# Investigating the mechanisms of Gene-by-SES interactions for education

Mirko Ruks[1]

[1] Faculty of Sociology, Bielefeld University

mirko.ruks@uni-bielefeld.de

April 11, 2022

**Abstract**

A growing body of research asks whether the opportunity to realize the genetic endowment for education varies by parental socio-economic status (G×SES). While the behavioral genetic Scarr-Rowe hypothesis (SRH) suggests stronger, the sociological compensatory advantage hypothesis (CAH) predicts weaker genetic effects for individuals with a high SES. Using data from the German TwinLife survey, I estimate biometric twin models to test for a G×SES and whether it can be accounted for by SES differences in the effect of genes associated with cognitive ability or by SES differences in the effect of genes independent of cognitive ability. While for secondary school track no G×SES can be found, there is a G×SES for tertiary enrolment in line with the CAH that is mainly accounted for by SES differences in the effect of genes associated with cognitive ability.

**Keywords:** Gene-Environment interaction, Scarr-Rowe hypothesis, Compensatory Advantage hypothesis, Behavioral Genetics, Intergenerational transmission of advantage, Social inequality of educational opportunity.

**Funding:** This study was funded by the German Research Foundation (DFG) as part of the long-term project TwinLife (grant number 220286500), awarded to Martin Diewald, Rainer Riemann and Frank M. Spinath.

# 1 Introduction

One important topic of the research on social mobility is the intergenerational transmission of educational advantage with numerous studies reporting considerable effects of the parental socio-economic status (SES) on educational attainment (e.g. Breen and Jonsson 2005; Jackson 2013). So, while there is abundant descriptive evidence of the social inequality of educational opportunity, the question which mechanisms give rise to this phenomenon, however, is still not fully answered. At the same time, behavioral genetics research reports considerable heritability estimates for educational attainment, with meta studies reporting mean heritability estimates of about 40% (Branigan *et al.* 2013; Silventoinen *et al.* 2020), meaning that genetic differences account for around 40% of the variance in educational attainment. So, if the aim is to shed light on the mechanisms of the intergenerational transmission of educational advantage, the presence of both SES and genetic effects suggests to combine sociological and behavioral genetic accounts.

Trascending the nature vs. nurture debate, a central branch of behavioral genetics research asks how genetic and environmental influences interact in shaping individual traits and behaviors. One part of this research on gene-environment interactions (G×E) investigates whether the opportunity to realize the genetic endowments for education varies by SES (G×SES). Here, two competing hypotheses can be distinguished. On the one hand, the Scarr-Rowe hypothesis (SRH, Rowe *et al.* 1999; Scarr-Salapatek 1971) assumes that the realization of an advantageous genetic disposition is enhanced in enriched environments so that stronger genetic effects can be expected for individuals with a high SES leading to a positive G×SES interaction. On the other hand, the sociological compensatory advantage hypothesis (CAH, Bernardi 2014) assumes that high-SES fam-

ilies compensate for early disadvantages of their children, e.g. disadvantageous genetic dispositions, so that weaker genetic effects can be expected for individuals with a high SES leading to a negative G×SES interaction.

Traditionally, the behavioral genetics literature tests the SRH primarily for cognitive ability (see e.g. Giangrande *et al.* 2019; Gottschling *et al.* 2019; Grasby *et al.* 2019; Hanscombe *et al.* 2012; Spengler *et al.* 2018; Turkheimer *et al.* 2003; Woodley of Menie *et al.* 2018). Here, results are mixed with a meta study showing that the SRH is mainly confirmed for the USA, while for Western Europe and Australia often null findings or negative moderations in line with the CAH are reported (Tucker-Drob and Bates 2016).

However, there is also a growing number of studies that test if the genetic effect on educational outcomes is moderated by SES with most studies focusing on the USA. Domingue *et al.* (2015) show that the genetic effect on the highest school degree is negatively moderated by SES while Conley *et al.* (2015) don't find a G×SES. Papageorge and Thom (2020) report mixed findings for different educational outcomes: A negative G×SES for the probability of obtaining a high school degree and a positive G×SES for obtaining a college degree. Using school socio-economic status as a moderator, Trejo *et al.* (2018) report mixed findings with the probability of postsecondary schooling as the outcome and a positive G×SES for college completion as the outcome. Finally, Lin (2020) finds a negative moderation of the genetic effect on years of education by parental education. Focusing on mathematics tracking in the US high school system and using school SES as a moderator, Harden *et al.* (2020) find a non-significant positive moderation of the genetic effect on the initial tracking level and a significant negative moderation of the genetic effect on the the persistence in a given track. Finally, Uchikoshi and Conley (2021) report stronger genetic effects on educational attainment and mathematics tracking for individuals with a higher SES while Figlio *et al.* (2017) don't find a G×SES for academic achievement. There are substantial fewer studies from outside the USA. For Germany, Baier and Lang (2019) report a positive moderation of the heritability of years of education by parental education. For Finland, Erola *et al.* (2021) find a similar pattern, whereas Isungset *et al.* (2021) don't find a G×SES with Norwegian data.

All the discussed studies have in common that they only test for the presence of a G×SES. They do not address the question through which mechansims such a G×SES may arise. The question of the mechanisms of a G×SES is addressed only in two studies by Tucker-Drob and Harden (2012a,b) who report a positive G×SES for academic achievement and show that this G×SES can be fully explained by a stronger realization of genes associated with different non-cognitive skills for high-SES individuals. However, especially against the heterogeneous state of research on G×SES for education where some studies confirm the SRH, some the CAH and some studies report mixed or null findings, a closer investigation of possible mechanisms giving rise to a G×SES is needed to provide a better theoretical understanding of the phenomenon.

Therefore, drawing on data from the German twin study TwinLife (Diewald *et al.* 2021), I follow Tucker-Drob and Harden (2012a,b) by investigating the mechanisms of G×SES for the secondary school track and tertiary enrolment as two central educational outcomes along the educational career. Specifically, I estimate moderated bivariate twin models which do not only allow to test for the presence of a G×SES in general, but also help to differentiate the particular genetic pathways through which such a G×SES may arise. Drawing on previous research that has consistently shown that cognitive ability is a central mediator of the genetic effect on educational outcomes (e.g. Bartels *et al.* 2002; Calvin *et al.* 2012; Johnson *et al.* 2009; Krapohl *et al.* 2014), the design allows decompose a G×SES into a moderation of the effect of genes associated with cognitive ability and a moderation of the effect of genes independent of cognitive ability. Tracing back the genetic pathways on education, thus, allows a more detailed understanding of the mechanisms of the G×SES for education.

# 2 Theoretical Framework

The aim of this section is to provide a better understanding of how the SES may affect the realization of the genetic endowment for education. In the first step, I discuss how a genetic endowment for education is realized via so-called gene-environment transactions

driven by cognitive ability. In the second step, I discuss how the SES may affect these ability-driven gene-environment transactions leading to a G×SES for education.

## 2.1 The genetic effect on education

One central finding of behavioral genetics research is that all human behavior is heritable (Turkheimer 2000). So while it is clear that education is heritable (see Branigan *et al.* 2013; Silventoinen *et al.* 2020), the question is how, i.e. through which mechanisms, the genetic endowment for education is actually realized (Tucker-Drob 2017).

An answer to this question can be derived from a transactional perspective on human development that conceptualizes individuals as active agents that shape their own environment but are also influenced by their environment (Sameroff 2009). From a behavioral genetics perspective, this idea can be used to explain the realization of the genetic endowment for education: Based on genetically influenced traits and behaviors, individuals are matched to environments which in turn have an effect on education, thus translating initial genetic differences into phenotypic educational differences[1]. This process is termed as gene-environment transaction (Tucker-Drob 2017; Tucker-Drob *et al.* 2013; Tucker-Drob and Harden 2017).

So, the process of gene-environment transaction can be decomposed into two parts: The *first part* refers to the matching of an individual into an environment based on genetically influenced traits and behaviors which is best conceptualized in terms of active and evocative gene-environment correlation (rGE, Plomin *et al.* 1977): Based on genetically influenced traits and behaviors, individuals select actively into environmental niches (active rGE) and/or evoke environmental reactions by others (evocative rGE). The *second part* refers to the effect of the experienced environment on education. This effect may be direct or indirect, e.g. involving several feedback loops between the genetically influenced trait and the environmental exposure (Dickens and Flynn 2001). So, the realization of both parts of the gene-environment transaction is a necessary condition for the realization

[1] Alternative versions of a gene are called *alleles*. While the *genotype* refers to an individual's combination of the alleles, the *phenotype* refers to the observed traits and behaviors (Knopik *et al.* 2017, p. 20)

of the the genetic endowment for education, i.e. the translation of genetic differences in a given trait into phenotypic differences in education (Tucker-Drob 2017, p. 473). One important implication of this perspective is that the realization of a genetic endowment for an outcome like education is not an immutable biological process where the genes operate directly and uniquely "through the ribosomes" but is an environmentally mediated process (Jencks 1980, p. 730; see also Freese 2008). As a consequence, a strong genetic effect or a high heritability does not imply the irrelevance of environmental influences. Quite the contrary, the realization of a genetic endowment *depends* on the opportunity to have proper environmental experiences (Jencks 1980; Scarr 1992).

What are the genetically influenced traits based on which individuals are matched via active and evocative rGE processes into educational relevant environments, i.e. what are the "driving forces" of the gene-environment transactions (Tucker-Drob and Harden 2017)? Tucker-Drob and Harden (2012a,b, 2017) argue to consider non-cognitive skills like motivations, academic self-concept or intellectual interest as possible driving forces and have empirically shown that a G×SES for academic achievement can be explained by SES differences in the effect of genes associated with different non-cognitive skills (Tucker-Drob and Harden 2012a,b).

Others argue that cognitive ability is an important driving force of the matching process into educational relevant environments (see Beam *et al.* 2015; Dickens and Flynn 2001) which is corroborated by empirical studies showing that cognitive ability is a central mediator of the genetic effect on educational outcomes (e.g. Bartels *et al.* 2002; Calvin *et al.* 2012; Johnson *et al.* 2009; Krapohl *et al.* 2014). So, since Tucker-Drob and Harden (2012a,b) already tested whether non-cognitive skills account for a G×SES for academic achievement and no study so far has tested whether a G×SES for education is accounted for by SES differences in the effect of genes associated with cognitive ability[2], in this article, I use cognitive ability as a candidate driving force to test whether SES differences

[2]Notice, that the finding of cognitive ability as a central mediator of the genetic effect on education does not necessarily imply that a G×SES for education is also accounted for by SES differences in the effect of genes associated with cognitive ability on education. It may well be that the effect of genes associated with cognitive ability does not vary by SES and the G×SES is completely accounted for by SES differences in the effect of genes independent of cognitive ability.

in ability-driven gene-environment transactions result in a G×SES for education.

How would the realization of the genetic endowment for education via ability-driven gene-environment transactions look like? The answer to this question depends on whether a person has a high or low level of genetic endowment for cognitive ability[3]. On the one hand, a high genetic endowment for cognitive ability may translate into a higher cognitive ability based on which individuals select into environments (e.g. choosing more challenging courses or spending more time in intellectually stimulating activities) or evoke environmental responses (e.g. receiving more attention by teachers and better grades) that in turn have a positive effect on education. On the other hand, a low genetic endowment for cognitive ability may be reflected in a lower cognitive ability based on which individuals select into environments (e.g. choosing less challenging courses or dropping out of school) or evoke environmental responses (e.g. receiving less attention by teachers and worse grades) that in turn have a negative effect on education. In both scenarios, initial differences in the genetic endowment for cognitive ability are translated into phenotypic differences in education via ability-driven gene-environment transactions. Based on the previous arguments, the first hypothesis can be formulated:
$H_1$ : *A part of the genetic influences on the educational outcomes can be explained by genes associated with cognitive ability.*

## 2.2 SES differences in the genetic effect on education

How may the SES facilitate or constrain these ability-driven gene-environment transactions and therefore the realization of the genetic endowment for education leading to a G×SES? In this section, I discuss two competing hypotheses about possible SES differences in the realization of the genetic endowment for education: The behavioral genetic *Scarr-Rowe hypothesis* (SRH, Rowe *et al.* 1999; Scarr-Salapatek 1971) and the sociological *compensatory advantage hypothesis* (Bernardi 2014).

[3]Here, I use a single dimensional concept of a genetic endowment ranging from low to high. Alternatively, one might think of a bi-dimensional case with two qualitatively different genetic dispositions affecting the same outcome: a positive genetic potential and a negative genetic risk. However, with the data at hand, it is not possible to differentiate nor model this bi-dimensional case. Therefore, the theoretical discussion is limited to the uni-dimensional case.

The empirical starting point of the SRH is a study conducted by Scarr-Salapatek (1971) with twin data from schools in Philadelphia (USA) where she found higher heritability estimates for cognitive ability for high-SES children than for low-SES children. The finding of stronger genetic effects on cognitive ability for individuals with a higher SES was replicated by Fischbein (1980) for Sweden and by Rowe *et al.* (1999) for the USA. The SRH can be linked to the more general G×E mechanism of *enhancement* from the G×E typology by Shanahan and colleagues (Shanahan and Boardman 2009; Shanahan and Hofer 2005) which refers to the facilitation of the realization of a positive or high genetic endowment within an enriched environment, meaning that the realization of an individual's genetic endowment is pushed "to her or his upper limit as defined by the genotype" (Shanahan and Boardman 2009, p. 222). Therefore, although the SRH originally focused on SES differences in the genetic effect on cognitive ability, the more general underlying enhancement mechanism allows the extension of the SRH to other outcomes like education. The enhancement effect should be most visible among individuals with a high genetic endowment since this group has the highest genetic limits, so that they can gain most from pushing the realization of their genetic endowment to their genetically defined upper limits.

Through which pathways might a high SES enhance the realization of a high genetic endowment for education? An answer to this question can be derived from the concept of primary SES effects (Boudon 1974) that refers to SES differences in academic skills and performance. Previous research suggests that compared to low-SES parents, high-SES parents have the necessary knowledge and resources to promote their children's skill development and learning processes by providing a cognitively stimulating environment, more relevant learning materials or engaging more in learning oriented interactions, e.g. reading books or holding discussions (cf. Cheadle 2008; Cheadle and Amato 2011; Cunha and Heckman 2008; Duncan and Magnuson 2012; for an overview see Becker 2019).

In other words, by providing a rich and cognitively stimulating environment, high-SES families facilitate the ability-driven gene-environment transactions outlined in the previous section: High ability individuals with a high SES have more opportunities to select

themselves into cognitively stimulating environments (e.g. reading books or participating in extracurricular activities) or to evoke cognitively stimulating responses (e.g. intellectually stimulating discussions) which results in a stronger link between genes associated with cognitive ability and the educational outcome leading to a G×SES for education. Thus, using cognitive ability as a candidate driving force, the SRH suggests that high-SES families enhance the realization of a high genetic endowment for cognitive ability by facilitating ability-driven gene-environment transactions leading to the following hypothesis:

*$H_2$: The SES differences in education are greatest among individuals with a high genetic endowment for cognitive ability.*

While the behavioral genetics literature on G×SES mainly concentrates on testing the SRH, the sociological literature on the intergenerational transmission of educational advantage offers an alternative perspective with the *compensatory advantage hypothesis* (CAH, Bernardi 2014). According to the CAH, high-SES parents compensate for early disadvantages of their children so that "life course trajectories of individuals from privileged backgrounds are less dependent on prior negative outcomes" (ibid., p. 75). Although the CAH was not originally formulated from a genetically informed perspective, it can be easily applied to the study of G×SES for education by conceptualizing a low genetic endowment for education as a specific form of the "prior negative outcomes" that high-SES parents try to compensate for [4]. Indeed, it is possible to match the CAH to the more general mechanism of *compensation* of the typology of G×E mechanisms by Shanahan and colleagues (Shanahan and Boardman 2009; Shanahan and Hofer 2005). Here, the realization of a disadvantageous genetic endowment or a "genetic diathesis" is constrained in an enriched environment.

Why and how do high-SES families compensate for a low genetic endowment for education? According to the theory of relative risk aversion (RRA, Breen and Goldthorpe

[4]Nielsen (2016) develops an argument that leads to the same expectation drawing on the work of Saunders (2010). However, unlike Bernardi (2014) who provides e.g. with the theory of relative risk aversion (Breen and Goldthorpe 1997) a discussion of the specific mechanisms giving rise to the expected interaction, Nielsen (2016) does not address the specific mechanisms giving rise to the expected pattern of G×SES. Therefore, my argument draws on the contribution by Bernardi (2014).

1997) educational choices are driven by the aspiration to avoid social downward mobility. In order to achieve this goal, high-SES children have to obtain higher levels of education than low-SES children where a low or intermediate educational level is sufficient to avoid social downward mobility. Thus, high-SES parents have an incentive to invest especially in their less endowed offspring since they face the highest risk of educational failure and social downward mobility and to make more ambitious educational choices irrespective of their children's ability and performance as described by the concept of secondary SES effects Boudon (1974). At the same time, they are also equipped with the resources necessary to successfully ensure the educational success for their less gifted children (Stienstra *et al.* 2021). For instance, high-SES families have the financial resources to afford private tutoring lessons (Bernardi 2014), dispose of the institutional knowledge to successfully guide their children through the difficulties of the educational system (Schulz *et al.* 2017) and are more likely to successfully challenge unfavorable decisions made by educational gatekeepers (Lareau 2011).

From a transactional perspective this means that high-SES parents restrict educational harmful gene-environment transactions for their low ability offspring, thus constraining the effect of genes associated with cognitive ability on educational outcomes. On the one hand, they may prevent their low ability children to select themselves into environmental niches (e.g. choosing a less demanding coursework or dropping out of school) or to evoke environmental responses (e.g. low educational recommendations or bad grades) that have negative educational consequences. On the other hand, high-SES parents might be able to buffer the negative educational consequences of an harmful environment already experienced, e.g. by ignoring bad grades in their educational decisions (Bernardi and Cebolla-Boado 2014; Breen and Goldthorpe 1997). So following the CAH, the compensatory practices of high-SES families restrict harmful ability-driven gene-environment transactions for their less endowed offspring, thus constraining the realization of genes associated with cognitive ability, leading to the following hypothesis:

*$H_3$: The SES differences in education are greatest among individuals with a low genetic endowment for cognitive ability.*

# 3 Data and Variables

For the empirical analysis I use data from the German twin study TwinLife (version 5.0.0, Diewald *et al.* 2021), a longitudinal study of four birth cohorts of same-sex twins and their families that started in 2014. Unlike convenience samples commonly used for twin studies, TwinLife applies a register-based probability sampling design. Previous studies have shown that the data includes families along the whole range of the socio-economic spectrum in Germany (Lang and Kottwitz 2017) and the socio-demographic structure of the twin families is comparable to non-twin families in other German probability-based surveys (Mönkediek *et al.* 2020). For the analysis I use data of the second (born in 2003-2004) and fourth (born in 1990-1993) birth cohort from the first two survey waves conducted in 2014-2016 (survey wave 1) and 2017-2019 (survey wave 2), covering an age range of 10-14 for the younger and 21-27 for the older cohort.

The German educational system is characterized by two central educational transitions: The transition into secondary schooling and tertiary education (Neugebauer and Schindler 2012, p. 22). The transition into secondary schooling is around age 10-12 so that the younger cohort is observed right after the transition into secondary schooling. As in other studies (Stienstra *et al.* 2021), the secondary school track is measured via an ordinal variable with four categories (1: Lower secondary school; 2: (Integrated) intermediate school; 3: Comprehensive school; 4: Higher secondary school)[5]. The transition into tertiary education is around age 18/19 (Neugebauer and Schindler 2012, p. 22) and the tertiary enrolment is measured by the twin's enrolment level in tertiary education. For twins that already finished tertiary education, the highest reported tertiary degree is considered. The resulting ordinal variable has three categories (1: Vocational training; 2: University of applied science; 3: University or higher). For both outcomes, the data of the second survey wave was only used if the information about the educational outcome in the first survey wave is missing.

Cognitive ability is measured in the first survey wave by a computer administrated

[5]For a short discussion of the ordinality assumption of the measurement of the secondary school track, please refer to the appendix.

version of the widely applied Culture Fair Test (CFT, Weiß 2006) that measures fluid intelligence. Compared to crystalized intelligence (i.e. learned knowledge and skills), fluid intelligence measures the ability to acquire skills and knowledge (Ortiz 2015). Sum scores of all four subdimensions of the CFT - figural reasoning (15 items), figural classification (15 items), matrices (15 items) and reasoning (11 items) - provided in the TwinLife data set were used to construct an indicator of cognitive ability by estimating a Confirmatory Factor Analysis (CFA) for each cohort. The fit of the measurement models is very good (younger cohort: RMSEA = 0.027; older cohort: RMSEA = 0.017).

SES is measured by the parental education operationalized via the CASMIN classification (Brauns *et al.* 2003). Here, again information from the second wave was used when information in the first wave was missing. For parents with different levels of education, I follow the dominance approach as outlined by Erikson (1984) considering only the highest level of education. To allow a better interpretation of the G×SES, parental education is z-standardized.

To avoid an overestimation of shared environmental influences, all twin models control for the twins' sex and age (using age from the first survey wave) (McGue and Bouchard 1984). Missing values for cognitive ability, secondary school track and tertiary enrolment are addressed with the Full Information Maximum Likelihood Estimator (FIML). Missing values on the covariates age and sex and the moderator parental education, however, need to be addressed in the analysis via listwise deletion[6]. This leads to a final sample size of 1039 twin pairs for the younger cohort and of 979 twin pairs for the older cohort. A descriptive overview of the used samples can be found in Table 1. Due to the low number of twins attending a lower secondary school track (5% of the MZ twins and 2% of the DZ twins) I conducted a robustness check of the moderation analysis for the younger cohort by merging the categories "Lower Secondary" and "(Integrated) Intermediate" together. The results of the robustness check support the reported results of the main analysis and can be found in the appendix.

[6]I use the `R` package `OpenMx` for the twin modeling, where the covariates and the moderator are specified as so called *definition variables* (cf. Mehta and Neale 2005) for which addressing missing data via FIML is not possible.

| Younger cohort | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | MZ | | | | | DZ | | | | |
| | Min | Max | Mean | SD | N | Min | Max | Mean | SD | N |
| **Cognitive Ability** | -2.68 | 2.32 | -0.03 | 0.87 | 832 | -3.02 | 2.25 | 0.02 | 0.85 | 1220 |
| **Secondary School Track** | | | | | | | | | | |
| *Lower Secondary* | 0 | 1 | 0.05 | 0.21 | 788 | 0 | 1 | 0.02 | 0.15 | 1155 |
| *(Integrated) Intermediate* | 0 | 1 | 0.30 | 0.46 | 788 | 0 | 1 | 0.22 | 0.42 | 1155 |
| *Comprehensive* | 0 | 1 | 0.19 | 0.39 | 788 | 0 | 1 | 0.23 | 0.42 | 1155 |
| *Higher Secondary* | 0 | 1 | 0.46 | 0.50 | 788 | 0 | 1 | 0.53 | 0.50 | 1155 |
| **Parental Education** | 1 | 9 | 6.66 | 2.24 | 842 | 1 | 9 | 6.97 | 2.14 | 1236 |
| **Age at $t_1$** | 10 | 12 | 11 | 0.36 | 842 | 10 | 12 | 11 | 0.28 | 1236 |
| **Sex** | | | | | | | | | | |
| *Male* | 0 | 1 | 0.45 | 0.50 | 842 | 0 | 1 | 0.50 | 0.50 | 1236 |
| *Female* | 0 | 1 | 0.55 | 0.50 | 842 | 0 | 1 | 0.50 | 0.50 | 1236 |
| **N (Twins)** | | | | | 842 | | | | | 1236 |
| **N (Twin pairs)** | | | | | 421 | | | | | 618 |
| **Older cohort** | | | | | | | | | | |
| | MZ | | | | | DZ | | | | |
| | Min | Max | Mean | SD | N | Min | Max | Mean | SD | N |
| **Cognitive Ability** | -3.25 | 1.64 | 0.07 | 0.83 | 1041 | -3.69 | 1.69 | -0.08 | 0.98 | 904 |
| **Tertiary Enrolment** | | | | | | | | | | |
| *Vocational Training* | 0 | 1 | 0.46 | 0.50 | 976 | 0 | 1 | 0.49 | 0.50 | 823 |
| *University of applied science* | 0 | 1 | 0.13 | 0.34 | 976 | 0 | 1 | 0.13 | 0.34 | 823 |
| *University or higher* | 0 | 1 | 0.41 | 0.49 | 976 | 0 | 1 | 0.38 | 0.49 | 823 |
| **Parental Education** | 1 | 9 | 5.82 | 2.34 | 1048 | 1 | 9 | 6.08 | 2.32 | 910 |
| **Age at $t_1$** | 21 | 25 | 23.06 | 0.83 | 1048 | 21 | 25 | 23.02 | 0.80 | 910 |
| **Sex** | | | | | | | | | | |
| *Male* | 0 | 1 | 0.41 | 0.49 | 1048 | 0 | 1 | 0.44 | 0.50 | 910 |
| *Female* | 0 | 1 | 0.59 | 0.49 | 1048 | 0 | 1 | 0.56 | 0.50 | 910 |
| **N (Twins)** | | | | | 1048 | | | | | 910 |
| **N (Twin pairs)** | | | | | 524 | | | | | 455 |

Table 1: The descriptive statistics are calculated for the long-formatted data (one row = one twin) before z-standardizing parental education.

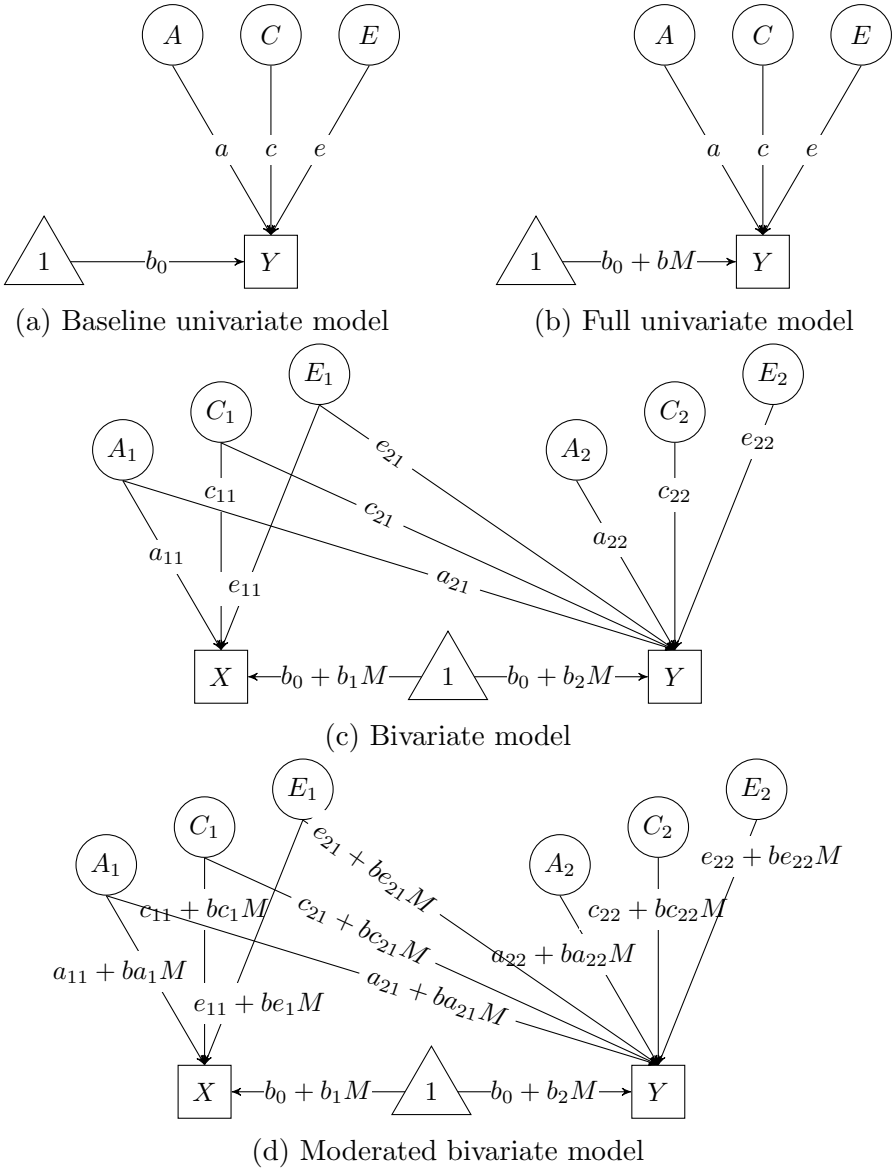
# 4 Methods

For the statistical analysis I estimate biometric twin models (Knopik *et al.* 2017; Neale and Maes 2004). Twin models build on the comparison of genetically identical monozy-

gotic twins (MZ) and dizygotic twins (DZ) that share on average 50% of their genes to decompose the variance of an outcome into three components that refer to additive genetic (A), shared environmental (C) and non shared environmental influences including the measurement error (E). The twin models are estimated as structural equation models with the ACE components being specified as latent factors with a variance fixed to 1 (Neale and Maes 2004). Both educational outcomes are ordinal variables which is properly addressed in the statistical analysis by estimating liability threshold models (c.f. ibid.). As every statistical model, the biometric twin model rests on a number of identifying assumptions. These assumptions and possible consequences of their violation are discussed in detail in the appendix. Furthermore, as shown in the appendix, the DZ correlations are not smaller than half the MZ correlations for cognitive ability and both educational outcomes which would indicate the presence of non-additive genetic effects (D). Therefore, the ACE model and not the ADE model is the right choice. All calculated models are parametrized as "path coefficients model", meaning that the *effects* and not the *variances* of the ACE factors are estimated (see ibid., p. 100).

The statistical analysis consists of three steps and a path diagram of the twin models used in each of them can be found in Fig. 1. The first analysis step consists of two univariate twin models. The baseline univariate model (see Fig. 1a) decomposes the variance of cognitive ability and the educational outcomes for both cohorts into the effects of the ACE components without controlling for parental education. While the shared environmental factor is often conceptualized as a proxy measure of SES, this is not necessarily the case (see Freese and Jao 2017). Therefore, in the full univariate model (see Fig. 1b) parental education is added as a covariate so that the comparison of the shared environmental effects between the baseline and the full univariate models allows to assess how much of the shared environmental effects is accounted for by parental education. Generally, the full univariate model already gives some first evidence about genetic, shared and non-shared environmental influences on cognitive ability and both educational outcomes.

In the second analysis step, a bivariate model (see Fig. 1c) in the Cholesky parametriza-




(a) Baseline univariate model

(b) Full univariate model

(c) Bivariate model

(d) Moderated bivariate model

Figure 1: Overview of twin models used in the analysis. Y = Educational outcome; X = Cognitive ability; M = Parental education. Note: For display reasons, only one twin per pair is shown and the control variables sex and age are not shown in the diagram.

tion (see Loehlin 1996) is estimated in order to decompose the genetic and environmental effects on the educational outcomes into effects of ACE components associated with cognitive ability ($a_{21}$, $c_{21}$ and $e_{21}$) and the effects of the ACE components independent of cognitive ability ($a_{22}$, $c_{22}$ and $e_{22}$). So, for the genetic pathway a significant effect $a_{21}$ means that genes associated with cognitive ability ($A_1$) affect the educational outcome, indicating ability-driven gene-environment transactions. A significant effect $a_{22}$ means that genes independent of cognitive ability ($A_2$) affect the educational outcome, indicating gene-environment transactions driven by other traits. Therefore, a significant $a_{21}$

would support hypothesis $H_1$ according to which a part of the genetic influences on the educational outcomes can be explained by genes associated with cognitive ability.

In the third step, a moderated bivariate twin model is estimated where the effects of all ACE components are allowed to vary by SES (see Fig. 1d). This parametrization is a bivariate adoption of the model proposed by Purcell (2002) and was used before by Tucker-Drob and Harden (2012a,b) to investigate the mechansisms of G×SES. Different to other approaches like a multigroup analysis, this parametrization has the advantage to test for G×E while controlling for rGE between the moderator and the outcome by estimating a main effect of the moderator ($b_1$ and $b_2$, see Purcell 2002)[7]. For the interpretation of the interaction effects, consider the effect of the genes associated with cognitive ability ($A_1$) on the educational outcome: $a_{21} + ba_{21} * M$. Here, $a_{21}$ represents the main effect of genes associated with cognitive ability on the educational outcome, conditional for $M = 0$, i.e. for an average parental education. $ba_{21}$ is the interaction effect which informs if and how much the effect of the genes associated with cognitive ability on the outcome varies by SES. A significant positive $ba_{21}$ would go in line with the SRH and hypothesis $H_2$ while a significant negative $ba_{21}$ supports the CAH and hypothesis $H_3$.

All twin models are calculated with `R` (version 4.1.1) using the package `OpenMx` (version 2.19.8, Neale *et al.* 2016). Only the best fitting models are reported which are obtained by dropping non-significant parameters and comparing the different models via Likelihood-Ratio tests (cf. Neale and Maes 2004). An overview of the model comparisons is provided in the appendix.

[7]Note that in the context of G×E, it is neccesary to distinguish two different rGE processes. First, there are "benign" rGE processes: As explained in section 2.1, rGE constitute the first part of the gene-environment transactions through which a genetic endowment is realized. So, a G×E may be explained by environmental differences in the opportunity for rGE. Second, there are "malign" rGE processes, i.e. a rGE between the moderator and the outcome. It is central to control for this second rGE process when estimating a G×E since any unmodelled rGE of this kind may appear as a G×E. Therefore, the Purcell model controls for rGE between the moderator and the outcome by estimating a main effect of the moderator to rule out the possibility that a G×E stems from systematic genetic variation between the groups defined by the moderator (Purcell 2002).

# 5 Results

In this section, the results of the biometric twin models are presented. First, I will discuss the results of the univariate ACE models. Then, the results of the unmoderated bivariate ACE models are presented and finally the results of the moderated bivariate ACE models are shown.

## 5.1 Univariate Analysis

The first analysis step consists of calculating univariate ACE models in order to test for genetic, shared environmental as well as non-shared environmental effects on cognitive ability and both educational outcomes. The results of the univariate models are shown in Tab. 2. While the baseline univariate model does not control for parental education, in the full univariate model parental education is added as a measured covariate. The decomposition of the variance of cognitive ability and the educational outcomes into the standardized variance components[8] based on the full univariate model is shown for both cohorts in Fig. 2.

For the younger cohort, in the baseline model there are significant effects of the ACE components on cognitive ability as well as the secondary school track. Controlling for parental education reduces the remaining shared environmental effects substantially: For cognitive ability by around 33% from 0.42 to 0.28 and for secondary school track by 23%. So while a substantial part of the shared environmental effects is accounted for by parental education, there are still unexplained shared environmental influences. As shown in Fig. 2, the heritability of cognitive ability is about 39% and most of the remaining variance is explained by non-shared environmental factors (37%), followed by parental education (13%) and shared environmental factors (11%). For the secondary school track, the heritability is substantially lower (27%), while most of the variance is explained by unmeasured shared environmental factors (42%) followed by parental

[8] All standardized variance components can be derived from the estimated path coefficients using path tracing rules or covariance algebra (cf. Neale and Maes 2004). E.g. for the full univariate model, the standardized genetic variance component, i.e. heritability ($h^2$), conditional on sex and age is calculated as: $h^2 = a^2/(a^2 + c^2 + e^2 + b^2)$.

| Younger Cohort | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Cogn. Abil. | | | | Sec. Educ. | | | |
| | Baseline | | Full | | Baseline | | Full | |
| | **Est.** | **SE** | **Est.** | **SE** | **Est.** | **SE** | **Est.** | **SE** |
| **a** | 0.53*** | 0.06 | 0.54*** | 0.05 | 0.53*** | 0.05 | 0.54*** | 0.05 |
| **c** | 0.42*** | 0.06 | 0.28** | 0.09 | 0.87*** | 0.05 | 0.67*** | 0.05 |
| **e** | 0.52*** | 0.02 | 0.52*** | 0.02 | 0.27*** | 0.02 | 0.27*** | 0.02 |
| **b** | - | - | 0.31*** | 0.02 | - | - | 0.52*** | 0.04 |
| **N** | 1039 | | 1039 | | 1039 | | 1039 | |
| **-2LL** | 4851.61 | | 4655.21 | | 3603.51 | | 3334.89 | |
| **Older Cohort** | | | | | | | | |
| | Cogn. Abil. | | | | Tert. Enrol. | | | |
| | Baseline | | Full | | Baseline | | Full | |
| | **Est.** | **SE** | **Est.** | **SE** | **Est.** | **SE** | **Est.** | **SE** |
| **a** | 0.74*** | 0.04 | 0.73*** | 0.02 | 0.71*** | 0.09 | 0.74*** | 0.09 |
| **c** | 0.25* | 0.12 | 0 | - | 0.58*** | 0.10 | 0.37* | 0.14 |
| **e** | 0.47*** | 0.02 | 0.47*** | 0.01 | 0.40*** | 0.03 | 0.39*** | 0.03 |
| **b** | - | - | 0.28*** | 0.02 | - | - | 0.43*** | 0.04 |
| **N** | 979 | | 979 | | 979 | | 979 | |
| **-2LL** | 4682.50 | | 4559.58 | | 3207.64 | | 3043.39 | |

Table 2: Results of the univariate models. Significance: $p < 0.001$ :***; $p < 0.01$ :**; $p < 0.05$ :*; $p < 0.10$ :$^{+}$. Est = Parameter estimate; SE = Standard Error. Note: Estimates of 0 without standard error (-) refer to parameters fixed to zero. Effects of covariates sex and age and intercept are not shown.

education (25%) and the non-shared environment only accounts for around 7% of the variance in secondary school track.

For the older cohort, the baseline models for cognitive ability as well as tertiary enrolment show significant effects of all ACE components. However, for cognitive ability, when controlling for parental education in the full model, the shared environmental effect disappears and can be dropped from the model, while for tertiary enrolment the shared environmental effect is reduced by 36%. This means that parental education explains the shared environmental effects on cognitive ability completely but only partially for tertiary enrolment. As shown in Fig. 2, the heritability of cognitive ability is about 64%, while the non-shared environment accounts for 27% and parental education for the remaining

9%. For tertiary enrolment, genes account for 54% of the variance, parental education for 18%, the non-shared environment for 15% and the remaining shared environment for the remaining 13%.

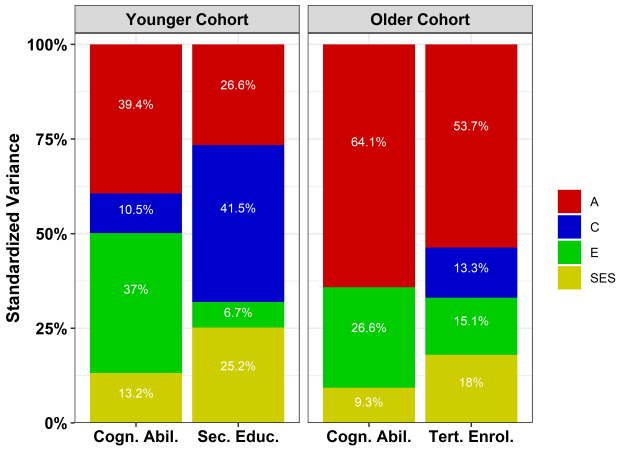



Figure 2: Variance decomposition based on the full univariate ACE models conditional on age and sex. See footnote 8 for an explanation of the calculation.

All in all, there are genetic and environmental influences on cognitive ability and the educational outcome for both cohorts. While parental education accounts for a substantial part of the shared environmental influences, there remains an unexplained part for most variables. Comparing the results between cohorts, shows that in the older cohort, genetic influences on both variables are stronger and (shared) environmental influences are weaker. However, it remains open which factors explain the genetic influences of the educational outcomes. Therefore, the second analysis step tests whether part of the genetic effect on the educational outcomes can be accounted for by genes associated with cognitive ability as assumed in the discussion of ability-driven gene-environment transactions and $H_1$.

## 5.2 Bivariate Analysis

The bivariate twin model allows to decompose the genetic and environmental effects on the educational outcomes into effects of ACE components associated with cognitive ability $(a_{21}, c_{21}, e_{21})$ and effects of ACE components independent of cognitive ability $(a_{22}, c_{22}, e_{22})$, while controlling for parental education. The model results are summarized in Tab. 3

and the variance decomposition[9] of the educational outcomes is plotted in Fig. 3.

| | Sec. Educ. | | Tert. Enrol. | |
|---|---|---|---|---|
| | **Est.** | **SE** | **Est.** | **SE** |
| $\mathbf{a_{11}}$ | $0.54^{***}$ | 0.05 | $0.73^{***}$ | 0.02 |
| $\mathbf{c_{11}}$ | $0.27^{**}$ | 0.09 | 0 | - |
| $\mathbf{e_{11}}$ | $0.52^{***}$ | 0.02 | $0.47^{***}$ | 0.01 |
| $\mathbf{a_{21}}$ | $0.23^{***}$ | 0.08 | $0.43^{***}$ | 0.05 |
| $\mathbf{c_{21}}$ | $0.25^{+}$ | 0.14 | $0.36^{**}$ | 0.13 |
| $\mathbf{e_{21}}$ | $0.07^{**}$ | 0.02 | $0.07^{*}$ | 0.03 |
| $\mathbf{a_{22}}$ | $0.50^{***}$ | 0.05 | $0.63^{***}$ | 0.09 |
| $\mathbf{c_{22}}$ | $0.63^{***}$ | 0.07 | 0 | - |
| $\mathbf{e_{22}}$ | $0.26^{***}$ | 0.02 | $0.38^{***}$ | 0.03 |
| $\mathbf{b_1}$ | $0.31^{***}$ | 0.02 | $0.28^{***}$ | 0.02 |
| $\mathbf{b_2}$ | $0.53^{***}$ | 0.04 | $0.45^{***}$ | 0.04 |
| **N** | 1039 | | 979 | |
| **-2LL** | 7865.02 | | 7432.22 | |

Table 3: Results of the bivariate models. Significance: $p < 0.001$ :$^{***}$; $p < 0.01$ :$^{**}$; $p < 0.05$ :$^{*}$; $p < 0.10$ :$^{+}$. Est = Parameter estimate; SE = Standard Error. Note: Estimates of 0 without standard error (-) refer to parameters fixed to zero. Effects of covariates sex and age and intercept are not shown.

For the younger cohort, the ACE and SES effects on cognitive ability remain virtually the same compared to the full univariate model. For the secondary school track, the bivariate model reveals two different genetic pathways: On the one hand, there is a significant effect of the genes associated with cognitive ability indicating ability-driven gene-environment transactions ($a_{21} = 0.23$, $p < 0.001$). However, genes associated with cognitive ability do not explain all the genetic effect, since on the other hand, genes independent of cognitive ability also have a (stronger) effect on the secondary school track indicating gene-environment transactions driven by other traits ($a_{22} = 0.50$, $p < 0.001$). This differentiation of the genetic pathways is reflected in Fig. 3: Nearly 17% of the heritability of the secondary school track is accounted for by genes associated with

[9]Again, the variance components can be calculated from the path coefficients following the path tracing rules or covariance algebra. For example, the total heritability of the educational outcome is calculated as: $h^2_{tot} = (a^2_{21} + a^2_{22})/(a^2_{21} + a^2_{22} + c^2_{21} + c^2_{22} + e^2_{21} + e^2_{22} + b^2_2)$. In order to see how much of the variance is accounted for by genes associated with cognitive ability, the numerator needs to be adjusted: $h^2_{ca} = (a^2_{21})/(a^2_{21} + a^2_{22} + c^2_{21} + c^2_{22} + e^2_{21} + e^2_{22} + b^2_2)$. In order to see how much of the heritability is accounted for by genes associated with cognitive ability, one needs to calculate: $p^{h^2}_{ca} = a^2_{21}/h^2_{tot}$.

cognitive ability ($100*0.046/(0.046+0.223)$), while genes independent of cognitive ability explain the remaining 83% of the heritability. Therefore, while a substantial part of the genetic influence on secondary school track can be explained by genes associated with cognitive ability, most of the genetic influences are accounted for by traits different than cognitive ability. With respect to the shared and non-shared environmental influences, the largest part is explained by factors independent of cognitive ability.

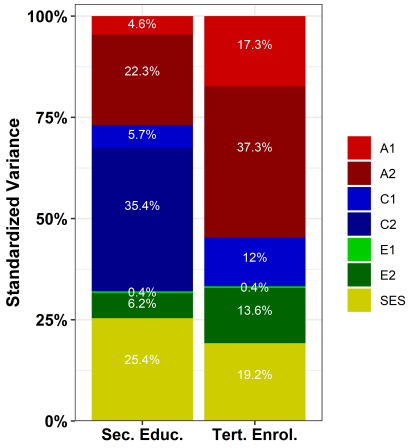



Figure 3: Decomposition of the variance of the educational outcomes based on the bivariate twin model conditional on age and sex. ACE components associated with cognitive ability (A1, C1, E1) and independent of cognitive ability (A2, C2, E2) can be distinguished. See footnote 9 for an explanation of the calculation.

For the older cohort, again the ACE and SES effects effects on cognitive ability remain stable compared to the full univariate model. Again, the genetic effect can be decomposed into the effect of genes associated with cognitive ability indicating ability-driven gene-environment transactions also for tertiary enrolment ($a_{21} = 0.43$, $p < 0.001$) and the effect of genes independent of cognitive ability ($a_{22} = 0.63$, $p < 0.001$) indicating gene-environment transactions driven by other traits. This is reflected in the heritability estimates shown in Fig. 3: 32% of the heritability of tertiary enrolment is accounted for by genes associated with cognitive ability, while the remaining 68% is accounted for by genes independent of cognitive ability. While again most of the genetic influences is due to factors different than cognitive ability, the effect of genes associated with cognitive ability is stronger than for the younger cohort. Also most of the non-shared environmental effects are due to factors independent of cognitive ability and there are no shared environmental

influences independent of cognitive ability as the parameter $c_{22}$ can be fixed to zero. However, while the shared environmental influences on cognitive ability are completely accounted for by parental education (like in the full univariate model), there are some shared environmental effects on tertiary enrolment that remain unexplained.

To sum up, according to the bivariate analysis, in line with previous reports in the behavioral genetics literature (e.g. Bartels *et al.* 2002; Calvin *et al.* 2012; Johnson *et al.* 2009; Krapohl *et al.* 2014), a considerable part of the genetic effect on both educational outcomes can be explained by genes associated with cognitive ability, thus confirming hypothesis $H_1$. However, especially for the younger cohort most of the genetic influences are due to genes independent of cognitive ability. Comparing the results between cohorts, shows that the effect of genes associated with cognitive ability is stronger for the older cohort suggesting stronger ability-driven gene-environment transactions.

## 5.3 Moderation Analysis

The third analysis step addresses the following questions: Does the realization of the genetic endowment for the educational outcomes depend on SES? If so, can the G×SES be explained by SES differences in the effect of genes associated with cognitive ability or is it accounted for by SES differences in the effect of genes independent of cognitive ability? The results of the moderated bivariate models are shown in Tab. 4 with the results for the younger cohort in the left panel and the results for the older cohort in the right panel. All G×SES are also visualized in prediction plots that are shown in Fig. 4 for the younger cohort and in Fig. 5 for the older cohort.

As can be seen in the left panel of Tab. 4, for the younger cohort all moderating paths of the genetic effects can be fixed to zero. There is only a negative moderation of the non-shared environmental effect on cognitive ability ($be_{11} = -0.03, p < 0.05$), meaning that a higher parental education is associated with a smaller non-shared environmental effect on cognitive ability. The main effects of the ACE components are comparable to those of the unmoderated bivariate twin model. So, there is no G×SES for the younger cohort - neither for cognitive ability nor for the secondary school track - which is reflected

| | Sec. Educ. | | Tert. Enrol. | |
|---|---|---|---|---|
| | **Est.** | **SE** | **Est.** | **SE** |
| $\mathbf{a_{11}}$ | $0.56^{***}$ | 0.05 | $0.72^{***}$ | 0.02 |
| $\mathbf{c_{11}}$ | $0.24^{*}$ | 0.10 | 0 | - |
| $\mathbf{e_{11}}$ | $0.52^{***}$ | 0.02 | $0.47^{***}$ | 0.01 |
| $\mathbf{ba_{11}}$ | 0 | - | $-0.13^{***}$ | 0.02 |
| $\mathbf{bc_{11}}$ | 0 | - | 0 | - |
| $\mathbf{be_{11}}$ | $-0.03^{*}$ | 0.01 | 0 | - |
| $\mathbf{a_{21}}$ | $0.22^{**}$ | 0.07 | $0.44^{***}$ | 0.05 |
| $\mathbf{c_{21}}$ | $0.28^{+}$ | 0.17 | $0.33^{*}$ | 0.14 |
| $\mathbf{e_{21}}$ | $0.07^{**}$ | 0.02 | $0.07^{*}$ | 0.03 |
| $\mathbf{ba_{21}}$ | 0 | - | $-0.14^{**}$ | 0.05 |
| $\mathbf{bc_{21}}$ | 0 | - | 0 | - |
| $\mathbf{be_{21}}$ | 0 | - | 0 | - |
| $\mathbf{a_{22}}$ | $0.50^{***}$ | 0.05 | $0.66^{***}$ | 0.09 |
| $\mathbf{c_{22}}$ | $0.62^{***}$ | 0.08 | 0 | - |
| $\mathbf{e_{22}}$ | $0.26^{***}$ | 0.02 | $0.38^{***}$ | 0.03 |
| $\mathbf{ba_{22}}$ | 0 | - | $-0.09^{+}$ | 0.06 |
| $\mathbf{bc_{22}}$ | 0 | - | 0 | - |
| $\mathbf{be_{22}}$ | 0 | - | 0 | - |
| $\mathbf{b_1}$ | $0.31^{***}$ | 0.02 | $0.28^{***}$ | 0.02 |
| $\mathbf{b_2}$ | $0.53^{***}$ | 0.04 | $0.49^{***}$ | 0.05 |
| **N** | 1039 | | 979 | |
| **-2LL** | 7859.36 | | 7385.28 | |

Table 4: Results of the moderated bivariate models. Significance: $p < 0.001$ :$^{***}$; $p < 0.01$ :$^{**}$; $p < 0.05$ :$^{*}$; $p < 0.10$ :$^{+}$. Est = Parameter estimate; SE = Standard Error. Note: Estimates of 0 without standard error (-) refer to parameters fixed to zero. Effects of covariates sex and age and intercept are not shown.

in the prediction plots shown in Fig. 4 where the constant SES gap in the predicted cognitive ability and liability of the secondary school track reflects the main effect of parental education. This SES gap, however, does not vary over the distribution of the genetic endowments associated with cognitive ability ($A_1$) and independent of cognitive ability ($A_2$) which would be the case with a G×SES. Therefore, for the younger cohort both G×SES-hypotheses, the SRH ($H_2$) and the CAH ($H_3$) need to be rejected: For the secondary school track, there are no SES differences in the realization of the genetic

endowment.

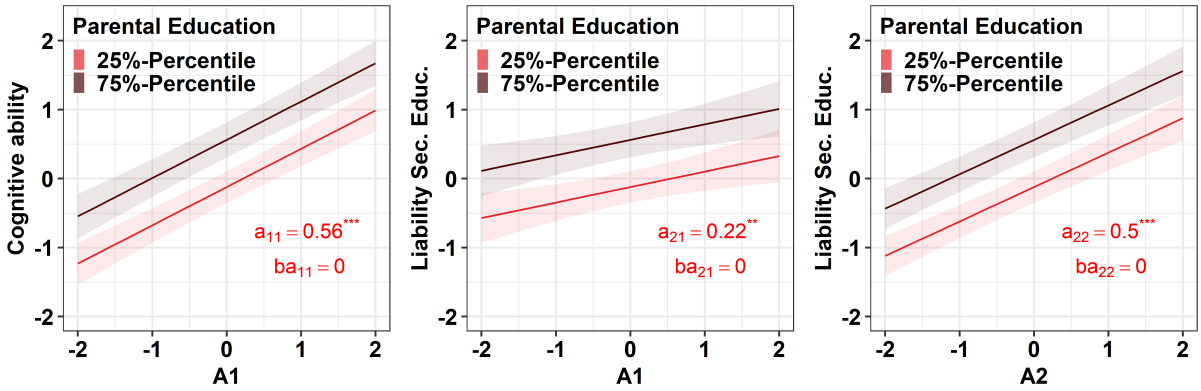



Figure 4: Prediction plots of the G×SES interactions for the younger cohort. A1 represents genes associated with cognitive ability, while A2 refers to genes independent of cognitive ability (see Fig. 1).

The right panel of Tab. 4 shows the results of the best fitting moderated bivariate model for the older cohort with tertiary enrolment as the educational outcome. There is a consistent negative moderation of all genetic effects in line with the CAH. For cognitive ability, the genetic effect for individuals with average educated parents ($a_{11}$) is about 0.72 ($p < 0.001$) and is reduced by about 0.13 ($p < 0.001$) with an increase of the parental education of 1 SD as indicated by the interaction effect $ba_{11}$. For tertiary enrolment, for individuals with parents with an average level of education, the effect of genes associated with cognitive ability ($a_{21}$) is 0.44 ($p < 0.001$) and the significant interaction effect ($ba_{21}$) suggests that an increase of the parental education of 1 SD is associated with a decrease of the genetic effect by 0.14 ($p < 0.01$). Finally, for individuals with average educated parents the effect of genes independent of cognitive ability ($a_{22}$) is 0.66 ($p < 0.001$) and the interaction effect $ba_{22}$ indicates a slight decrease of the genetic effect of about 0.09 ($p < 0.1$) for an increased parental education of 1 SD. However, since this interaction effect is only significant on a 10% level it has to be interpreted cautiously. Nevertheless, the results show that there is a G×SES for tertiary enrolment which can be decomposed into a (stronger) moderation of the effect of genetic endowments associated with cognitive ability and a (weaker) moderation of the effect of genetic endowments independent of cognitive ability.

The G×SES interactions for the older cohort are visualized in Fig. 5. Here, the com-

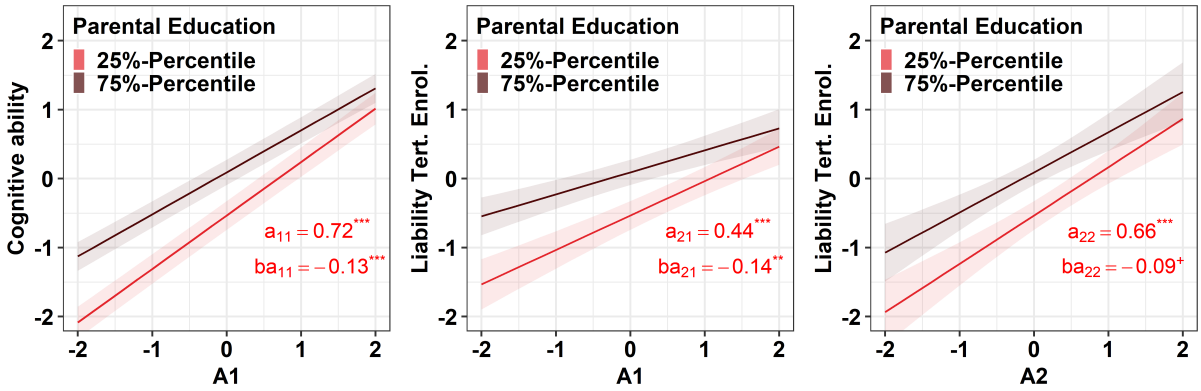



Figure 5: Prediction plots of the G×SES interactions for the older cohort. A1 represents genes associated with cognitive ability, while A2 refers to genes independent of cognitive ability (see Fig. 1).

pensatory pattern of the G×SES becomes clearly visible: high-SES parents compensate for a low genetic endowment associated cognitive ability ($A_1$) which results in a greater SES gap in the predicted cognitive ability as well as the predicted liability of tertiary enrolment among individuals with a low genetic endowment for cognitive ability (see the left and middle part of Fig. 5). In other words, for individuals with a high SES, having a low genetic endowment associated with cognitive ability is less detrimental for the cognitive development and tertiary enrolment level than for individuals with a low SES. At the same time, there is some evidence that high-SES parents compensate for disadvantageous genes independent of cognitive ability ($A_2$) which results in a greater SES gap in the predicted liability of tertiary enrolment among individuals with a low genetic endowment independent of cognitive ability (see the right part of Fig. 5). Therefore, for the older cohort, the hypothesis the SRH ($H_2$) needs to be rejected, while the results are in line with the CAH ($H_3$).

# 6 Discussion

While sociological research often reports high levels of intergenerational transmission of educational advantage (e.g. Breen and Jonsson 2005; Jackson 2013), the mechanisms leading to this result often remain unclear. In this article, I combined sociological and behavioral genetics perspectives in order to investigate whether high-SES families ensure

the educational success of their offspring by enhancing the realization of a high genetic endowment for education as suggested by the SRH (Rowe *et al.* 1999; Scarr-Salapatek 1971) or by compensating for a low genetic endowment for education as predicted by the CAH (Bernardi 2014).

Conceptually, this article contributes to the literature of G×SES for education by investigating the mechanisms of G×SES for secondary school track and tertiary enrolment as two central educational outcomes along the educational career. Using cognitive ability as a candidate driving force for the gene-environment transactions relevant for education, I provide a more detailed picture of the mechanisms through which a G×SES for education may arise. In particular, I follow the approach by Tucker-Drob and Harden (2012a,b) and calculate moderated bivariate twin models that allow to decompose a G×SES for education into a moderation of the effect of genes associated with cognitive ability and a moderation of the effect of genes independent of cognitive ability.

What are the empirical main results of this article? First, in line with previous research (see Bartels *et al.* 2002; Calvin *et al.* 2012; Johnson *et al.* 2009; Krapohl *et al.* 2014), the bivariate twin models showed that genes associated with cognitive ability account for a substantial part of the genetic influences on secondary school track and tertiary enrolment suggesting that ability-driven gene-environment transactions play a central part in the realization of the genetic endowment for both educational outcomes. This means that part of the heritability of the educational outcomes can be explained by individuals selecting themselves into educational relevant environments (e.g., more or less challenging course work, more or less cognitively stimulating activities like reading) based on their genetically influenced cognitive ability, thus translating differences in the genetic endowment for cognitive ability into phenotypic educational differences. At the same time, results show that genes associated with cognitive ability account only for 17% of the heritability of the secondary school track and 32% of the heritability of tertiary enrolment, suggesting that the largest part of the genetic influences of the educational outcomes is accounted for by genes independent of cognitive ability (see e.g. Krapohl *et al.* 2014).

Second, the moderated bivariate twin models show that there is no G×SES for the secondary school track, while the genetic effects on tertiary enrolment are negatively moderated by SES. Decomposing the genetic pathway on tertiary enrolment into the effect of genes associated with cognitive ability and the effect of genes independent of cognitive ability allowed a more detailed picture of the mechanisms of this G×SES. I found a consistently negative moderation of the genes associated with cognitive ability, suggesting that high-SES parents compensate for a low genetic endowment for cognitive ability. This results in a G×SES for cognitive ability, but also explains the greatest part of the G×SES for tertiary enrolment. Therefore, cognitive ability is not only a central mediator of the additive genetic effect on education, but also of the G×SES for education. Thus, the G×SES results for the older cohort are in line with the CAH and can be understood in terms of SES differences in the possibility for ability-driven gene-environment transactions: In order to avoid social downward mobility (Bernardi 2014; Breen and Goldthorpe 1997), high-SES parents constrain the risk for their less gifted offspring to select themselves into environments or evoke environmental responses based on their cognitive ability that have negative consequences for the tertiary enrolment. For instance, they prevent their children from selecting a less demanding coursework or dropping out of school and they may successfully intervene into the decisions of educational gatekeepers like teachers, e.g. in order to prevent bad grades (Lareau 2011). At the same time, high-SES parents may buffer the negative consequences of an educational harmful experience already made, e.g. by supporting their less endowed children to continue their academic degree despite of a poor initial performance.

Third, the SES differences in the effect of genes associated with cognitive ability on education don't completely account for the G×SES for tertiary enrolment. There is also some evidence for SES differences in the effect of genes independent of cognitive ability on tertiary enrolment. However, this finding has to be interpreted cautiously due to the significance level of $\alpha = 0.10$. Nevertheless, it can be used as a point of departure for future research to identify the particular traits and behaviors that are associated with these genes independent of cognitive ability in order to provide a better understanding

of the driving forces of a G×SES for education.

Another interesting result are the different results across cohorts, especially that I found a G×SES for tertiary enrolment but not for the secondary school track. How can these differences be explained? Clearly, with the cross-sectional data at hand, it is empirically impossible to answer this question since cohort or age effects cannot be differentiated. However, there are reasons to explain the differences from a life-course perspective. First, the birth years of the two cohorts are not so far apart that one might expect the differences in the results to be due to cohort-specific experiences. Second, rising genetic effects over the life course are a well documented phenomenon in behavioral genetics research for a range of educational relevant traits (Bartels *et al.* 2002; Bergen *et al.* 2007; Briley and Tucker-Drob 2013; Kandler and Papendick 2017; Petrill *et al.* 2004; Plomin *et al.* 1988; Zheng *et al.* 2019) which is commonly explained from a life course perspective: As individuals grow older and become more independent, they increasingly select themselves into environmental niches and evoke environmental responses based on genetically influences traits (Scarr and McCartney 1983) so that genetic differences become increasingly important over the life course. So, the differences in the G×SES results could be explained from a life-course perspective: As in younger ages genes play only a minor role, SES differences in the realization of the genetic endowments may also play a minor role. But when individuals grow older and increasingly make environmental experiences relevant for education based on genetically influenced traits and behaviors, the relevance of genes increases and SES differences in the realization of the genetic endowment also become more visible, leading more pronounced G×SES interactions in the later educational life course. However, as already meantioned, although this conclusion may be backed by empirical and theoretical arguments, it is not possible to test them with the data at hand and future research may test for a life-course graded G×SES pattern more explicitly.

This study has also some limitations: First, with cognitive ability I only consider one candidate driving force of the gene-environment transactions relevant for education. However, the results suggest that there is a substantial effect of genes independent of

cognitive ability and also a part of the G×SES for tertiary enrolment can be attributed to SES differences in the effect of genes independent of cognitive ability. Future research could investigate these unmeasured driving forces. Previous research suggests that different non-cognitive skills and personality traits may be promising candidates (Krapohl *et al.* 2014; Tucker-Drob and Harden 2012a,b) Second, with parental education I used a general proxy of overall SES which is not informative about the concrete mechanisms through which SES constrains the realization of the genetic endowments. Future research could investigate this part of the G×SES in more detail.

All in all, this study suggests that one mechanism of the intergenerational transmission of educational advantage is the ability of high-SES families to constrain unfavorable gene-environment transactions for their less endowed offspring so that their low genetic endowment for cognitive ability does not lower their probabilities of educational success.

# References


Autorengruppe Bildungsberichterstattung (2020). *Bildung in Deutschland 2020*. DE: wbv Media.

Baier, Tina and Volker Lang (2019). "The Social Stratification of Environmental and Genetic Influences on Education: New Evidence Using a Register-Based Twin Sample". In: *Sociological Science* 6, pp. 143–171. DOI: 10.15195/v6.a6.

Bartels, Meike, Marjolein J. H. Rietveld, G. Caroline M. Van Baal, and Dorret I. Boomsma (2002). "Heritability of Educational Achievement in 12-year-olds and the Overlap with Cognitive Ability". In: *Twin Research* 5.6, pp. 544–553. DOI: 10.1375/136905202762342017.

Beam, Christopher R., Eric Turkheimer, William T. Dickens, and Deborah Winders Davis (2015). "Twin Differentiation of Cognitive Ability Through Phenotype to Environment Transmission: The Louisville Twin Study". In: *Behavior Genetics* 45.6, pp. 622–634. DOI: 10.1007/s10519-015-9756-0.

Becker, Birgit (2019). "Early educational inequality: growing up in different learning environments". In: *Research Handbook on the Sociology of Education*. Edward Elgar Publishing, pp. 233–252. DOI: 10.4337/9781788110426.00023.

Bergen, Sarah E., Charles O. Gardner, and Kenneth S. Kendler (2007). "Age-Related Changes in Heritability of Behavioral Phenotypes Over Adolescence and Young Adulthood: A Meta-Analysis". In: *Twin Research and Human Genetics* 10.3, pp. 423–433. DOI: 10.1375/twin.10.3.423.

Bernardi, Fabrizio (2014). "Compensatory Advantage as a Mechanism of Educational Inequality: A Regression Discontinuity Based on Month of Birth". In: *Sociology of Education* 87.2, pp. 74–88. DOI: 10.1177/0038040714524258.

Bernardi, Fabrizio and Héctor Cebolla-Boado (2014). "Previous School Results and Social Background: Compensation and Imperfect Information in Educational Transitions". In: *European Sociological Review* 30.2, pp. 207–217. DOI: 10.1093/esr/jct029.

Bleidorn, Wiebke, Anke Hufer, Christian Kandler, Christopher J. Hopwood, and Rainer Riemann (2018). "A Nuclear Twin Family Study of Self–Esteem". In: *European Journal of Personality* 32.3, pp. 221–232. DOI: 10.1002/per.2136.

Boudon, Raymond (1974). *Education, Opportunity, and Social Inequality. Changing Prospects in Western Society*. New York.

Branigan, A. R., K. J. McCallum, and J. Freese (2013). "Variation in the Heritability of Educational Attainment: An International Meta-Analysis". In: *Social Forces* 92.1, pp. 109–140. DOI: 10.1093/sf/sot076.

Brauns, Hildegard, Stefani Scherer, and Susanne Steinmann (2003). "The CASMIN Educational Classification in International Comparative Research". In: *Advances in Cross-National Comparison*. Ed. by Jürgen H.P. Hoffmeyer-Zlotnik and Christof Wolf. New York: Springer US.

Breen, Richard and John H. Goldthorpe (1997). "Explaining Educational Differentials: Towards a formal rational action theory". In: *Rationality and Society* 9.3, pp. 275–305. DOI: 10.1177/104346397009003002.

Breen, Richard and Jan O. Jonsson (2005). "Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility". In: *Annual Review of Sociology* 31.1, pp. 223–243. DOI: 10.1146/annurev.soc.31.041304.122232.

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". In: *Psychological Science* 24.9, pp. 1704–1713. DOI: 10.1177/0956797613478618.

Calvin, Catherine M. *et al.* (2012). "Multivariate genetic analyses of cognition and academic achievement from two population samples of 174,000 and 166,000 school children". In: *Behavior Genetics* 42.5, pp. 699–710. DOI: 10.1007/s10519-012-9549-7.

Cheadle, Jacob E. (2008). "Educational Investment, Family Context, and Children's Math and Reading Growth from Kindergarten Through the Third Grade". In: *Sociology of Education* 81.1, pp. 1–31. DOI: 10.1177/003804070808100101.

Cheadle, Jacob E. and Paul R. Amato (2011). "A Quantitative Assessment of Lareau's Qualitative Conclusions About Class, Race, and Parenting". In: *Journal of Family Issues* 32.5, pp. 679–706. DOI: 10.1177/0192513X10386305.

Conley, Dalton, Emily Rauscher, Christopher Dawes, Patrik K. E. Magnusson, and Mark L. Siegal (2013). "Heritability and the Equal Environments Assumption: Evidence from Multiple Samples of Misclassified Twins". In: *Behavior Genetics* 43.5, pp. 415–426. DOI: 10.1007/s10519-013-9602-1.

Conley, Dalton *et al.* (2015). "Is the Effect of Parental Education on Offspring Biased or Moderated by Genotype?" In: *Sociological Science* 2, pp. 82–105. DOI: 10.15195/v2.a6.

Cunha, Flavio and James J. Heckman (2008). "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation". In: *Journal of Human Resources* 43.4, pp. 738–782. DOI: 10.1353/jhr.2008.0019.

Dickens, William T. and James R. Flynn (2001). "Heritability estimates versus large environmental effects: The IQ paradox resolved." In: *Psychological Review* 108.2, pp. 346–369. DOI: 10.1037/0033-295X.108.2.346.

Diewald, Martin *et al.* (2021). *TwinLifeTwinLife*. DOI: 10.4232/1.13747.

Domingue, Benjamin W., Daniel W. Belsky, Dalton Conley, Kathleen Mullan Harris, and Jason D. Boardman (2015). "Polygenic Influence on Educational Attainment: New Evidence From the National Longitudinal Study of Adolescent to Adult Health". In: *AERA Open* 1.3, p. 233285841559997. DOI: 10.1177/2332858415599972.

Duncan, Greg J. and Katherine Magnuson (2012). "Socioeconomic status and cognitive functioning: moving from correlation to causation: Socioeconomic status and cognitive functioning". In: *Wiley Interdisciplinary Reviews: Cognitive Science* 3.3, pp. 377–386. DOI: 10.1002/wcs.1176.

Erikson, Robert (1984). *Social Class of Men, Women and Families - Robert Erikson, 1984*.

Erola, Jani, Hannu Lehti, Tina Baier, and Aleksi Karhula (2021). "Socioeconomic Background and Gene–Environment Interplay in Social Stratification across the Early Life Course". In: *European Sociological Review*, jcab026. DOI: 10.1093/esr/jcab026.

Figlio, David N., Jeremy Freese, Krzysztof Karbownik, and Jeffrey Roth (2017). "Socioeconomic status and genetic influences on cognitive development". In: *Proceedings of the National Academy of Sciences* 114.51, pp. 13441–13446. DOI: 10.1073/pnas.1708491114.

Fischbein, Siv (1980). "IQ and social class". In: *Intelligence* 4.1, pp. 51–63. DOI: 10.1016/0160-2896(80)90006-9.

Freese, Jeremy (2008). "Genetics and the Social Science Explanation of Individual Outcomes". In: *American Journal of Sociology* 114 (S1), S1–S35. DOI: 10.1086/592208.

Freese, Jeremy and Yu-Han Jao (2017). "Shared Environment Estimates for Educational Attainment: A Puzzle and Possible Solutions: Shared Environment Estimates for Education". In: *Journal of Personality* 85.1, pp. 79–89. DOI: 10.1111/jopy.12226.

Giangrande, Evan J. *et al.* (2019). "Multivariate analysis of the Scarr-Rowe interaction across middle childhood and early adolescence". In: *Intelligence* 77, p. 101400. DOI: 10.1016/j.intell.2019.101400.

Gottschling, J. *et al.* (2019). "Socioeconomic status amplifies genetic effects in middle childhood in a large German twin sample". In: *Intelligence* 72, pp. 20–27. DOI: 10.1016/j.intell.2018.11.006.

Grasby, Katrina L., William L. Coventry, Brian Byrne, and Richard K. Olson (2019). "Little Evidence That Socioeconomic Status Modifies Heritability of Literacy and Numeracy in Australia". In: *Child Development* 90.2, pp. 623–637. DOI: 10.1111/cdev.12920.

Gualtieri, C. Thomas (2013). "Husband-Wife Correlations in Neurocognitive Test Performance". In: *Psychology* 04.10, pp. 771–775. DOI: 10.4236/psych.2013.410109.

Hanscombe, Ken B. *et al.* (2012). "Socioeconomic Status (SES) and Children's Intelligence (IQ): In a UK-Representative Sample SES Moderates the Environmental, Not Genetic, Effect on IQ". In: *PLOS ONE* 7.2, e30320. DOI: 10.1371/journal.pone.0030320.

Harden, K. Paige *et al.* (2020). "Genetic associations with mathematics tracking and persistence in secondary school". In: *npj Science of Learning* 5.1, pp. 1–8. DOI: 10.1038/s41539-020-0060-2.

Isungset, Martin *et al.* (2021). *Social and Genetic Effects on Educational Performance in Early Adolescence*. w28498. Cambridge, MA: National Bureau of Economic Research, w28498. DOI: 10.3386/w28498.

Jackson, Michelle (2013). "Introduction". In: *Determined to Succeed?* Ed. by Michelle Jackson. Stanford University Press, pp. 1–33. DOI: 10.11126/stanford/9780804783026.003.0001.

Jencks, Christopher (1980). "Heredity, environment, and public policy reconsidered". In: *American Sociological Review* 45.5, pp. 723–736.

Johnson, Wendy, Ian J. Deary, and William G. Iacono (2009). "Genetic and Environmental Transactions Underlying Educational Attainment". In: *Intelligence* 37.5, pp. 466–478. DOI: 10.1016/j.intell.2009.05.006.

Kandler, Christian and Michael Papendick (2017). "Behavior genetics and personality development". In: *Personality Development Across the Lifespan*. Elsevier, pp. 473–495. DOI: 10.1016/B978-0-12-804674-6.00029-6.

Knopik, Valerie S., Jenae M. Neiderhiser, J. C. DeFries, and Robert Plomin (2017). *Behavioral genetics*. Seventh edition. New York: Worth Publishers, Macmillan Learning. 508 pp.

Krapohl, Eva *et al.* (2014). "The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence". In: *Proceedings of the National Academy of Sciences* 111.42, pp. 15273–15278. DOI: 10.1073/pnas.1408777111.

Lang, Volker and Anita Kottwitz (2017). "The sampling design and socio-demographic structure of the first wave of the TwinLife panel study: A comparison with the Microcensus". In: *TwinLife Technical Report Series* 3.

Lareau, Annette (2011). *Unequal childhoods: class, race, and family life.* 2nd ed., with an update a decade later. Berkeley: University of California Press. 461 pp.

Lin, Meng-Jung (2020). "The social and genetic inheritance of educational attainment: Genes, parental education, and educational expansion". In: *Social Science Research* 86, p. 102387. DOI: 10.1016/j.ssresearch.2019.102387.

Loehlin, John C. (1996). "The Cholesky approach: A cautionary note". In: *Behavior Genetics* 26.1, pp. 65–69. DOI: 10.1007/BF02361160.

McGue, Matt and Thomas J. Bouchard (1984). "Adjustment of twin data for the effects of age and sex". In: *Behavior Genetics* 14.4, pp. 325–343. DOI: 10.1007/BF01080045.

Mehta, Paras D. and Michael C. Neale (2005). "People are variables too: Multilevel structural equations modeling." In: *Psychological Methods* 10.3, pp. 259–284. DOI: 10.1037/1082-989X.10.3.259.

Meulemann, Heiner and Ilona Relikowski (2016). "Chancengleichheit und Leistung im sozialen Wandel. Primäre und sekundäre Effekte der sozialen Herkunft am Gymnasialübergang in Hessen 1969 und 2007". In: *Zeitschrift für Erziehungswissenschaft* 19.2, pp. 443–462. DOI: 10.1007/s11618-015-0654-2.

Mönkediek, Bastian, Wiebke Schulz, Harald Eichhorn, and Martin Diewald (2020). "Is there something special about twin families? A comparison of parenting styles in twin and non-twin families". In: *Social Science Research* 90, p. 102441. DOI: 10.1016/j.ssresearch.2020.102441.

Neale, Michael C. and Hermine Maes (2004). *Methodology for Genetic Studies of Twins and Families.* Dordrecht: Kluwer Academics.

Neale, Michael C. *et al.* (2016). "OpenMx 2.0: Extended Structural Equation and Statistical Modeling". In: *Psychometrika* 81.2, pp. 535–549. DOI: 10.1007/s11336-014-9435-8.

Neugebauer, Martin and Steffen Schindler (2012). "Early transitions and tertiary enrolment: The cumulative impact of primary and secondary effects on entering university in Germany". In: *Acta Sociologica* 55.1, pp. 19–36. DOI: 10.1177/0001699311427747.

Nielsen, François (2016). "The Status-Achievement Process: Insights from Genetics". In: *Frontiers in Sociology* 1. DOI: 10.3389/fsoc.2016.00009.

Ortiz, Samuel O. (2015). "CHC Theory of Intelligence". In: *Handbook of Intelligence.* Ed. by Sam Goldstein, Dana Princiotta, and Jack A. Naglieri. New York, NY: Springer New York, pp. 209–227. DOI: 10.1007/978-1-4939-1562-0_15.

Papageorge, Nicholas W and Kevin Thom (2020). "Genes, Education, and Labor Market Outcomes: Evidence from the Health and Retirement Study". In: *Journal of the European Economic Association* 18.3, pp. 1351–1399. DOI: 10.1093/jeea/jvz072.

Petrill, Stephen A. *et al.* (2004). "Genetic and Environmental Contributions to General Cognitive Ability Through the First 16 Years of Life." In: *Developmental Psychology* 40.5, pp. 805–812. DOI: 10.1037/0012-1649.40.5.805.

Plomin, Robert, J. C. DeFries, and J. C. Loehlin (1977). "Genotype-environment interaction and correlation in the analysis of human behavior". In: *Psychological Bulletin* 84.2, pp. 309–322.

Plomin, Robert, John C. DeFries, and David W. Fulker (1988). *Nature and Nurture during Infancy and Early Childhood*. 1st ed. Cambridge University Press. DOI: 10.1017/CBO9780511527654.

Purcell, Shaun (2002). "Variance Components Models for Gene–Environment Interaction in Twin Analysis". In: *Twin Research* 5.6, pp. 554–571. DOI: 10.1375/136905202762342026.

Rijsdijk, F. V. (2002). "Analytic approaches to twin data using structural equation models". In: *Briefings in Bioinformatics* 3.2, pp. 119–133. DOI: 10.1093/bib/3.2.119.

Rowe, David C., Kristen C. Jacobson, and Edwin J. C. G. Van den Oord (1999). "Genetic and Environmental Influences on Vocabulary IQ: Parental Education Level as Moderator". In: *Child Development* 70.5, pp. 1151–1162. DOI: https://doi.org/10.1111/1467-8624.00084.

Sameroff, Arnold, ed. (2009). *The transactional model of development: How children and contexts shape each other*. Washington: American Psychological Association. DOI: 10.1037/11877-000.

Saunders, Peter (2010). *Social mobility myths*. London: CIVITAS. 171 pp.

Scarr, Sandra (1992). "Developmental Theories for the 1990s: Development and Individual Differences". In: *Child Development* 63.1, p. 1. DOI: 10.2307/1130897.

Scarr, Sandra and K. McCartney (1983). "How people make their own environments: a theory of genotype greater than environment effects". In: *Child Development* 54.2, pp. 424–435. DOI: 10.1111/j.1467-8624.1983.tb03884.x.

Scarr-Salapatek, Sandra (1971). "Race, social class, and IQ". In: *Science* 174.4016, pp. 1285–1295. DOI: 10.1126/science.174.4016.1285.

Schulz, Wiebke, Reinhard Schunck, Martin Diewald, and Wendy Johnson (2017). "Pathways of Intergenerational Transmission of Advantages during Adolescence: Social Background, Cognitive Ability, and Educational Attainment". In: *Journal of Youth and Adolescence* 46.10, pp. 2194–2214. DOI: 10.1007/s10964-017-0718-0.

Schwartz, Christine R. (2013). "Trends and Variation in Assortative Mating: Causes and Consequences". In: *Annual Review of Sociology* 39.1, pp. 451–470. DOI: 10.1146/annurev-soc-071312-145544.

Shanahan, Michael J. and Jason D. Boardman (2009). "Genetics and Behavior in the Life Course". In: *The Craft of Life Course Research*. Ed. by Glen H. Elder and Janet Z. Giele. New York: The Guilford Press.

Shanahan, Michael J. and Scott M. Hofer (2005). "Social Context in Gene–Environment Interactions: Retrospect and Prospect". In: *The Journals of Gerontology: Series B* 60 (Special_Issue_1), pp. 65–76. DOI: 10.1093/geronb/60.Special_Issue_1.65.

Silventoinen, Karri *et al.* (2020). "Genetic and environmental variation in educational attainment: an individual-based analysis of 28 twin cohorts". In: *Scientific Reports* 10.1, p. 12681. DOI: 10.1038/s41598-020-69526-6.

Spengler, Marion *et al.* (2018). "Does the heritability of cognitive abilities vary as a function of parental education? Evidence from a German twin sample". In: *PLOS ONE* 13.5. Ed. by Jelte M. Wicherts, e0196597. DOI: 10.1371/journal.pone.0196597.

Stienstra, Kim, Ineke Maas, Antonie Knigge, and Wiebke Schulz (2021). "Resource Compensation or Multiplication? The Interplay between Cognitive Ability and Social Origin in Explaining Educational Attainment". In: *European Sociological Review* 37.2, pp. 186–200. DOI: 10.1093/esr/jcaa054.

Trejo, Sam *et al.* (2018). "Schools as Moderators of Genetic Associations with Life Course Attainments: Evidence from the WLS and Add Health". In: *Sociological Science* 5, pp. 513–540. DOI: 10.15195/v5.a22.

Tucker-Drob, Elliot M. (2017). "Motivational factors as mechanisms of gene-environment transactions in cognitive development and academic achievement". In: *Handbook of competence and motivation: Theory and application, 2nd ed.* New York, NY, US: The Guilford Press, pp. 471–486.

Tucker-Drob, Elliot M. and Timothy C. Bates (2016). "Large Cross-National Differences in Gene × Socioeconomic Status Interaction on Intelligence". In: *Psychological Science* 27.2, pp. 138–149. DOI: 10.1177/0956797615612727.

Tucker-Drob, Elliot M., Daniel A. Briley, and K. Paige Harden (2013). "Genetic and Environmental Influences on Cognition Across Development and Context". In: *Current Directions in Psychological Science* 22.5, pp. 349–355. DOI: 10.1177/0963721413485087.

Tucker-Drob, Elliot M. and K. Paige Harden (2012a). "Intellectual Interest Mediates Gene × Socioeconomic Status Interaction on Adolescent Academic Achievement: **Intellectual** Interest and G×E". In: *Child Development* 83.2, pp. 743–757. DOI: 10.1111/j.1467-8624.2011.01721.x.

— (2012b). "Learning motivation mediates gene-by-socioeconomic status interaction on mathematics achievement in early childhood". In: *Learning and Individual Differences* 22.1, pp. 37–45. DOI: 10.1016/j.lindif.2011.11.015.

— (2017). "A Behavioral Genetic Perspective on Non-Cognitive Factors and Academic Achievement". In: *Genetics, Ethics and Education*. Ed. by Elena L. Grigorenko, Mei Tan, Stephen R. Latham,

and Susan Bouregy. Current Perspectives in Social and Behavioral Sciences. Cambridge: Cambridge University Press, pp. 134–158. DOI: 10.1017/9781316340301.007.

Turkheimer, Eric (2000). "Three Laws of Behavior Genetics and What They Mean". In: *Current Directions in Psychological Science* 9.5, pp. 160–164. DOI: 10.1111/1467-8721.00084.

Turkheimer, Eric, Andreana Haley, Mary Waldron, Brian D'Onofrio, and Irving I. Gottesman (2003). "Socioeconomic Status Modifies Heritability of IQ in Young Children". In: *Psychological Science* 14.6, pp. 623–628. DOI: 10.1046/j.0956-7976.2003.psci_1475.x.

Uchikoshi, Fumiya and Dalton Conley (2021). "Gene-environment interactions and school tracking during secondary education: Evidence from the U.S." In: *Research in Social Stratification and Mobility*, p. 100628. DOI: 10.1016/j.rssm.2021.100628.

Weiß, Rudolf H. (2006). *Grundintelligenztest Skala 2*. Göttingen: Hogrefe.

Woodley of Menie, Michael A., Jonatan Pallesen, and Matthew A. Sarraf (2018). "Evidence for the Scarr–Rowe Effect on Genetic Expressivity in a Large U.S. Sample". In: *Twin Research and Human Genetics* 21.6, pp. 495–501. DOI: 10.1017/thg.2018.63.

Zheng, Anqing *et al.* (2019). "Genetic and Environmental Influences on Achievement Goal Orientations Shift with Age". In: *European Journal of Personality* 33.3, pp. 317–336. DOI: 10.1002/per.2202.

# A Appendix

## A.1 Measuring the secondary school track

In the German educational system, traditionally three school tracks are differentiated: The upper secondary track ("Gymnasium"), the intermediate track ("Realschule") and the lower track ("Hauptschule"). However, comprehensive schools ("Gesamtschule") that integrate all three traditional tracks became increasingly popular. However, there are at least two reasons to rank the comprehensive school between the upper and intermediate track in the assumed ordinality of the measurement of secondary track ($upper > comprehensive > intermediate > lower$): 1) A reason for the ordinality assumption of $upper > comprehensive$ is that the upper secondary school ("Gymnasium") is still linked with a higher prestige (Meulemann and Relikowski 2016, p. 450). 2) A reason for the ordinality assumption of $upper > comprehensive > intermediate$ is the probability of obtaining the highest school degree ("Abitur") conditional on the visited track.

While it is not possible to obtain the Abitur in the intermediate track, in 2018 25% graduated from the comprehensive school with an Abitur, while 85% graduated from the upper track with an Abitur (Autorengruppe Bildungsberichterstattung 2020), therefore: $P(Abitur|Upper) > P(Abitur|Comprehensive) > P(Abitur(Intermediate)$.

## A.2 Model assumptions

The identification of the biometric twin model rests on the following assumptions (Rijsdijk 2002, pp. 129 sqq.): 1) The equal environment assumption (EEA) states that the environmentally caused similarity is the same for MZ and DZ twins. This assumption is the reason why the correlations between the shared environmental factors of the same trait is fixed to 1 for both MZ and DZ twins ($Cor_{MZ}(C_1, C_2) = Cor_{DZ}(C_1, C_2) = 1$). A violation of the EEA in the form of a more similar treatment of MZ twins would lead to an overestimation of the genetic effects and an underestimation of the shared environmental effects. The assumption has been tested for various traits and seems to be reasonable (Knopik *et al.* 2017, p. 86) and there is also accumulating evidence that the results of the twin models are quite robust against a violation of the EEA since it does not change the parameter estimates substantially (see e.g. Conley *et al.* 2013; Mönkediek *et al.* 2020). 2) There is no assortative mating with respect to the studied phenotypes. Research suggests that there is substantial assortative assortative mating for education and cognitive ability (Gualtieri 2013; Schwartz 2013), which is why this assumption is likely violated leading to an underestimation of the genetic effects (Rijsdijk 2002, p. 130). Therefore, the estimates of the genetic effects should be considered as lower bounds or "conservative" estimates.

Further, it is assumed that 3) there are no gene-environment interactions or correlations. A unmodelled A×C would inflate the genetic effects, while an unmodelled A×E would act like E. A correlation between A and C acts like C and between A and E like E (Purcell 2002). However, since I conduct a G×E analysis I explicitly test for the assumption of no G×E and per design I control for rGE between the predictor and the educational outcomes and by controlling for the main effect of the moderator, I also

control for rGE between parental education and both variables in the model.

Finally, it is assumed that 4) there are only additive genetic effects and thus no interactions of alleles of the same gene locus (dominance) or different gene loci (epistasis). One indicator for non-additive genetic effects is a DZ correlation being smaller than half the MZ correlation ($Cor_{DZ}(X_1, X_2) < 0.5 * Cor_{MZ}(X_1, X_2)$) (cf. Bleidorn *et al.* 2018, pp. 222 sq.). As displayed in the next section, the DZ correlations are not smaller than half the MZ correlations for any of the variables studied.

## A.3 Twin correlations

| | **Younger Cohort** | | **Older Cohort** | |
|---|---|---|---|---|
| | **MZ** | **DZ** | **MZ** | **DZ** |
| **Cogn. Abil.** | $0.63^{***}$ | $0.41^{***}$ | $0.63^{***}$ | $0.38^{***}$ |
| **Sec. Educ.** | $0.93^{***}$ | $0.77^{***}$ | - | - |
| **Tert. Enrol.** | - | - | $0.84^{***}$ | $0.59^{***}$ |

Table 5: Twin correlations for cognitive ability and the educational outcomes. Significance: $p < 0.001$ :$^{***}$; $p < 0.01$ :$^{**}$; $p < 0.05$ :$^{*}$; $p < 0.10$ :$^{+}$. For secondary school track and tertiary enrolment, polychoric correlations are shown. For cognitive ability Pearson correlations are shown.

## A.4 Model comparisons

### A.4.1 Baseline univariate twin models

For the baseline univariate twin models, no model comparisons were needed since there were no non-significant parameters in the full model.

### A.4.2 Full univariate twin models

For the full univariate twin models, only the model of cognitive ability for the older cohort contained non-significant parameters. Here the shared environmental effect is eliminated:

### A.4.3 Bivariate twin models

For the bivariate models, in the younger cohort no model comparisons were needed since there were no non-significant parameters in the full model. The following table refers to

| Baseline | Comparison | No. Par. | -2LL | df | AIC | Δ 2-LL | p |
|---|---|---|---|---|---|---|---|
| Full ACE | - | 7 | 4559.58 | 1938 | 4573.58 | - | - |
| Full ACE | Drop c11 | 6 | 4559.58 | 1939 | 4571.58 | 0 | p = 1 |

Table 6: Model comparisons: Full univariate twin model for cognitive ability in the older cohort

the older cohort:

| Baseline | Comparison | No. Par. | -2LL | df | AIC | Δ 2-LL | p |
|---|---|---|---|---|---|---|---|
| Baseline | - | 17 | 7431.84 | 3727 | 7465.84 | - | - |
| Baseline | Drop c11, c22 | 15 | 7432.22 | 3729 | 7462.22 | 0.38 | p = 0.83 |

Table 7: Model comparisons: Bivariate twin model of older cohort

### A.4.4 Moderated bivariate twin models

| Baseline | Comparison | No. Par. | -2LL | df | AIC | Δ 2-LL | p |
|---|---|---|---|---|---|---|---|
| Full ACE | - | 27 | 7850.20 | 3968 | 7904.20 | - | - |
| Full ACE | Drop bc11, bc21, bc22 | 24 | 7853.27 | 3971 | 7901.27 | 3.06 | p = 0.38 |
| Drop bc11, bc21, bc22 | Drop ba11, ba21, ba22 | 21 | 7857.33 | 3974 | 7899.33 | 4.06 | p = 0.25 |
| Drop ba11, ba21, ba22 | Drop be21, be22 | 19 | 7859.36 | 3976 | 7897.36 | 2.04 | p = 0.36 |

Table 8: Model comparisons: Moderated bivariate twin model of younger cohort

| Baseline | Comparison | No. Par. | -2LL | df | AIC | Δ 2-LL | p |
|---|---|---|---|---|---|---|---|
| Full ACE | - | 26 | 7381.54 | 3718 | 7433.54 | - | - |
| Full ACE | Drop bc11, bc21, bc22 | 23 | 7382.40 | 3721 | 7428.40 | 0.86 | p = 0.83 |
| Drop bc11, bc21, bc22 | Drop be11, be21, be22 | 20 | 7385.25 | 3724 | 7425.25 | 2.85 | p = 0.42 |
| Drop be11, be21, be22 | Drop c11, c22 | 18 | 7385.28 | 3726 | 7421.28 | 0.03 | p = 0.98 |

Table 9: Model comparisons: Moderated bivariate twin model of older cohort

## A.5 Robustness check

| Baseline | Comparison | No. Par. | -2LL | df | AIC | Δ 2-LL | p |
|---|---|---|---|---|---|---|---|
| Full ACE | - | 26 | 7517.73 | 3969 | 7569.73 | - | - |
| Full ACE | Drop bc11, bc21, bc22 | 23 | 7521.91 | 3972 | 7567.91 | 4.18 | p = 0.24 |
| Drop bc11, bc21, bc22 | Drop ba11, ba21, ba22 | 20 | 7526.08 | 3975 | 7566.08 | 4.17 | p = 0.24 |
| Drop ba11, ba21, ba22 | Drop be21, be22, c21 | 17 | 7530.38 | 3978 | 7564.38 | 4.3 | p = 0.23 |

Table 10: Model comparisons: Moderated bivariate twin model of younger cohort (Robustness check)

| | **Secondary School Track** | |
|---|---|---|
| | **Est.** | **SE** |
| **a11** | 0.57*** | 0.05 |
| **c11** | 0.20$^{+}$ | 0.12 |
| **e11** | 0.52*** | 0.02 |
| **ba11** | 0 | - |
| **bc11** | 0 | - |
| **be11** | -0.03* | 0.01 |
| **a21** | 0.33*** | 0.05 |
| **c21** | 0 | - |
| **e21** | 0.05* | 0.02 |
| **ba21** | 0 | - |
| **bc21** | 0 | - |
| **be21** | 0 | - |
| **a22** | 0.45*** | 0.06 |
| **c22** | 0.68*** | 0.06 |
| **e22** | 0.22*** | 0.02 |
| **ba22** | 0 | - |
| **bc22** | 0 | - |
| **be22** | 0 | - |
| **b1** | 0.31*** | 0.02 |
| **b2** | 0.53*** | 0.04 |
| **N** | 1039 | |
| **-2LL** | 7530.38 | |

Table 11: Results of the best fitting moderated bivariate twin models for the robustness check. Significance: $p < 0.001$ :***; $p < 0.01$ :**; $p < 0.05$ :*; $p < 0.10$ :$^{+}$. Est = Point estimate of parameter; SE = Standard Error. Note: Estimates of 0 without standard error (-) indicates parameters fixed to zero. Effects of covariates sex and age are not shown.