were most readily attributable to effects on academic performance that were then carried through to later educational outcomes.

References

1. Turkheimer E (2000) Three laws of behavior genetics and what they mean. *Curr Dir Psychol Sci* 9(5):160–164.
2. Haugaard JJ, Hazan C (2003) Adoption as a natural experiment. *Dev Psychopathol* 15:909–926.
3. McGue MK, Rustichini A, Iacono WG (2017) Cognitive, non-cognitive and family background contributions to college attainment: A behavioral genetic perspective. *J Pers* 85(1):65–78.
4. Kendler KS, Turkheimer E, Ohlsson H, Sundquist J, Sundquist K (2015) Family environment and the malleability of cognitive ability: A Swedish national home-reared and adopted-away cosibling control study. *Proc Natl Acad Sci* 112(15):4612–4617.
5. Neiss M, Rowe DC (2000) Parental education and child's Verbal IQ in adoptive and biological families in the National Longitudinal Study of Adolescent Health. *Behav Genet* 30(6):487–495.
6. Bjorklund A, Lindahl M, Plug E (2006) The origins of intergenerational associations: Lessons from Swedish adoption data. *Q J Econ* 121(3):999–1028.

7. Davies S (2004) School choice by default? Understanding the demand for private tutoring in Canada. *Am J Educ* 110(3):233–255.
8. Sandefur GD, Meier AM, Campbell ME (2006) Family resources, social capital, and college attendance. *Soc Sci Res* 35(2):525–553.
9. Johnson W, McGue MK, Iacono WG (2007) How parents influence school grades: Hints from a sample of adoptive and biological families. *Learn Individ Differ* 17(3):201–219.
10. Poropat AE (2009) A meta-analysis of the five-factor model of personality and academic performance. *Psychol Bull* 135(2):322–38.
11. Bergen SE, Gardner CO, Kendler KS (2007) Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: a meta-analysis. *Twin Res Hum Genet* 10(3):423–33.
12. John OP, Srivastava S (1999) Big Five Inventory (Bfi). *Handb Personal Theory Res* 2:102–138.
13. Chamorro-Premuzic T, Furnham A (2004) A possible model for understanding the personality--intelligence interface. *Br J Psychol* 95(Pt 2):249–64.
14. Nandrup AB, Beuchert-Pedersen LV (2018) The Danish national tests at a glance. *Natl Tidsskr* 2018(1). Available at: http://www.forskningsdatabasen.dk/en/catalog/2394771282 [Accessed August 8, 2018].
15. Sievertsen HH, Gino F, Piovesan M (2016) Cognitive fatigue influences students' performance on standardized tests. *Proc Natl Acad Sci* 113(10):2621–2624.
16. Matteson LK, McGue MK, Iacono WG (2013) Shared environmental influences on personality: a combined twin and adoption approach. *Behav Genet* 43(6):491–504.
17. Jaffee SR, Price TS, Reyes TM (2013) Behavior genetics: Past, present, future. *Dev Psychopathol* 25(4 PART 2):1225–1242.
18. Turkheimer E, Haley A, Waldron M, D'Onofrio BM, Gottesman II (2003) Socioeconomic status modifies heritability of IQ in young children. *Psychol Sci* 14(6):623–628.
19. Branigan AR, Mccallum KJ, Freese J (2013) Variation in the heritability of educational attainment: An international meta-analysis. *Soc Forces* 92(1):109–140.
20. Rodgers JL, et al. (2008) Education and cognitive ability as direct, mediating, or spurious influences on female age at first birth: Behavior genetic models fit to Danish twin data. *Am J Sociol* 114(S1):S202–S232.
21. Danmarks Statistik (2013) *Befolkningens Løn* (Copenhagen).
22. Heckman JJ, Kautz T (2014) Fostering and measuring skills: Interventions that improve character and cognition. *The Myth of Achievement Tests: The GED and the Role of Character in American Life*, eds Heckman JJ, Humphries JE, Kautz T (University of Chicago Press, Chicago).
23. John OP, Naumann LP, Soto CJ (2008) Paradigm shift to the integrative Big Five trait

taxonomy. *Handb Personal Theory Res*:114–158.

24. Goldberg LR, et al. (2006) The International Personality Item Pool and the future of public-domain personality measures. *J Res Pers* 40(1):84–96.

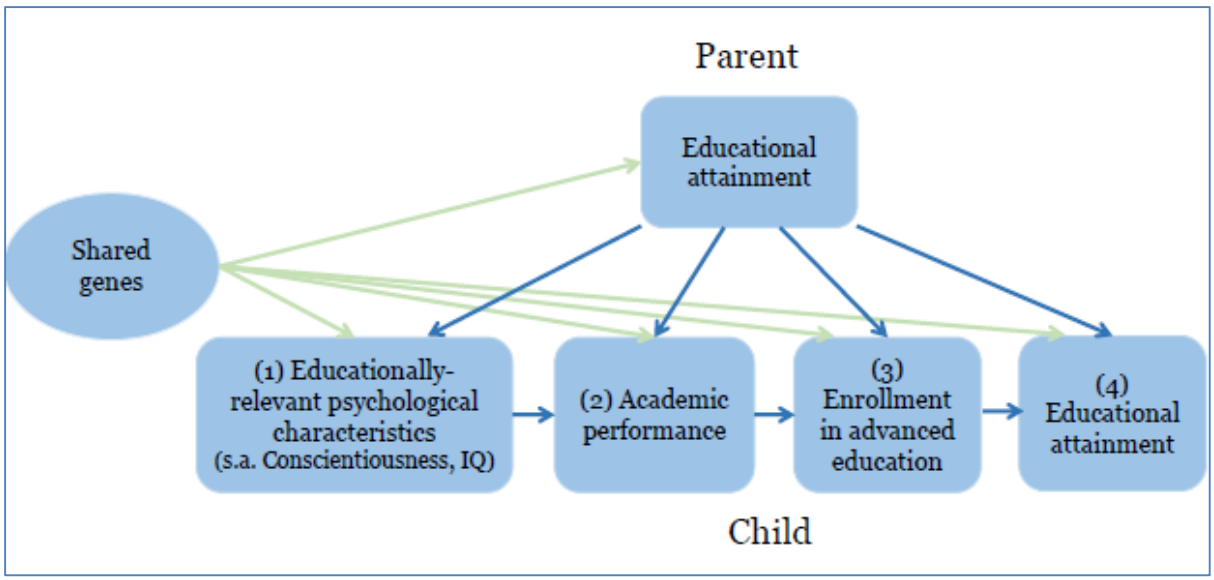

***Figure 1*. Model of potential connections among parental educational attainment and four stages of the child's education.**

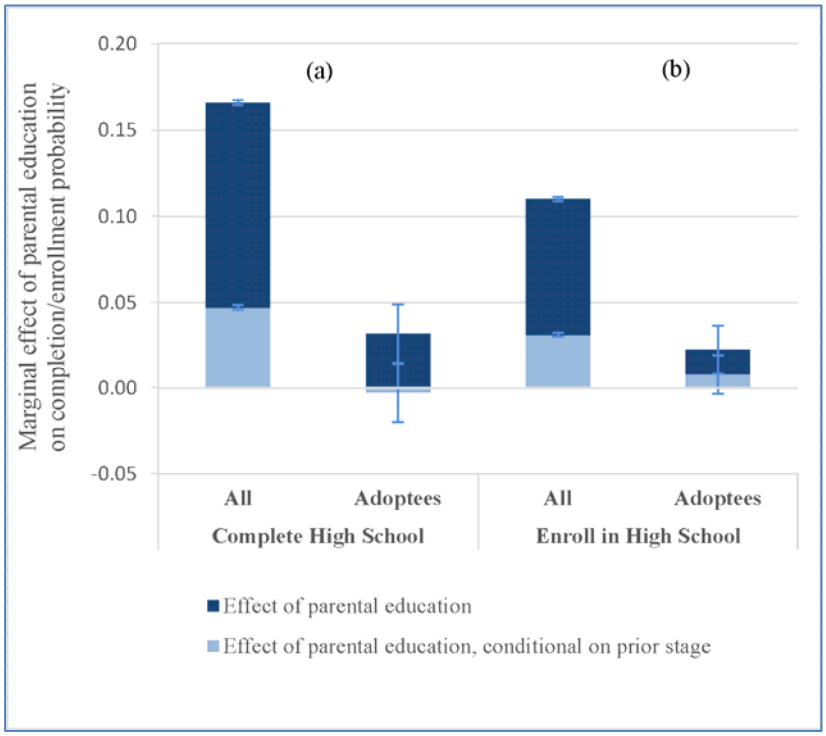

***Figure 2*. Effect of parental education on high school completion and enrollment probability.**

The effects shown are marginal effects of one standard deviation of parental education after estimating the binary outcome of completion/enrollment. The light blue column shows the partial effect of parental education that remains when conditioning the outcome on the prior stage, i.e. conditioning on enrollment for the completion outcome, and conditioning on high school exit exam scores for the enrollment outcome.

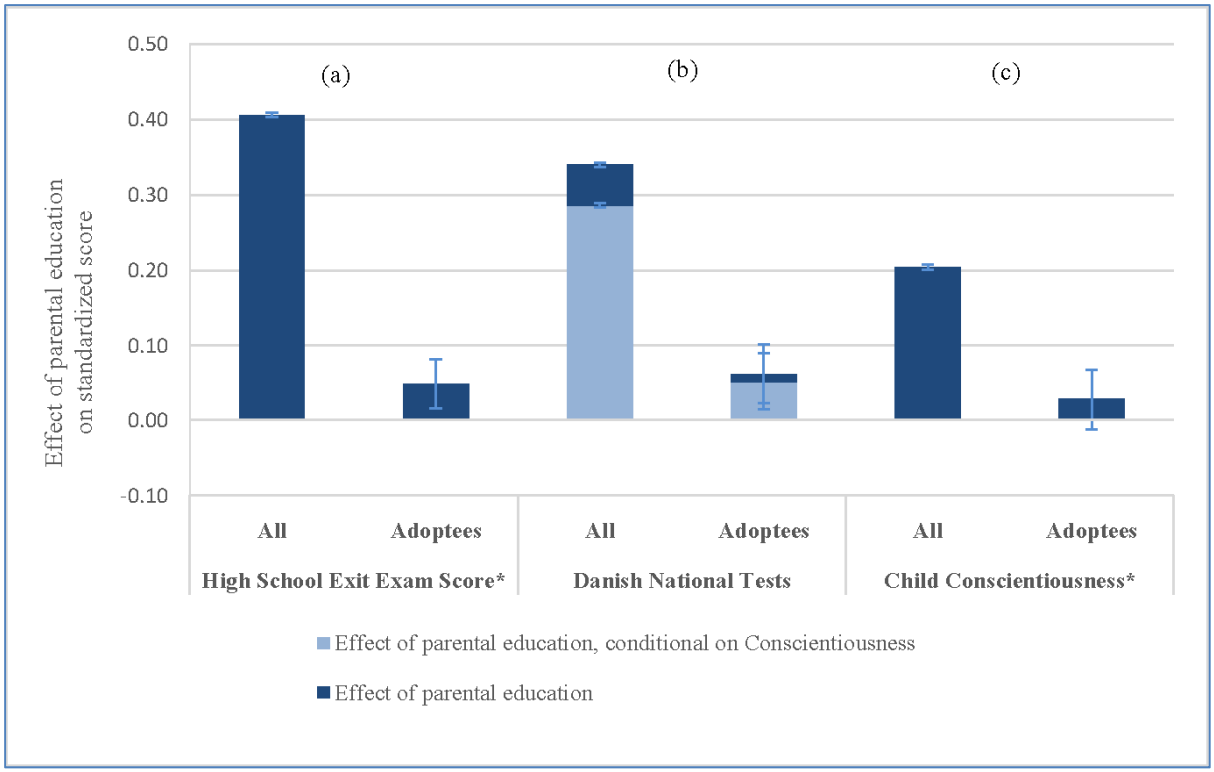

Figure 3. **Effect of parental education on academic performance and conscientiousness.**

The effects shown are regression coefficients of standardized parental education on standardized scores of Samples 2a (panel a) and 3a (panels b and c). The light blue column shows the partial effect of parental education that remains when conditioning the outcome on the prior stage. * denotes a characteristic for which the reported effects are never conditioning on a prior stage.

Appendix A1

Measures

Further details on Conscientiousness

In 2015 we collected additional data on a sample of Danish public school students. Schools were invited to participate in a project aimed at validating the questionnaire used in the annual national assessment of wellbeing. 6 schools participated, with 197 to 603 students, and 2,474 students in total. To reduce the length of the validation survey, students were randomly assigned to different previously validated instruments alongside the items used in the national assessment. Thus, a group of students was assigned to the Big Five Inventory (N = 399; John, Naumann, & Soto, 2008) and IPIP Conscientiousness facets (Ns = 429-440; Goldberg et al., 2006).

Our measure of Conscientiousness consists of three items: "How often can you complete what you set out to do?"; "Can you concentrate during class?"; "If interrupted during lessons, I can quickly concentrate again." They are given in Danish, with the original phrasing of "Hvor tit kan du klare det, du sætter dig for?", "Kan du koncentrere dig i timerne?" and " Hvis jeg bliver forstyrret i undervisningen, kan jeg hurtigt koncentrere mig igen."

This measure clearly assesses some components of Conscientiousness better than others: analyzed alongside the six facets of IPIP Conscientiousness, it exhibited pronounced links with the Self-Discipline, Self-Efficacy, and Achievement facets (rs > .50), and much more modest links with Cautiousness, Dutifulness, and Orderliness (rs < .32). In a multiple regression, the first three facets retained some predictive power (betas > .25) whereas the second set did not (betas < .09).

Further details on exit exams

The exit exams acquired their present form in 2007, dividing into a stable component assessed each year (on written Danish, oral Danish, reading, spelling, problem-solving (math), skills (math), oral science (physics, chemistry, biology and geography), and oral English) as well as in additional topics drawn by lot each year. To make consistent comparisons across years, we limit our analyses to the stable exams used from 2007 onward.

The grading and preparation of each exam is dependent on the manner of assessment. Written exams are prepared by the Danish Ministry of Education and then graded by external examiners. Oral exams are instead prepared by the teacher, which are then conducted at the school and evaluated by the teacher and an external examiner. Students are allowed to retake their 9th grade exit exams in 10th grade (an optional, additional pre-gymnasium year in the Danish educational system), but in our analyses, we use each student's first attempt at each exam to make consistent comparisons across all students.

Further details on National Tests

The national tests are given each year, with Danish tests given to grades 2, 4, 6, and 8. The overall score of this adaptive computerized test is based on three sub-parts:

Language comprehension: A key element in language comprehension is word awareness. The language comprehension test contains questions that covers word awareness on multiple levels. For example, the pupil has to match a picture with a word in one exercise. Whereas other exercises focus on knowledge of homonyms and language use which for example is knowledge of idioms.

Decoding: This tests the pupil's ability to identify written words and letters. For example, by setting up words in a long chain without spaces, where the pupil must identify and separate the words from each other. Another exercise could be where the pupil reads a word and must choose a corresponding picture among five possibilities. The pupil may be very good at decoding without understanding the context or meaning of the text.

Reading comprehension: It tests the pupil's abilities in the comprehension of written words, which means the questions test whether the pupil understands the context and meaning of a text. The exercises are designed to test if the pupil can obtain information from a given text. The pupil has to read a text and afterwards answer questions in relation to the text. Another exercise is where the pupil must fill out the empty space in a given text by choosing between four similar words. Thus, he must select the word, which fits best into the context.

Higher educational enrollment and attainment

Because approximately half of eventual high school enrollees first pursue an optional, pre-high school 10th grade year to improve their academic background before enrolling in high school, to analyze enrollment decisions we use registry information on student enrollment status not one but instead two years after they completed 9th grade. The overwhelming majority (96.5%) of 2009/10 9th graders who enrolled in high school by 2016 had done so by 2012.

Table A1

Breakdown of participant loss in each sample

	Sample 1		Sample 2		Sample 3	
	Born between 1988 and 1993		Enrolled in 9th grade between August 2007 and June 2014		Enrolled in 4th-9th grade January 1st 2014-2017	
	Full population	Non-family adoptees	Full population	Non-family adoptees	Full population	Non-family adoptees
Initial Sample Size	477,384	3,348	545,792	4,375	536,593	4,222
Enrolled in public school January 1st	NA	NA	NA	NA	458,901	3,362
Parental education data for at least one parent available	412,295	3,297	539,865	4,359	452,422	3,361
Conscientiousness data available	NA	NA	NA	NA	404,751	2,940
Academic performance measure available	NA	NA	465,358[a]	3,505[a]	392,163[b]	2,799[b]

Final N for analysis	412,295	3,297	465,358[a]	3,505[a]	392,163[b]	2,799[b]

Note. [a] = 9th grade exit exams; [b] = National test in Danish language. NA indicates the restriction does not apply to the sample, thus leading to no loss of participants.

Samples 1a and 1b – High school completion

Members of samples 1a and 1b were all those born between 1988 and 1993, a date restriction resulting from constraints on availability of registry data and the ages of the relevant populations. First, as discussed below in the adoptee section, information on adoptive status is not available for adoptions occurring prior to 1988. Second, registry information on high school attainment was available to us only through 2016. We sought to reduce the frequency with which we would classify an individual as having not completed high school when they would in fact eventually complete that degree at a later date; to this end, we analyzed all those born in 1988 who had completed high school by 2016 (i.e., by age 28), looking for the year by which 95% of completers had finished their degree. For sample members who were adoptees, that year was 2011, which is one year later than the same result for the general population (2010). We accordingly set our latest birth year for samples analyzing high school completion to be 1993, as among both adoptees and the general population at least 95% of those born in this year who will complete high school by age 28 might be expected to have completed their degree by 2016.

Samples 2a and 2b – Academic performance (exit exams) and high school enrollment

Our restrictions for these samples were (a) the availability of scores on the exit exams taken in the years that followed a single common procedure, and (b) information on academic enrollment choices two years subsequent to the completion of 9th grade. Because of delays in the availability of registry data, the latest year for which high school enrollment data is available at the time of writing is 2016, with the result that Samples 2a and 2b concern those taking the exit exams between 2007 (the first year the procedure for this period was adopted) and 2014.

The 14% of those enrolled in 9th grade for whom we do not have a score on their exit exams, a fifth are excluded because of missing one exam, another 30% were missing between two and seven exams, and half were missing scores on all eight exams. Missing exams is discouraged but (during the years for this sample) did not prevent students from continuing their education. Those missing *all* exams were a heterogeneous group, and included students in international schools (which are exempted from these exams, with international exams substituted for the Danish ones), students in schools for those with special needs or treatment programs, as well as students enrolled in schools for whom such testing is required (i.e., students for whom the missing tests likely reflected truancy).

Appendix A2

Table A 2

Detailed results for main analyses

(a)

	High School Completion				High School Enrollment			
	Direct Effects		Conditional on enrollment		Direct Effects		Conditional on exit exam	
	(a) Full	(b) Adoptees	(c) Full	(d) Adoptees	(e) Full	(f) Adoptees	(g) Full	(h) Adoptees
Parent Education	0.166 [0.164 ; 0.167]	0.029 [0.0117 ; 0.0463]	0.047 [0.046 ; 0.048]	-0.0026 [-0.02 ; 0.01]	0.11 [0.109 ; 0.111]	0.0223 [0.008 ; 0.036]	0.0312 [0.030 ; 0.032]	0.00791 [-0.003 ; 0.019]
Boy	-0.152 [-0.154 ; -0.149]	-0.136 [-0.169 ; -0.104]	-0.041 [-0.045 ; -0.038]	-0.0648 [-0.098 ; -0.032]	-0.119 [-0.122 ; -0.117]	-0.0911 [-0.118 ; -0.064]	-0.0738 [-0.076 ; -0.072]	-0.0396 [-0.062 ; -0.018]
Exit Exam Scores							0.199 [0.198 ; 0.200]	0.239 [0.229 ; 0.249]
N	412,295	3,297	255,787	1,921	465,358	3,505	465,358	3,505

(b)

	High School Exit Exam 2007-2014		Danish National Test 2014-2017				Child Conscientiousness	
			Direct Effects		Conditional on Child Conscientiousness			
	(i) Full	(j) Adoptees	(k) Full	(l) Adoptees	(m) Full	(n) Adoptees	(o) Full	(p) Adoptees
Parent Education	0.406 [0.403 ; 0.409]	0.0487 [0.0155 ; 0.0818]	0.34 [0.337 ; 0.342]	0.0623 [0.023 ; 0.102]	0.286 [0.284 ; 0.289]	0.052 [0.014 ; 0.0896]	0.204 [0.201 ; 0.207]	0.028 [-0.012 ; 0.068]
Boy	-0.234 [-0.239 ; -0.228]	-0.227 [-0.290 ; -0.164]	-0.208 [-0.213 ; -0.202]	-0.326 [-0.398 ; -0.255]	-0.218 [-0.223 ; -0.213]	-0.29 [-0.358 ; -0.222]	0.0351 [0.029 ; 0.041]	-0.129 [-0.200 ; -0.058]
Child C					0.264 [0.262 ; 0.267]	0.316 [0.282 ; 0.351]		
N	465,358	3,505	392,163	2,799	392,163	2,799	405,762	2,952

Note. Columns (a) to (h) show marginal effects from probit estimations, while columns (i) to (p) show regression coefficients on standardized dependent variables. Parental Education is measured in years, but standardized to have mean zero and standard deviation 1, as are Exit Exam Scores. 95% confidence intervals are given in square brackets.

Appendix A3

We use the administrative registers described in the main text for two further analyses: a) the share of high school students who ever enroll who eventually graduate from high school, and b) the share of the population that will eventually obtain a university degree.

High school graduation

To study the size of the dropout problem from academic high school, we use sample 1a as described in the main text. We ask how many percent of those enrolled will graduate within twice the normative length for finishing high school, which is three years. To have most recent figures for this sample, we thus analyze graduation by 2016 (the latest year in our sample) for those students who enrolled in the year 2010. We do not condition on any covariates, just tabulate graduation status among sample 1a participants who enrolled in 2010.

College graduation

Participants in sample 1a (the oldest of our samples) are, unfortunately, still too young to analyze college graduation. For example, participants born in 1993 can be expected to have graduated from high school by 2016, but many will have barely enrolled in college. Furthermore, Danish completion time of university is traditionally longer than the normed time. Prudence thus dictates to allow for at least 8 years after high school completion to study university completion. Otherwise, by using younger cohorts who have less time, we may underestimate completion. The youngest cohort who would have 8 those years by 2016 would have to have completed high school by 2008. Of the cohort born in 1987, 94% of eventual completers (by 2016) will have completed high school by 2008. Thus, we analyze the 10 birth cohorts from 1978 to 1987, using the full range of the oldest cohorts available together with the most recent cohort that has enough time to graduate.

We use the same definitions for high school completion as in the main text. Our definition of college completion comprises courses at the bachelor's or master's level. The Danish definition includes "medium-length further education" (2-4 years after high school, including engineers, nurses, and teachers), "bachelor-education," and "long further education."

Among the birth cohorts 1978-1987, 62.4% of individuals who completed high school will have graduated one such education by 2016. When not conditioning on high school completion, the graduation rate is 30.5%.