Educational quality and inequality

The interplay between schools, families, and genes

Kim Stienstra

Educational quality and inequality

The interplay between schools, families, and genes

Kim Stienstra

Educational quality and inequality: The interplay between schools, families, and genes
Stienstra, K.

ISBN 978-90-393-7607-2
DOI 10.33540/2045

Printing Ridderprint | www.ridderprint.nl
Layout and design Indah Hijmans | persoonlijkproefschrift.nl
Cover design Sanne Stienstra | www.artbysanny.com

Educational quality and inequality

The interplay between schools, families, and genes

Onderwijskwaliteit en ongelijkheid

De wisselwerking tussen scholen, gezinnen en genen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 19 januari 2024 des middags te 2.15 uur

door

Linnsey Kimberly Stienstra

geboren op 11 augustus 1995
te Harderwijk

Promotor:
Prof. dr. I. Maas

Copromotor:
Dr. A. Knigge

Beoordelingscommissie:
Dr. E. van Bergen
Prof. dr. M.M. Jæger
Prof. dr. M.H.D. van Leeuwen
Prof. dr. D.L. Oberski
Prof. dr. I. de Wolf

Dit proefschrift werd (mede) mogelijk gemaakt door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) Talent Grant [nummer 406.18.557]. Dit proefschrift is afgedrukt met financiële steun van de J.E. Jurriaanse Stichting.

Table of Contents

CHAPTER 1

Examining the interplay between schools, families, and genes: a synthesis

Kim Stienstra is the sole author of this chapter.

This chapter benefited from valuable feedback provided by Ineke Maas and Antonie Knigge.

1.1 Introduction

Education increasingly determines who gets ahead in society (Noord et al., 2021; Tannock, 2008). Children who perform well in school are more likely to continue their education and obtain higher degrees, which in turn has numerous social and economic benefits. These include, among others, higher income, higher occupational status, better health, and increased civic engagement for individuals, and economic growth and social cohesion at the societal level (OECD, 2017; UNESCO, 2018). Because of this, the importance of education has been emphasized in international conventions and development agendas (UNESCO, 2018). For example, the Sustainable Development Goal on quality education calls for action so that all countries by 2030 "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (United Nations, 2015). This requires ongoing attention for both developing and developed countries.

In a developed country as the Netherlands, many aspects of education seem to be going well. Nearly all children attend primary education, and the Netherlands ranks among the highest in terms of compulsory instruction hours among OECD countries (OECD, 2014). Additionally, most schools meet the minimal statutory requirements for basic quality based on the assessments by the Dutch Inspectorate of Education, which is part of the ministry of education (Inspectie van het Onderwijs, 2017). Nevertheless, the quality and equality of education are under pressure. There are indications that Dutch primary education does not succeed in making full use of the potential of pupils (Inspectie van het Onderwijs, 2021a). National and international research reveals a decline in the average performance levels of Dutch pupils, with only a small proportion of pupils achieving high levels in reading and mathematics (De Wolf, 2023; Inspectie van het Onderwijs, 2020; Swart et al., 2023). Additionally, pupils' performance levels increasingly depend on the socioeconomic status (SES) of their parents (De Wolf, 2023; Inspectie van het Onderwijs, 2016). Moreover, for pupils' educational outcomes it matters which school they attend. There seem to be large differences between schools in pupils' performance, even if schools have similar pupil populations (Bolhaar & Scheer, 2019; Inspectie van het Onderwijs, 2017). Recent developments put further pressure on the quality of education. The COVID-19 pandemic led to a delay in learning, and widened the differences between families and between schools (Engzell et al., 2021; Haelermans, Korthals, et al., 2022; Haelermans, van Wetten, et al., 2022). Also, the growing shortage of teachers and school leaders has a negative impact on educational quality (De Wolf, 2023).

Similar situations apply to other countries, including other high-income countries. Despite the increase in access and duration of schooling in the past decades, too many children go through primary school without learning enough foundational skills including literacy and numeracy skills (UNICEF, 2020). Based on data from 2015-2019, 10 to 30% of the children in most high-income countries did not meet the minimum required proficiency in reading at the end of lower-secondary school (United Nations, 2019). Additionally, international assessments show that there is a socioeconomic achievement gap which increased in numerous countries between 1964 and 2015 (Chmielewski, 2019). The COVID-19 pandemic worsened the existing problems in

many countries: children's learning progress delayed, especially among children from low-SES backgrounds (Betthäuser et al., 2023).

Schools play a dual role in explaining educational performance and socioeconomic differences therein. First, schools could be part of the problem. There are between-school differences in performance, also when pupils' background characteristics are taken into account (Rjosk, 2022; Thrupp et al., 2002). Studies indicate that children who attend schools of a lower quality (e.g., presence of less effective teachers, more disruptive climate, and low expectations) perform less well (Hanushek & Rivkin, 2006; Opdenakker & Van Damme, 2007; Scheerens & Bosker, 1997; Thapa et al., 2013). If there are large quality differences between schools, children can develop and show their educational potential in some schools while they face barriers to do so in other schools. There are also reasons to expect that schools contribute to the SES gap in educational performance. Children from high-SES backgrounds more often attend higher-quality schools (Borghans et al., 2015a; Robert, 2010). They may also benefit more from their school environment. High-SES children enter the school with better academic preparation, allowing them to reap greater rewards from good learning opportunities than their lower-SES counterparts (Hanselman, 2018).

Second, schools could also be part of the solution. Higher-quality schools have, for example, better and more experienced teachers, high expectations, and an orderly climate (Scheerens & Bosker, 1997). Such a school environment may positively affect the performance of all pupils. But it may also be especially advantageous for those pupils who are less likely to find stable, stimulating, and resourceful learning environments at home (Hanselman, 2018; Rumberger & Palardy, 2005). Learning opportunities within schools overlap with those within socioeconomically advantaged families and may substitute for each other. Therefore, schools may reduce SES differences in performance.

It is thus important to study the role of schools to understand how performance differences are produced and can potentially be reduced. Investigating the role of schools in (SES gaps in) educational performance has been a core focus in sociological research, especially since the Coleman Report (Coleman et al., 1966). Despite the numerous studies, it remains unclear whether for educational inequality schools are part of the problem or of the solution (Downey & Condron, 2016). Two reasons contribute to the unclarity on this matter.

The first reason relates to the conceptual question of what educational inequality entails. The term 'inequality' is used in different ways and there are different perspectives on which sources of differences are part of educational inequality and problematized as such. First, there is the mere existence of differences in performance, which has been labeled 'inequality as dispersion' (Van de Werfhorst & Mijs, 2010). Many find dispersion of educational performance quite acceptable (Ferreira & Gignoux, 2014). The variation in performance levels between pupils is even needed because the labor market demands differently skilled workers (Strello et al., 2021). However, dispersion in performance may be more problematic if it has consequences for health outcomes and civic participation, for example. Despite whether dispersion in itself is seen as problematic, dispersion results from different sources that can be seen as more or less problematic. Dispersion in performance does not completely reflect individual choices or actions of pupils, but also circumstances beyond their control (Ferreira & Gignoux, 2014). This relates to

the second way how 'inequality' is used, namely, to refer to differences in performance along the lines of ascribed characteristics. Most notably, differences by children's SES background, which has been commonly viewed as unfair differences and has been labeled 'inequality of opportunity' or 'social inequality' (Strello et al., 2021; Strietholt, 2014; Van de Werfhorst & Mijs, 2010). But for performance differences along the lines of other characteristics there is less consensus on to what extent it is problematic.

The school children go to is, just as their SES background, a circumstance beyond children's control. The Dutch Inspectorate of Education also warned about school differences in educational outcomes, implying that school differences are problematic (Inspectie van het Onderwijs, 2017). But it is unclear if school differences are still seen as problematic if it turns out that schools play a compensatory role and reduce the SES gap in performance, for example. Differences in performance related to genetic endowment are also debatable. On the one hand, genetic differences reflect differences in innate ability. The influence of innate ability on performance is often seen as legitimate (Dias Pereira, 2021; Tannock, 2008). If there are no social barriers that hinder the expression of innate talent, a greater proportion of performance will be explained by genetic differences. Therefore, the genetic contribution to educational performance can be seen as an indication of opportunity for achievement (Nielsen, 2006). On the other hand, individuals cannot control their innate ability any more than their social background. Being lucky in the genetic lottery can therefore also be seen as an unfair source of advantage, and inconsistent with equality of opportunity (Harden, 2021; Roemer, 1998).[1] Moreover, genetic endowment does not only reflect positive influences of genetic potential on educational performance. It also involves genetically influenced characteristics that negatively affect performance (i.e., genetic risks), such as behavioral problems and health issues (Krapohl et al., 2014). If the genetic contribution to education performance is larger in higher-quality schools because genetic potential is more realized in these schools, this can be seen as higher-quality schools providing more equal opportunities. However, if the genetic contribution in high-quality schools is larger because genetic risks are more expressed, most would agree that this is inconsistent with equality of opportunity.

The second reason for why it is still unclear whether schools increase or decrease educational inequality is that only a limited number of studies investigate all three aspects: schools, families, and genes. Even if there is agreement what influences on educational performance are part of educational inequality, it is still an empirical question how schools, families, and genes together play a role. Not studying the role of schools, families, and genes simultaneously is problematic for two reasons. First, the influences may be mistaken for each other because they are correlated. For example, parents who performed well in school themselves did so partly because of their genetic endowment. Parents partly transmit these genes to their children, but also tend to provide a stimulating learning environment at home and tend to choose higher-quality schools for their

1 As both family background influences and genetic influences can be seen as unfair, some argue that only differences resulting from effort (e.g., motivation, interest, good study habits) are justified. Yet, this is also not without problems (see Swift & Marshall, 1997). Effort is partly shaped by family background and genetic predisposition, implying that effort also reflects influences beyond an individual's control (Meyer, 2016; Swift & Marshall, 1997).

children. Second, the different factors may interact with each other. For example, high-quality schools may reduce family background differences in performance, but could simultaneously increase genetic differences between children. Prior studies tell only part of the story because they ignore the interaction with either genetic or environmental factors.

In this dissertation, I investigate three reasons why children differ in their educational performance (i.e., standardized test scores, grade-point-average): differences in children's school environment, differences in children's family background, and genetic differences between children. I ask:

How does the interplay between schools, families, and genes shape educational inequality?

While sociological research considers the influences of schools, families, and their interplay on educational outcomes, the role of genetic differences between pupils is often not considered. Behavioral genetics research investigates genetic differences in educational outcomes and its dependency on the family environment, but only to a lesser extent the role of the school environment. Moreover, these studies tend to overlook the complexity and diversity of environmental conditions (Baier, 2019). Behavioral genetics studies examine how genetic influences are dependent on either the family or school environment, but seldom consider that children are simultaneously embedded in both families and schools. As a result, it remains largely unknown how schools, families, and genes operate together. By combining sociological theories with methods and models from behavioral genetics, I aim to gain a better understanding of the interplay between school characteristics, family SES, and genes in shaping educational performance and its implications for educational inequality. While my focus is on the theoretical and empirical part of this question, I also discuss the conceptual part since one's definition of educational inequality has consequences for the interpretation of the results.

1.2 Theoretical perspectives

1.2.1 Environmental and genetic influences on performance

Family influences

Between-family differences, most notably related to families' SES, affect children's educational performance (Sirin, 2005). Two complementary social mechanisms can be distinguished to explain the association between parental SES and children's educational performance. The first mechanism relates to the resources available within the family. High-SES families are more able to invest in children's educational success, because parents' SES provides access to financial, cultural, and social resources that parents can invest in or transmit to their children (Blau & Duncan, 1967; Bourdieu & Passeron, 1977). Second, families differ in their motivations and ambitions. According to the Relative Risk Aversion theory, high-SES parents have more incentives to invest in the educational careers that lead to higher degrees than parents with a low social standing, because they want to avoid that their children will be downwardly mobile (Breen & Goldthorpe, 1997). Indeed, empirical research has shown that family SES is an important predictor of educational achievement and attainment (see, e.g., Breen & Jonsson, 2005; Sirin, 2005).

School influences

Similarly to the family environment, socialization takes place in the school environment, and some schools are a more advantageous environment for educational achievement than others (Parcel & Dufur, 2001). There are various ways through which the school environment could affect educational performance. The first way is via *resources*. School resources refer to aspects that can (potentially) be bought either directly (e.g., educational materials) or more indirectly (e.g., educational time, pupil-to-teacher ratio) (Grubb, 2009; Hofflinger & Von Hippel, 2020). Aspects related to *climate* provide another way. School climate characteristics refer to aspects that cannot readily be bought and are more difficult to change. These include norms, values, and expectations (e.g., academically oriented culture, high expectations), relationships (e.g., cohesion), and larger organizational structures (e.g., educational leadership) (Cohen et al., 2009; Grubb, 2009). Additionally, the *quality of teachers* is important. On the classroom level, teachers play a key role in the educational performance of pupils via teaching and learning practices (e.g., structured instruction, differentiation), but also via their contribution to a favorable educational climate (e.g., having high expectations of pupils, creating a safe and orderly classroom climate) (Hanushek & Rivkin, 2006). On the school level, a large share of high-quality teachers may contribute to a shared achievement-oriented culture (cf. Bosker & Scheerens, 1994). Also, the lack of high-quality teachers due to teacher shortage or high turnover rates may contribute to an instable school environment (Ingersoll, 2001). Lastly, the *composition* of the pupil population may play a role. Partly, the socioeconomic composition of the pupil body exerts an influence via other school characteristics. For example, high-SES schools may more easily attract good and experienced teachers and have more rigorous curricula (Armor et al., 2018; Sykes & Kuyper, 2013). On top of that, there could also be an influence of the SES composition via peer interactions (Armor et al., 2018). For instance, high-SES students with higher aspirations, better study habits, and less disruptive classroom behavior may have a positive influence on the performance of their peers (Gutiérrez, 2023).

Genetic influences

Differences in educational performance also reflect genetic differences between children. DNA is passed on from parents to offspring. Hence, children's genetic makeup consists of two independently inherited DNA sequences, one from the mother and one from the father (Fagerness & Nyholt, 2008). Humans are largely similar in their genetic makeup. The DNA sequence between any two individuals is 99.9% identical (Collins & Mansoura, 2001). It is the small part of 0.1% of genetic variation that contributes to differences in observable characteristics of an individual (i.e., phenotypes) such as educational performance (Fagerness & Nyholt, 2008). Performing well in school relies on many cognitive skills (e.g., intelligence, working memory, attention) as well as non-cognitive skills (e.g., conscientiousness, perseverance, self-control). Brain activity is an important cause of the development of these skills, which in turn depends on neural functioning and neural connectivity. Brain formation and the functioning of its neurons develop under certain genetic programs, meaning that genes exert at least some influence on the cognitive and non-cognitive functions involved in learning (Bueno, 2019). To quantify how much of the variation in an outcome among people in a certain population is related to genetic variation among them, the statistical

concept 'heritability' is used. This is often estimated using the classical twin design (see Section 1.3). Behavioral genetics studies applying this design show that educational performance is substantially heritable, both in the early school years and at later stages of compulsory education (De Zeeuw et al., 2016; Krapohl et al., 2014). The study by Krapohl et al. (2014) shows that the heritability in educational performance can largely be explained by genetically influenced cognitive skills (general intelligence) and non-cognitive characteristics (personality, behavioral problems, self-efficacy, and well-being). While people often mistakenly believe that genetic effects do not change after conception, the opposite is true. Genetic factors become increasingly important over the life course (Knopik et al., 2016). For example, it has been found that the heritability of cognitive ability is lower in infancy than in middle childhood and adolescence (Plomin & Spinath, 2004). The most likely explanation for this is the presence of gene-environment correlations (Knopik et al., 2016), which I will discuss next.

1.2.2 Correlations between schools, families, and genes

The influences of the schools, families, and genes on educational performance are not independent but are associated with each other in different ways. Genetic influences are not fixed and do not operate in a vacuum. Instead, they are expressed in correlation and interaction with the environment (Sameroff, 2009; Tucker-Drob et al., 2013). The behavioral genetics literature identifies three types of gene-environment correlation (rGE), that is, passive, evocative, and active gene-environment correlations (Knopik et al., 2016; Plomin et al., 1977). A *passive gene-environment correlation* refers to the association between the genotype that a child inherits from their parents and the environment the child is raised in (Hart et al., 2021). For example, when children have parents with a high cognitive ability, they inherit genes that are favorable to the development of cognitive skills but are also typically raised in an intellectually stimulating environment conducive to the full development of cognitive abilities (Plomin et al., 1977). These children are passively exposed to a social environment that fits their genetic predisposition. It is called passive because it occurs independently of the observed characteristics or activities of the child, in contrast to the other two types of rGE (Plomin et al., 1977). An *evocative gene-environment correlation* exists when genetically influenced characteristics bring about specific reactions from others in the environment. For example, highly talented children may be recognized by their parents and teachers as such, and accordingly may receive an enriched environment to maximize their talents (Diewald et al., 2015; Plomin et al., 1977). This is an example of a positive evocative rGE, but the correlation could also be negative. For example, children with low reading ability may receive more tutoring and increased literacy exposure from their parents and/or teachers (Pennington et al., 2009). An *active gene-environment correlation* is present when individuals actively seek and select themselves into environments conducive to their genotype. For example, children with a high cognitive ability may seek peers or certain activities that foster their cognitive growth (Plomin et al., 1977). Also here a negative correlation is possible, but this is often not considered by researchers (Carey, 2003). An example would be children with a genetic liability to low cognitive ability seek out help with learning themselves, which has the potential to increase children's performance and reduce the expression of the underlying genetic predisposition.

Besides these gene-environment correlations, there is also an environment-environment correlation, namely between the family and school environment. Higher-SES parents more often choose higher-quality schools for their children than low-SES parents (Borghans et al., 2015a; Robert, 2010). This selection effect into schools leads to a positive association between family SES and school quality. Better educational outcomes of children attending higher-quality schools may therefore not reflect the influences of the school environment but influences of the family. These influences of the family can both be environmental (i.e., parental investment of time and resources in children's education) and genetic, since parents do not only pass on educationally relevant environments but also genes. Hence, children in the same school are both more socially and genetically similar. This dissertation covers compulsory schooling, where school enrollment is (at least in the countries that I study) not restricted by ability or performance on an entrance test. Moreover, children are too young to influence their own school choice. Therefore, it is unlikely that children are selected into schools based on genetically influenced traits (i.e., evocative rGE) or that children actively self-select in schools that match their genotype (i.e., active rGE) (cf. Smith-Woolley et al., 2018). Instead, it is in this case more likely that a correlation between children's genotype and the quality of their school operates via parents' SES.

1.2.3 Interactions between schools, families, and genes

Multiplication

The influences of schools, families, and genes are not only correlated with each other, they could also interact with each other in predicting educational performance. One way is by strengthening each other's influence on educational performance. I refer to this as *multiplication*, but it has also been referred to as a Matthew effect, accumulative advantage, enhancement, or amplification (e.g., DiPrete & Eirich, 2006; Erola & Kilpi-Jakonen, 2017). Multiplication can take place between two environmental influences (e.g., an advantaged school environment is more beneficial for high-SES children), but also between genetic and environmental influences (e.g., positive genetic influence is stronger for children from high-SES families and in high-quality schools).

In the sociological literature, different arguments are provided for why the family and school environment may strengthen each other in predicting educational outcomes. The cumulative nature of skill formation where skills beget skills, suggests that children who gained more skills early in life (e.g., in high-SES home environments) may increasingly gain more during school (Heckman, 2000). This could be because they benefit more from the school resources and academic climate in high-quality schools (Hanselman, 2018). For example, high-SES children may learn more from instructional materials because they enter school with more language skills (cf. Bradley & Corwyn, 2002). Additionally, the cultural correspondence between the home and school environment for high-SES children may play a role. From a cultural reproduction perspective (Bourdieu, 1986; Bourdieu & Passeron, 1977), it can be expected that the high-status cultural signals of high-SES children (e.g., behaviors, tastes, and attitudes) are positively evaluated by teachers and that these children may experience a greater sense of belonging in their class, leading to better educational performance (De Graaf et al., 2000). This may be especially the case in higher-quality schools, where the educational environment is more ambitious and academically

oriented and coincides with high-SES parents' expectations and ambitions for educational success.

The behavioral genetics literature provides insights on how genetic and environmental influences may strengthen each other. The bioecological model suggests that enriched environments exacerbate the influence of genetic potential (Bronfenbrenner, 1992; Bronfenbrenner & Ceci, 1994). Enriched environments, such as high-SES families and high-quality schools, are more resourceful and stable and have higher levels of proximal processes. Proximal processes are enduring forms of (reciprocal) interactions characterized by increasing complexity. For example, interactive language practices or guided play activities between parents (or teachers) and children (Hadley et al., 2023). Therefore, it can be derived that in higher-SES families and higher-quality schools, the influence of genetic potential is stronger (Haughbrook et al., 2017; Rowe et al., 1999; Scarr-Salapatek, 1971).

Compensation

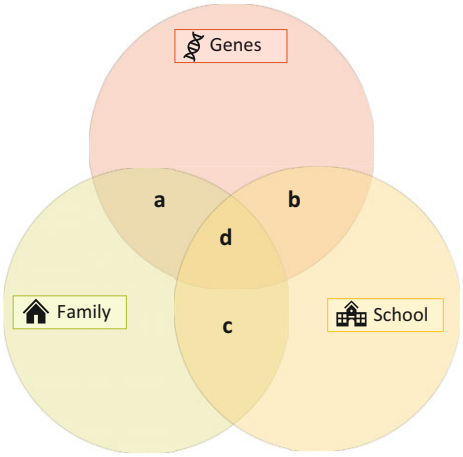
Influences of schools, families, and genes could also compensate or substitute for each other. Again, this could apply to two environmental aspects (e.g., a higher-quality school environment having a stronger influence for lower levels of parental SES) or to genetic and environmental influences (e.g., a higher-SES environment compensating for genetic influences that lower educational performance). A reason to expect such a negative interaction between the family and school environment, is that learning opportunities within schools overlap with those within socioeconomically advantaged families and can substitute for each other (Chiu & Khoo, 2005; Hanselman, 2018). Lower-quality school environment may be less harmful to high-SES students because the fewer learning opportunities in such schools can be substituted by parental resources (e.g., providing tutoring), while low-SES parents cannot provide such compensation (Hanselman, 2018). In the same way, low-SES children may benefit more from high-quality schools because these schools provide environmental inputs (e.g., resources, academic climate, higher levels of motivation and aspiration) that these children are less likely to find at home (Coleman et al., 1966; Rumberger & Palardy, 2005).

If there is a compensation interaction between genes and the environment, genetic influence on educational performance becomes weaker the higher the SES of the family and/or the higher the quality of the school. According to the diathesis-stress model, the realization of a diathesis (i.e., genetic vulnerability) is more likely when the level of environmental risks and stressors is higher (Rende & Plomin, 1992; Shanahan & Hofer, 2005). The absence of stressors can neutralize the realization of genetic vulnerability that would otherwise lead to lower educational performance. Moreover, the presence of positive features in the environment, such as in high-SES families and high-quality schools, could compensate for the expression of genetic risks (see Shanahan & Hofer, 2005). This gene-environment interaction pattern aligns with the sociological compensatory advantage mechanism. This mechanism suggests that prior negative outcomes (e.g., health and cognitive endowments at birth, poor school performance) are less consequential for the educational performance of children from higher-SES backgrounds (Bernardi, 2014). Given that these prior outcomes are partly genetically influenced, high-SES parents may compensate for the realization of genetic risks. Put differently, children from higher-SES families would perform well

in school irrespective of their genetic potential (Lin, 2020). Hence, the influence of genes can be expected to be weaker in higher-SES families and higher-quality schools.

1.2.4 An overview

Figure 1.1 shows the conceptual model of the influences of schools, families, and genes. Their overlap consists of gene-environment correlations and interactions. Evocative and active gene-environment correlations are thought to be pathways through which environments enhance high genetic potential for educational performance (or compensate for low genetic potential or genetic risks) (Tucker-Drob et al., 2013). That is, these gene-environment correlations underly gene-environment interactions. The presence of adequate opportunities in the environment influences gene-environment correlations. Family SES, but also school quality, can be seen as indicators of the quality of environmental opportunity (cf. Tucker-Drob et al., 2013). For example, in high-SES families, children with high genetic potential for education may evoke more cognitively stimulating responses (e.g., challenging educational materials), because high-SES parents are more likely to have the necessary knowledge and resources to provide this (Ruks, 2022).[2] While the interdependency between genes and environments could reflect both correlations and interactions, I only empirically investigate the interactions given the model and data constraints.



a	Gene-family environment correlation E.g., active rGE: children with a high reading ability may be more likely to pick up a book, which are more often present in high-SES families Gene-family environment interaction E.g., multiplication: genetic potential is more strongly expressed in high SES families
b	Gene-school environment correlation E.g., evocative rGE: especially high-quality teachers recognize children's mathematics difficulties and provide them with additional exercises to practice Gene-school environment interaction E.g., compensation: genetic risks are mitigated in high quality schools
c	Family-school environment correlation E.g., high-SES parents choose high-quality schools Family-school environment interaction E.g., high-quality school compensates for low-SES background
d	Gene-family-school environment interaction E.g., gene-family SES interaction is less pronounced in high-quality schools, because the school environment can substitute the family environment

Figure 1.1 Conceptual model of the influence of schools, families, and genes, and their overlap (a-d) which consists of gene-environment correlations (rGE) and interactions.

2 As Ruks (2022) points out, this can also be derived from the sociological literature on primary effects of SES (Boudon, 1974) which discusses how SES affects educational skills and performance.

1.3 Methodological approach

1.3.1 Classical Twin Design

In this dissertation, I utilize a quantitative genetics approach, leveraging its established tradition to unravel the interplay of genetic and environmental factors in educational outcomes (e.g., Behrman & Taubman, 1989; Eckland, 1967; Scarr-Salapatek, 1971; Taubman, 1976). More specifically, I study the interplay between families, schools, and genes using different applications of the Classical Twin Design (CTD).[3] This design is based on comparing the twin correlations of an outcome (in my case, educational performance) for identical (i.e., monozygotic; MZ) and fraternal (dizygotic; DZ) twins reared together in the same family (Knopik et al., 2016). MZ twins share all of their genes at conception, while DZ twins share on average half of their segregating genes.[4] Therefore, if MZ twins are more similar in their educational performance (as indicated by a greater twin correlation) than DZ twins, this implies that genetic influences are at play. Based on this idea, and further assumptions (see Table 1.1), the total variance in educational performance can be decomposed into three components capturing variance associated with additive genetic influences (*A*), shared environmental (between-family) influences (*C*), and non-shared (within family) influences that also include measurement error (*E*) (see Figure 1.2). More technical details on the CTD are described in Appendix A.

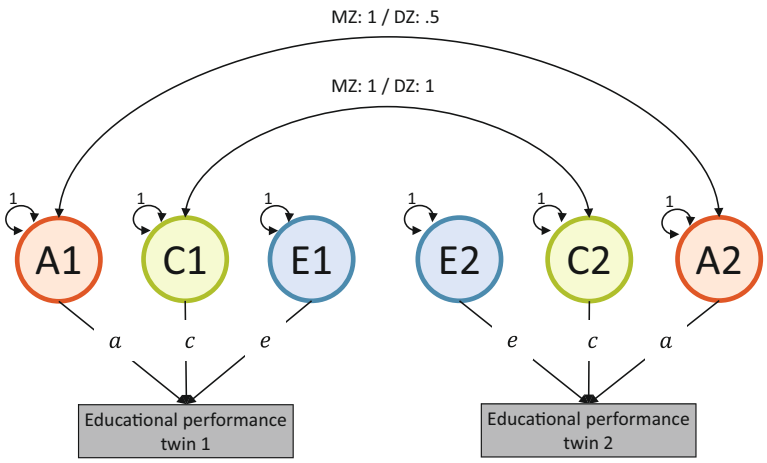


Figure 1.2 The Classical Twin Design.

3 There are other designs to study the interplay between schools, families, and genes that I do not consider in this dissertation. These include other quantitative genetic designs that apply the same idea as the CTD and use information of individuals who differ in genetic and/or environmental similarity to infer genetic and environmental effects (e.g., adoption design, pedigree design) (see, e.g., Knopik et al., 2016). Alternatively, molecular data can be used to identify specific genes related to an outcome. For example, polygenic scores (PGSs) can be constructed that summarize the associations between genetic variants and an outcome (see, e.g., Domingue et al., 2020).

4 DZ twins share half of their *segregating* genes because all humans are 99.9% identical in their genetic makeup (Collins & Mansoura, 2001)

Table 1.1 provides an overview of the assumptions underlying the CTD, bias if the assumption is violated, evidence on the violation likelihood, and consequences of violation. Importantly, my primary focus is not on univariate performance decomposition, but on comparative investigation of *ACE* components across various levels of family SES, school quality, and over time. The potential violation of most assumptions is unlikely to depend on family SES, school quality, or grade. Loehlin et al. (2009) have also shown that interaction effect estimates are less affected by assumption violations compared to main effect estimates.

Having a clear understanding of the information that can be obtained from the *ACE* components as derived from the CTD, as well as their limitations, is essential. The CTD conceptualizes genetic and environmental influences as unobserved latent variables, instead of the specific influences of thousands of genes and of the many aspects of children's family background, for example. While this approach avoids the need to measure all these influences, it also creates black boxes. For instance, the *A* component indicates a link between genetic differences and educational performance but does not provide details about the number or nature of these genetic influences (see Medland & Hatemi, 2009). Yet, this is also not the goal of the CTD, as the focus of the CTD is on understanding the variation in an outcome, similar to analysis of variance (ANOVA) (Verhulst & Hatemi, 2013).

Focusing on the variance of educational performance also implies that there would be limited shared environmental variance if all children are raised in equally supportive families and attend equally good schools. Genetic differences between children would then largely account for the variance (Asbury & Plomin, 2014). In that case, supportive families and good schools increase *average* educational performance but do not explain much of the *differences* between children, because these environments are rather universal. A small shared environmental component thus does not mean that the environment does not matter.

The environment could play a role in more complex ways than modeled in the CTD. Human characteristics like educational performance are too complex to perfectly divide into genetic and environmental components. The CTD should be seen as a baseline model providing an initial decomposition of individual differences (Verhulst & Hatemi, 2013). More complexities, including the consequences of assortative mating and gene-environment correlations and interactions, are confounded with the variance components estimates (see Table 1.1). Such complexities can be modeled by extending the CTD, as I will describe in the next section (1.3.2 Extensions of the CTD). The design remains a simplification of reality, even with extensions. Therefore, the estimates of the *ACE* components should be interpreted with their limitations in mind and should not be seen as definitive predictive estimates (Verhulst & Hatemi, 2013).

Table 1.1 Assumptions of the Classical Twin Design (CTD).[a]

Possible bias if violated	Evidence
Equal environments assumption (EEA) – MZ and DZ twins share their environment to the same extent	
If MZ twins have a more similar environment than DZ twins: overestimation of *A* and underestimation of *C* (the opposite is also possible, but unlikely)	▪ Different research designs provided results that are inconsistent with violation of the EEA. ▪ First, if MZ twins are treated more similar, DZ twins who are misclassified by parents as MZ should be more similar than correctly classified DZ twins (and vice versa). Little or no effect of misclassification was found (Conley et al., 2013). ▪ Second, if twins look more alike, this often leads to misperception of their zygosity, which in turn may lead to different treatment and a violation of the EEA. If greater physical resemblance leads to greater similarity in an outcome after controlling for zygosity, this indicates that the EEA is violated. For educational performance, this has not been found, suggesting no violation of the EEA (Mönkediek, 2021). ▪ Third, one could directly measure if MZ twins are treated differently by their parents than DZ twins. Parenting styles generally did not differ between MZ and DZ twins (Mönkediek et al., 2020), neither did the effects of different parenting dimensions on school grades (Mönkediek, 2021). ▪ Moreover, even if MZ are treated more similar than DZ twins, this is not necessarily problematic. For example, MZ twins are more likely than DZ twins to share a room and to dress alike, but this is uncorrelated with twin similarity in educational related outcomes (Evans & Martin, 2000). Even if more similar treatment is related to twin similarity in performance, this relationship should be strong in order to lead to a meaningful overestimation of *A* (Mönkediek, 2021). Moreover, various studies suggest that the more similar treatment of MZ twins results from their shared genetic makeup and the similar responses it evokes from the environment, rather than their greater phenotypic similarity or expectations of the social environment (see Evans & Martin, 2000). Hence, it reflects gene-environment correlations (rGE) which are part of the twin model interpretation if this is unmodeled (see the discussion on rGE later in this table) ▪ *ACE* estimates of educational outcomes based on other designs than the CTD, that do not rely on the EEA, are very similar. These include, e.g., designs based on pedigree data (Schwabe et al., 2017) and twins reared apart (Bouchard et al., 1990). **Conclusion:** The probability of violation of this assumption and the resulting bias are likely limited.
Non-additive genetic variance (dominance; *D*) – is assumed to be zero	
If *D*>0: overestimation of *A* and underestimation of *C*	▪ The CTD does not allow to simultaneously estimate *C* and *D*. Hence, an *ACE* decomposition implicitly assumes that *D* is zero. *D* results from interactions within a gene or between genes. Non-additive genetic effects are not expected for educational outcomes, which is influenced by many genes with small effects (i.e., polygenic). A high degree of polygenicity is generally accompanied by a lack of *D* (Hivert et al., 2021). Indeed, *D* is typically detected in highly biological traits, but seldom for social traits (Medland & Hatemi, 2009). ▪ An *ADE* model instead of an *ACE* model will usually be preferred if the DZ twin correlation is smaller than half of that of MZ twins, which is generally not the case for educational outcomes. **Conclusion:** This assumption is expected to hold, given that this dissertation focuses on educational outcomes and given that the comparison of twin correlations of the several studies in this dissertation does not indicate the presence of *D*.

Table 1.1 Assumptions of the Classical Twin Design (CTD).[a] (continued)

Possible bias if violated	Evidence
Assortative mating – is assumed to be absent	
If present: underestimation of *A* and overestimation of *C*	▪ There is assortative mating based on education. For Dutch married and cohabiting different sex couples, the correlation between male and female educational attainment has been found to be .47 (Verbakel & Kalmijn, 2014). This leads to increased environmental and genetic correlations between relatives, and thereby a genetic correlation among DZ twins larger than the assumed 0.5. **Conclusion:** This assumption is violated but does not substantially influence the conclusions of this dissertation. Violation results in a larger genetic correlation among DZ twins than assumed. This does not affect the variance component of interest in Chapter 2 (i.e., the classroom component). Taking the increased correlation into account (by using different values for the assumed genetic relatedness in Chapters 3, 4, and 5) does not substantially change the conclusions.
Gene-environment correlations (rGE) – are assumed to be absent for the outcome	
Passive rGE	
If present and positive: underestimation of *A* and overestimation of *C*. (a negative correlation is in theory also possible, but less likely) If present and a measured shared-environmental covariate (e.g., parental SES) is included in the CTD: effect of the measure is partly genetically confounded	▪ A positive passive rGE is expected. High-educated parents transmit genes that shape their children's educational performance and are more likely to expose their children to a home-environment conducive to higher performance. If the association between a family environment measure (e.g., provision of toys/books) and children's performance is greater in non-adoptive families than adoptive families, passive rGE is present (Rijsdijk & Sham, 2002). A Dutch study using parental and children's educational attainment indicated that this is indeed the case (Scheeren et al., 2017). ▪ A simulation study by Verhulst & Hatemi (2013) gives an idea on the size of the bias. If the correlation between genes and the shared environment has a moderate effect size (.3), *A* and *C* are of equal size, and *E* is small, the bias in *A* is about 10-15% of the total variance. Although this is a non-negligible amount, it would lead to an unbiased *A* estimate that usually still falls within the confidence interval of the biased *A* within the CTD. ▪ The size of the bias drops quickly when *A* and *C* are not of equal size but more different, and when *A* or *C* approaches zero, among others (Verhulst & Hatemi, 2013). This is usually the case for educational performance in the contexts that I study. *A* is typically large and *C* small or sometime even absent (De Zeeuw et al., 2016). Hence, the underestimation of *A* and overestimation of *C* can be expected to be small. **Conclusion:** This assumption is violated as a positive passive rGE can be expected, which decreases *A* and increases *C* in the univariate CTD. However, the bias is expected to be small for educational performance. Moreover, possible bias is likely not very problematic in the context of this dissertation, given that I am mostly interested in the *ACE* decomposition across different environments or over time. Violation of this assumption would only be problematic if the *A* and *C* are more biased in one context than the other (Knigge et al., 2022). This would require parental genes to influence children's shared environment differently for children in, e.g., high- vs. low-quality schools. Nevertheless, while interpreting the results it should be kept in mind that the *C* component may also capture passive rGE and that the effects of measured shared environments (e.g., parental SES) are genetically confounded.

Table 1.1 Assumptions of the Classical Twin Design (CTD).[a] (continued)

Possible bias if violated	Evidence
Evocative and active rGE	
If present and positive: overestimation of *A* and underestimation of *E* If present and negative: underestimation of *A* and overestimation of *E*	▪ In usual development beyond early childhood, especially positive correlations are expected (Scarr & McCartney, 1983). Yet, negative correlations are not unusual, especially for personality outcomes (Plomin et al., 1977). ▪ A twin study using longitudinal data on parental stimulation (specific to each twin) and cognitive ability of young children (2-4 years old) showed that genetic differences in children's initial ability levels predicted subsequent levels of cognitive stimulation by parents, that is, an evocative rGE (Tucker-Drob & Harden, 2012a). Active rGE is expected to play a role as well when children grow older, as children become more independent and more able to actively select environments for themselves (Tucker-Drob & Harden, 2012a). ▪ Another way to show the presence of evocative/active rGE, is by using longitudinal data and including the effects of polygenic scores (PGS)[b] in within-sibling analyses. A recent study on the development of educational performance (age 7-16) found evidence for evocative/active rGE based on their genetic propensity towards non-cognitive skills, but not for cognitive skills (Malanchini et al., 2023). ▪ The simulation study by Verhulst & Hatemi (2013) shows that presence of rGE affects the estimate of *A* more than that of *E*, and provides indications of the amount of bias under different situations. In general, the amount of bias is difficult to interpret, and dependent on the correlation between the specific environment and outcome under study. They also emphasize that calling it 'bias' is misleading, given that genes are still influencing the outcome but that the environment is mediating this influence. Hence, the authors state that heritability in the presence of rGE should be interpreted as the impact of genetic factors on the *development* of the outcome. **Conclusion**: This assumption is likely violated, which may be overestimate *A* (and to a smaller degree in *E*) in case of positive correlations. In case of negative rGE, which may be less often present, the opposite would be the case. Hence, while interpreting the results it should be kept in mind that the *A* component may also reflect evocative and active rGE.
Gene-environment interactions (GxE) – are assumed to be absent for the outcome	
If AxC is present: estimated in *A*. *A* reflects average genetic variance across all values of the measured *C* If AxE is present: estimated in *E*. *A* reflects average genetic variance across all values of the measured *E*	▪ If present, the genetic and environmental components are not independent and the CTD is then not the appropriate model. In general, the presence of GxE should be investigated with explicitly measured environmental indicators, preferably with theoretically driven hypotheses, on a trait-by-trait basis (Rijsdijk & Sham, 2002; Verhulst & Hatemi, 2013). **Conclusion:** This assumption is expected to be violated. Consequently, the estimates provided by the CTD provide a simplified representation. Investigating GxE is the focus of this dissertation, and I model several of them (including the shared-environmental moderators family SES, school SES, school quality). However, there might be other shared environmental moderators, as well as non-shared environmental moderators, that are not modeled in this dissertation.

Table 1.1 Assumptions of the Classical Twin Design (CTD).[a] (continued)

Possible bias if violated	Evidence
Sibling effects – are assumed to be absent, meaning that twins' educational performances have no influence on each other	
If present and positive: estimated in *C* If present and negative: estimated in *D*	▪ If the educational performance of twin 1 affects the educational performance of twin 2, and vice versa, there is a sibling effect. A positive effect reflects sibling cooperation or imitation effects, whereas a negative effect reflects competition or contrast effects. The CTD can be extended by adding paths for the direct effects of the performance of twin 1 on twin 2, and twin 2 on twin 1. For educational attainment, it has been shown that including these effects did not improve model fit (Nielsen & Roos, 2015). ▪ Cooperation and competition effects among twins affect the MZ and DZ covariance patterns that are consistent with patterns reflecting shared environmental effects (in case of cooperation) and non-additive genetic effects (in case of competition). Additionally, sibling effects also affect the total variance in educational performance. Neale and Maes (2004) mathematically show that cooperation increases the variance for MZ and DZ twins, and competition lowers the variance for MZ and DZ twins, and these effects are greatest on the MZ variance given their higher degree of relatedness. Therefore, the presence of sibling effects can be indicated by differences in MZ and DZ variances. It is standard practice in twin studies to test if the variance in an outcome can be equalized for both types of twins. **Conclusion:** This assumption is expected to hold. In all chapters, the variance does not differ by zygosity in a statistically significant way, providing evidence for the absence of sibling effects.
Generalizability – twins are not different from the general population in terms of the outcome that is studied	
If violated: results can only be generalized to twins	▪ For the Dutch population, twins were found to have slightly lower cognitive ability between age 6 and 8 than singletons, which could be related to lower birth weight and shorter gestation (Webbink et al., 2008). However, these differences had disappeared when children were 12 years old. Another study with Dutch data showed that twins had lower educational performance in primary school than their older non-twin sibling, but the same or even higher performance levels than their younger non-twin sibling (De Zeeuw et al., 2012). The difference in educational performance between twins and non-twins could largely be explained by birth other. ▪ Also Schwabe et al. (2017) show that results of twin studies can be generalized. They compared the variance components of educational performance of the Netherlands Twin Register with those derived from pedigree data (including the performance of full siblings, half siblings, full cousins, and half cousins). ▪ When the average performance and total variance in performance for the full population is known, differences between the twin sample and general population can easily be tested. **Conclusion:** This assumption is expected to hold, and results are generalizable beyond twins. Average performance levels and standard deviations are not substantially different between the twin sample and full population in the studies I conducted.

Note: [a] See for a general overview of the assumptions, e.g., Evans et al. (2002), Rijsdijk and Sham (2002), Verhulst and Hatemi (2013) [b] PGS is a composite measure based on the correlation between many genetic variants and an outcome, reflecting individual's genetic predisposition for this outcome.

1.3.2 Extensions of the CTD

The CTD can be extended in several ways. Two modeling extensions are most important for this dissertation. The first extension is the inclusion of a moderator. In this way, gene-environment interactions can be investigated.[5] For example, school quality can be included as a continuous moderating variable, which affects the estimates of the average educational performance and the underlying genetic and environmental components (Purcell, 2002). The genetic path estimate *a* becomes $a + b_a M$, for instance (see Figure 1.3). It is also possible to investigate moderation by modelling the CTD for different subgroups (e.g., boys and girls). This is also known as non-parametric gene-environment interaction analysis.

In the behavioral genetics literature, the *ACE* moderation model is referred to as a univariate model even though it includes multiple variables (in this example, educational performance and school quality). This is because only educational performance differs between twins and can be decomposed into the *ACE* components. The moderator is measured on the family level and always shared between twins from the same pair, hence, it only reflects shared environmental variance (Turkheimer et al., 2005).

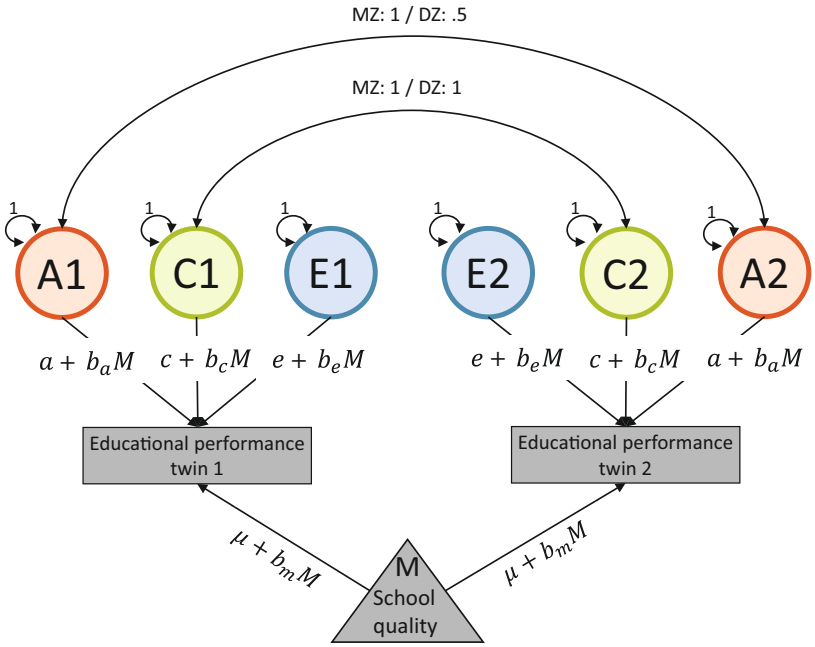


Figure 1.3 ACE moderation model.

5 Note that gene-environment interaction is a term from the behavioral genetics literature which refers to the whole class of moderation effects of the *ACE* model (Purcell, 2002). That is, it includes the interactions between a measured environment (M) and additive genetic effects, shared environmental effects, and non-shared environmental effects.

The second extension is the bivariate *ACE* model, which allows for decomposing the variance of two variables, as well as their covariance, into *ACE* components (see Figure 1.4). The so-called Cholesky decomposition reflects this. The genetic and environmental factors that influence variable 1 can be partly the same as those influencing variable 2 (captured by the cross-paths a_{21}, c_{21}, and e_{21}) but there could also be genetic and environmental factors that uniquely affect variable 2 (see paths a_{22}, c_{22}, and e_{22}). While this kind of bivariate modelling is often applied to two measured variables, I apply it to two latent variables. These are the estimated initial performance and learning growth, which will be described in more detail in section 1.5. This approach allows to investigate the genetic and environmental influences on performance over time.

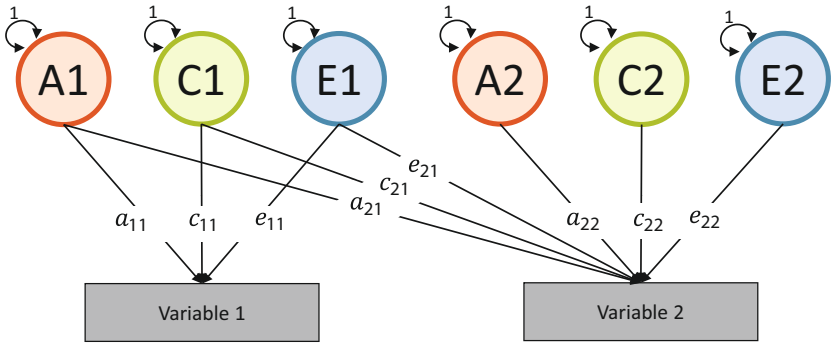


Figure 1.4 Bivariate ACE model with Cholesky decomposition.

Note: Only one twin is showed to avoid clutter.

1.3.3 Applications of the CTD to twin data and administrative data

In this dissertation, I apply (extensions of) the CTD to two data sources: twin registry data and administrative data. Twin registry data come with the advantage that zygosity of the twins is known but has the disadvantage that it depends on locating and recruiting twins and keeping them to participate (Figlio et al., 2017). This could result in a smaller and potentially selective sample of twins. Parents of twins that participate in twin registries have on average a higher SES, which leads to less variance in educational performance and potentially to an over- or underestimation of the relative contributions of genes and the environment (Schwabe et al., 2017). Administrative data have the disadvantage that important variables are not always available. In our case, it is a disadvantage that the zygosity of twins is unknown. An advantage of administrative data is that they do not suffer from self-selection bias. Moreover, administrative data provide many observations, which is advantageous given that this increases the power to detect genetic and environmental variation and their dependency on the family and school environment (Posthuma & Boomsma, 2000).

The zygosity of twins is often determined with questionnaire items and/or based on genetic markers such as DNA or blood group polymorphism. Zygosity questionnaires are very accurate. For example, in the Netherlands Twin Registry (NTR) the questionnaire allows for zygosity determination with 97% accuracy (Ligthart et al., 2019). When zygosity is known, it is possible to divide the twin data into a group of MZ twins with a genetic relatedness of 1, and a group of DZ twins with an average genetic relatedness of .50. When zygosity is unknown, these groups cannot be distinguished. Instead, two groups of twin types with different genetic relatedness can be created based on the sex composition of twins. This still allows to decompose the variance in educational performance into the genetic and environmental components, just as the CTD when zygosity is known. Opposite-sex (OS) twins are always DZ and thus have an average genetic relatedness of .50. Same-sex (SS) twins are a mixture of DZ and MZ twins, and hence their average genetic relatedness lies in between the genetic relatedness of DZ and MZ twins (i.e., in between .50 and 1). In any case, SS twins are on average genetically more similar than OS twins. The same logic as in the CTD applies: if the educational performance of SS twins is more similar than that of OS twins, this is indicative of educational performance being genetically influenced. However, to identify the model, it is needed to impose an assumed genetic relatedness of SS twins. There are different ways in the literature to do this and it requires additional assumptions. More details on this are provided in Appendix A. Importantly, these are testable assumptions and one can check the extent to which the results are dependent on the assumed genetic relatedness of SS twins. Prior studies showed that assumptions that are needed to fit the model hold (De Zeeuw & Boomsma, 2017; Figlio et al., 2017). Also, my own robustness analyses show that the conclusions are not dependent on the assumed genetic relatedness of SS twins.

1.4 The twin design and educational inequality

The absolute and relative variance components as derived from the twin design can be used to study educational inequality. I focus on the two types of educational inequality that are distinguished by Van de Werfhorst and Mijs (2010): inequality as dispersion and inequality of opportunity (see also Section 1.1). Inequality as dispersion is studied by investigating the distribution, for example, by focusing on the total variance, standard deviation, or different percentiles of the performance distribution (Hanushek & Wößmann, 2006; Van de Werfhorst & Mijs, 2010). Inequality of opportunity in education is often narrowed to differences in average performance between socioeconomic groups (i.e., the influence of parental SES on children's performance). By doing this, the considerable variation within groups is overlooked (Ziegler et al., 2021). Focusing on the *ACE* components of the twin design provide another way to investigate both inequality as dispersion and inequality of opportunity.

The relative contribution of genetic variance (i.e., heritability) is often seen as an index of (in) equality of opportunity (Asbury & Plomin, 2014; Nielsen, 2016; Nielsen & Roos, 2015; Pokropek & Sikora, 2015). In a meritocratic society where children are not constrained by their family background and all children have the opportunities to develop their full genetic potential, genetic differences between children will become more visible resulting in a high heritability (Asbury &

Plomin, 2014). Comparing the heritability of educational performance across contexts, such as schools of different quality, can thus be used to investigate in which context there is more equality of opportunity. If high-quality schools would provide more equal opportunities, the relative contribution of *A* is expected to be larger and the relative contribution of *C* is expected to be smaller. The absolute *A*, *C*, and *E* components can be used to investigate inequality as dispersion. The total variance is a measure of inequality as dispersion, and the *ACE* components show the underlying genetic, shared environmental, and non-shared environmental sources of dispersion.

It is important to study both the relative and absolute variance components together. First, if one would rely on the relative components only, no conclusions on inequality as dispersion can be made. Dispersion is disregarded since relative components add up to 100% by definition. Second, relative components conceal that *A* and *C* could change independently. Given one's view on whether shared environmental differences, genetic differences, or both are problematic (see section 1.1), it is important to know what drives the pattern. For example, a larger relative contribution of *A* and smaller relative contribution of *C* in higher-quality schools could be driven by an absolute increase of *A*, absolute decrease of *C*, or both. If there is both a relative and absolute decrease of *C*, this reflects the common idea of more equality of opportunity (i.e., less family background differences). Lastly, relying on only the absolute components would also not be ideal because the total variance may differ between contexts while the effects of genes and the environment do not differ. For example, in some schools children may be more similar to each other concerning their genetic makeup and/or (non-)shared environmental characteristics (Knigge et al., 2022). Using the relative, standardized, *ACE* components takes this into account.

To conclude, both relative and absolute variance components can be used to study educational inequality. Inequality as dispersion, and the underlying sources of the dispersion, can be investigated by means of the absolute variance components. Inequality of opportunity can be examined by both the absolute and relative variance components. Since most people agree that there is more equality of opportunity if children's performance is less determined by their family background, a smaller relative and absolute *C* component capturing environmental between-family influences can be used as an indicator of (in)equality of opportunity. For the *A* component, it is more debatable. If the absolute *A* component is larger, this indicates more genetic inequality, which could be seen as problematic. Also, the *A* component does not only capture genetic influences of characteristics that are positively valued in school, such as genetic potential. It also captures genetic differences of characteristics that have a negative effect on performance, for example, behavioral and health problems (Krapohl et al., 2014). Therefore, I only use a higher relative *A* and lower relative *C* as an indicator of equality of opportunity if this is driven by a lower absolute *C*.

Table 1.2 Overview of the empirical chapters.

Chapter and focal interaction		Data	Cohort	N_{pairs}	Outcome	Moderator	Analytical strategy
2	To what extent classroom influences and family SES multiply or compensate for each other	NTR	1991-2002	1,476 MZ twins 2,740 DZ twins	Cito score end of primary education (~age 12)	Parental education (ISLED)	Univariate *ACE* moderation model extended with a classroom component
3	To what extent school quality multiplies or compensates for genetic and shared environmental influences	Register data (Statistics Netherlands) + Inspectorate of Education	1994-2007	18,384 SS twins 11,050 OS twins	Cito score end of primary education (~age 12)	▪ School quality (factor score) ▪ School SES (factor score) ▪ Parental SES (factor score)	Univariate *ACE* moderation model
4	To what extent parental SES, school SES, gender, and their intersection, multiply or compensate for genetic and shared environmental influences	Register data (Statistics Denmark)	1986-1998	5,010 SS twins 32,283 SS siblings	GPA end of compulsory education (~age 16)	▪ Parental education (binary: college degree) ▪ School SES (binary: median cut average parental education) ▪ Gender (binary)	Univariate *ACE* model with non-parametric moderation
5	To what extent genetic and shared environmental influences, and the total dispersion in performance, are multiplied or compensated during primary education	NCO-LVS + register data (Statistics Netherlands) + Inspectorate of Education	2004-2014	3,544 SS twins 2,155 OS twins	Tests scores throughout primary education (~age 6 to 11)	n/a	Biometric latent growth model (i.e., bivariate *ACE* model for the latent intercept and slope)

Note: NTR = Netherlands Twin Register, NCO-LVS = National Cohort Study on Education – learning monitoring system, MZ = monozygotic, DZ = dizygotic, SS = same-sex, OS = opposite-sex, GPA = grade point average, ISLED = international standard level of education, *A* = genetic variance, *C* = shared environmental variance, *E* = non-shared environmental variance.

1.5 Four empirical studies

1.5.1 Overview of the studies

In order to provide more insight into the interplay between genes, families, and schools in explaining educational performance and inequality, I conducted four empirical studies. Table 1.2 provides an overview of the four empirical chapters of this dissertation, summarizing the interactions that are central in the chapter, data, core measurements, and analytical strategy. Table 1.3 presents a summary of the results.

Chapter 2: Are classrooms equalizers or amplifiers of inequality?

Chapter 2 examines if there is a compensatory or multiplicative interplay between families and classrooms in explaining educational performance. The classroom environment (e.g., teacher quality, pedagogical climate, peer influence) could be more important for the performance of lower-SES children. In this case, favorable classroom contexts would equalize inequality with respect to socioeconomic background (i.e., compensation). Conversely, if the classroom environment is more important for higher-SES children, the classroom context potentially increases educational inequalities (i.e., multiplication). I investigate the two alternatives by analyzing the scores on a standardized achievement test (i.e., *Cito* test) of 4,216 twin pairs from the Netherlands Twin Register. Given that some twins are in the same classroom while others are in separate classes, I extend the CTD and estimate a latent classroom component next to the *ACE* components. To examine whether classrooms and parental SES multiply or compensate for each other, I include parental education as a moderator.

On average, only a small part (2%) of performance differences can be accounted for by the classroom environment. The classroom accounts for almost 8% of the variance in educational performance for children whose parents finished at most primary education. This decreases to 1% for children with postdoctoral educated parents. This suggests that family and classroom influences compensate for each other. Still, the role of classrooms is relatively small, and it can be expected that especially children from high-SES families are in high-quality classes given the socio-economic selection into schools (Borghans et al., 2015a). Therefore, classroom environments have a modest potential to reduce SES differences in performance but are, based on the results of this study, not 'great equalizers'.

The twin design provides some additional conclusions on inequality of opportunity and inequality as dispersion. The relative variance components show that heritability is relatively large (around 75%) for all children regardless of their social origin. This can be interpreted as realized genetic endowment, reflecting the relative opportunity for achievement, being independent of children's socioeconomic background. However, while the relative contribution of genes is the same across all levels of parental education, the absolute genetic variance is not. The higher the level of parental education, the less absolute genetic variance, and, to smaller degrees, less non-shared environmental and less classroom variance. Altogether, this contributes to less inequality as dispersion with higher levels of parental education. This is consistent with a compensatory interaction between parental education and genetic, non-shared environmental, and classroom influences.

Chapter 3: Does school quality decrease educational inequality?

In Chapter 3, I investigate to what extent genetic and environmental influences on children's educational performance depend on the quality of children's school environment. The behavioral genetics literature on gene-environment interactions and sociological literature on the relationship between schools and inequality, provide arguments for both a larger and smaller role of genes and the shared environment in higher-quality school environments. I investigate whether there are multiplicative or compensatory interactions between school quality and genetic and shared environmental influences. Given the selection of high-SES children in higher-quality schools, I expect a school quality interaction effect to be at least partly explained by SES. To examine if the school environment and/or the family environment plays a role, I incorporate multiple moderators (i.e., factor scores for school quality, school SES, and parental SES) in the CTD simultaneously (Purcell, 2002). I use Dutch administrative data on 18,384 same-sex and 11,050 opposite-sex twin pairs, enriched with many indicators on the quality of primary schools as obtained from the Dutch Inspectorate of Education.

The results show that children in higher-quality schools, as measured by many indicators related to school resources and school climate, have somewhat higher educational performance (i.e., *Cito* test score) but that the effect size is small. The variance decomposition indicates that inequality of opportunity does not depend on school quality. Heritability is not moderated by school quality or school SES, neither are the relative and absolute shared environmental variance components. At first sight it seems as if there is less absolute genetic variance (and less non-shared environmental variance) in higher-quality schools. This suggests less inequality as dispersion in higher-quality schools. However, these results appear to be related to school SES and parental SES rather than school quality. Hence, it is not the quality of schools that compensate for genetic influences, but high-SES families and schools. Altogether, these results indicate that school quality does not decrease (and neither increase) inequality of opportunity nor inequality as dispersion. Reducing quality differences between schools would likely not be sufficient to reduce educational inequality.

Chapter 4: The intersection between gender, family background, and school context

Chapter 4 focuses on the gene-environment interplay in educational performance while considering further complexities. Family and school environments could moderate genetic and environmental influences and it is important to investigate families and schools simultaneously given the socioeconomic selection into school (see also Chapter 3). The role of families and schools are not necessarily independent and could interact. Moreover, a gene-environment interaction in educational performance could work differently for boys than for girls. A stronger interaction for boys can be expected, because investments in high-SES families and schools may be higher for boys and/or because boys may be more sensitive to their environment (e.g., Autor et al., 2019; Legewie & DiPrete, 2012). Therefore, I investigate whether gene-SES interactions in educational performance follow a multiplicative or compensatory pattern with respect to family SES, school SES, and their intersection, as well as whether such interaction patterns are stronger for boys. I use data from Statistics Denmark on the educational performance (i.e., GPA at the end of compulsory school) of 5,010 same-sex twin pairs and 32,283 same-sex sibling pairs. SES

is measured by whether at least one parent has obtained a college degree. For school SES, the share of children with college-educated parents is calculated by school and graduation cohort, and divided into two equally sized groups (cut at the median).

I find a gene-environment interaction for family SES, but not for school SES. In high-SES families, heritability is lower. This gene-family SES interaction is mainly driven by less absolute genetic variance in high-SES families. Second, the gene-family SES interaction is moderated by the children's gender. In high-SES families, heritability is considerably lower for boys than for girls. The absolute variances show that this is related to both lower genetic variance for boys and higher environmental variance. Third, the moderating effect of family SES for boys is almost entirely driven by children attending low-SES schools.

These findings can be interpreted as high-SES families compensating for lower genetic potential or genetic risk for boys, especially when they attend low-SES schools. The finding that the gene-family SES environment interaction is more pronounced in low-SES schools corresponds with the idea that family and school influences substitute or compensate for each other's influence. In general, school SES does not seem to play a large role in educational inequality. I do not find evidence that inequality of opportunity depends on school SES. Inequality as dispersion is a bit lower in high-SES schools, driven by smaller absolute genetic variance in these schools. But when the multiplicity of social contexts is considered, school SES does turn out to play a role. The gene-family SES interaction appears to be heterogenous, as it is largely concentrated in low-SES schools. Interestingly, this is the case for boys, but not for girls. A potential explanation could be that boys are more sensitive to their environment (Legewie & DiPrete, 2012).

Chapter 5: The development of inequality during primary education

Chapter 5 investigates how dispersion in educational performance develops over the primary school career and why. I study if initial genetic and environmental differences in performance at the start of formal education are reproduced, exacerbated, or compensated over the primary school career. I also examine if new genetic and environmental sources of dispersion are coming into play during schooling. Lastly, I test to what extent dispersion can be explained by school quality, school SES, and family SES. I apply biometric latent growth models to reading and mathematics tests scores of around 5,500 same-sex and opposite-sex twin pairs, which I identified in data of the Netherlands Cohort Study on Education (Haelermans et al., 2020). These models decompose the variance in initial performance (i.e., the intercept) and learning growth (i.e., the slope) in genetic and environmental components, as well as the covariance between initial performance and growth.

Table 1.3 Overview of the results with respect to inequality of opportunity and inequality as dispersion.

Chapter	Moderator	Relative components				Absolute components				
		A	C	E	CL[b]	V	A	C	E	CL[b]
2	**Parental SES**	-	-	-	↓CL	↓ V	↓ A	[c]	↓ E	↓CL
3	**School quality**	-	-	-	-	↓ V	↓ A	[c]	-	-
	controlled for parental SES and school SES	-	-	-	-	-	-	-	-	-
	School SES	-	-	-	-	↓ V	↓ A	[c]	↓ E	-
	controlled for parental SES and school quality	-	-	-	-	↓ V	↓ A	↓ C[d]	-	-
	Parental SES	-	-	-	-	↓ V	↓ A	[c]	↓ E	-
	controlled for school SES and school quality	-	-	-	-	↓ V	↓ A	↑ C[d]	↓ E	-
4	**School SES**	-	-	-	-	↓ V	↓ A	-	-	-
	boys	-	-	-	-	↓ V	-	-	-	-
	girls	-	-	-	-	↓ V	↓ A	-	-	-
	Parental SES	↓ A	-	↑ E	-	↓ V	↓ A	-	↑ E	-
	boys	↓ A	↑ C	↑ E	-	↓ V	↓ A	-	↑ E	-
	girls	-	-	-	-	↓ V	↓ A	-	-	-
	low-SES school, boys	↓ A	↑ C	↑ E	-	-	↓ A	↑ C	↑ E	-
	low-SES school, girls	-	-	-	-	↓ V	-	-	-	-
	high-SES school, boys	-	-	-	-	↓ V	↓ A	-	↑ E	-
	high-SES school, girls	-	-	-	-	↓ V	↓ A	-	-	-
5	**Time**[a] (total variance reading)	-	-	-	-	↑ V	-	-	-	-
	initial variance	-	-	-	-	↓ V	↓ A	-	-	-
	new variance	-	-	-	-	↑ V	↑ A	-	↑ E	-
	Time[a] (total variance math)	-	-	-	-	↑ V	↑ A	-	-	-
	initial variance	↑ A	↓ C	↓ E	-	↓ V	-	↓ C	↓ E	-
	new variance	-	-	-	-	↑ V[e]	↑ A[e]	-	↑ E[e]	-

Note: V = total variance, A = genetic variance, C = shared environmental variance, E = non-shared environmental variance, CL = classroom variance, ↓/↑ = decreasing/increasing with increasing levels on the moderator. Empty cells indicate no statistically significant change.

[a] Time is strictly speaking not included as a moderator. Instead, the development of the *ACE* components is followed over time as estimated via latent growth modelling. Moderation in this case refers to whether there is a statistically significant change over time. For the relative components capturing new variance, a change over time is not applicable

[b] Only included in Chapter 2.

[c] After including the moderator, shared environmental variance is very minimal or absent.

[d] Although statistically significant, the change is substantively very small

[e] While the underlying path estimate is statistically significant, the estimated variance component does not reach statistical significance at the five percent level.

The results show, first of all, that existing dispersion at the start of formal education is compensated over the primary school career. For both reading and mathematics performance, genetic differences account for a large part of the dispersion in initial performance. For reading, these genetic differences decrease while environmental differences are largely reproduced over the school career. For mathematics, the decrease in dispersion results from decreasing environmental differences, while genetic differences are reproduced. Second, new sources of dispersion are coming into play over time, mostly new genetic influences. Combining the development of pre-existing sources of dispersion with these new influences results in an increase in the total dispersion in educational performance over time. Third, measured school characteristics did not account for (the development of) dispersion, suggesting that quality differences between schools likely play a limited role.

Concerning inequality of opportunity, the results suggest that over the school career inequality of opportunity in reading performance is reproduced. Both the absolute and relative shared environmental variance do not change during primary school. Inequality of opportunity in mathematics, on the other hand, does decrease during primary education. Although these changes occur over the primary school career, more research is needed to investigate whether it is actually produced by the school environment or not.

1.6 Conclusion and discussion

1.6.1 Lessons learned

In this dissertation, I conducted four empirical studies to answer the question: *How does the interplay between schools, families, and genes shape educational inequality?* Four overarching conclusions can be drawn from this.

1.6.2 Compensation rather than multiplication

In general, schools, families, and genes turned out to play a role in a compensatory way rather than in a multiplicative way. The compensation interactions between the school and family environment (Chapters 2, 3, and 4) and schools and genes (Chapters 3, 4, and 5) show that schools have the potential to reduce shared environmental and genetic differences in educational performance. The compensatory interaction between family SES and genes (Chapters 2, 3, and 4) indicates that dispersion in performance related to genetic variance is smaller in high-SES families. Compensation is stronger in some domains than others. Generally, the gene-environment compensation interactions involving families are stronger than those involving schools.

The compensation interaction between genes and SES at the family and school level supports the diathesis-stress model (Rende & Plomin, 1992; Shanahan & Hofer, 2005) and the compensatory advantage mechanism (Bernardi, 2014). In environments with more risks and stressors, such as low-SES families and schools, the realization of genetic risks toward lower performance is more likely. In high-SES environments, where stressors are more often absent and more positive features are present, the influence of genetic risks (including lower genetic potential) is neutralized or compensated.

Families and schools not only compensate for the influence of genes on educational performance but also for each other. This is in line with the argument that children from low-SES families could substitute school resources for family resources. Because these children are less likely to find supportive influences at home, for example, they may be more susceptible to a supportive environment in school (Coleman et al., 1966; Jencks & Mayer, 1990; Rumberger & Palardy, 2005). Conversely, a high-SES family environment may compensate for a less advantageous school environment (e.g., by providing tutoring) (Hanselman, 2018).

Another compensation pattern is found when investigating how the initial sources of dispersion evolve during the primary school period (Chapter 5). Initial dispersion in reading and mathematics performance decreases. Moreover, inequality of opportunity in mathematics decreases. This suggests that schools might contribute to reducing inequality, although more research is needed to know if these changes are indeed induced by the school environment.

Prior empirical studies provided mixed results for both gene-environment interactions and family-school interactions in education. One explanation that has been offered for the inconsistent findings for gene-environment interactions is country differences in welfare state provisions, among others (Tucker-Drob & Bates, 2016). I find support for compensatory gene-environment interactions in educational performance in the Netherlands and Denmark, countries where the welfare state arrangements ensure that the living standard of low-SES families is not as low as in other countries (e.g., the U.S.). Baier et al. (2022) provide arguments for both a stronger and weaker genetic influence on educational outcomes when welfare states have more protective features. Also, differences in welfare state provisions are thought to affect the *strength* of gene-SES interactions (Baier et al., 2022; Tucker-Drob & Bates, 2016). However, it is unclear how they may affect the *direction* of gene-SES interactions, while this is most important for explaining the mixed results. Next to country differences in welfare state arrangements, other reasons for inconsistent findings that have been proposed are children's age, operationalization of family background, and the 'selectivity' of the educational outcome[6] (Ghirardi & Bernardi, 2023; Mönkediek et al., 2023). More theory development and comparative research is needed to provide more insight into under which conditions compensation is more likely than multiplication. Since this is currently underdeveloped, empirical support for the compensation interactions should be interpreted in the study contexts covered in this dissertation and should not be overgeneralized.

1.6.3 School quality plays no role, but school SES and classrooms do

The school environment plays a role in shaping educational inequality, but the quality of schools is not as important as often argued (Inspectie van het Onderwijs, 2017; Jennings et al., 2015). The degree of inequality of opportunity and inequality as dispersion did not depend on the quality of schools (Chapter 3). Instead, the socioeconomic composition of the school and the classroom environment turn out to be of (modest) importance. School SES differences and classroom differences were related to educational inequality in a compensatory way (Chapters 2 and 3).

I expected that the school environment affects educational performance and inequality therein via school resources, school climate, composition of the student body, and teacher quality.

6 With selectivity it is meant how difficult and costly it is to obtain a certain educational outcome.

School quality, school SES, and the classroom environment show weak associations with average educational performance (Chapters 2-5). Children who attend higher-quality schools, as measured by indicators related to school resources and climate, perform on average slightly better. However, school quality does not multiply or compensate for genetic and environmental influences on performance (Chapter 3). Thus, school quality is not related to inequality of opportunity or inequality as dispersion. This does not mean that differences between schools are not important for educational inequality at all. The SES composition of schools is related to less inequality of dispersion, especially less genetic dispersion. The literature suggests that a larger share of high-SES pupils in school contributes to an environment that is more conducive to learning, for example, because there are fewer classroom disruptions and a more learning-oriented climate (Kahlenberg, 2001; Rumberger & Palardy, 2005). Moreover, high-SES pupils may have a positive influence on the aspirations, study habits, and performance of their peers (Gutiérrez, 2023). This may be beneficial for all pupils, but especially for pupils with lower performance levels (e.g., related to the presence of genetic risks).

Not only differences between schools but also within schools play a role. While resources (e.g., educational materials) and climate (e.g., academically oriented, high expectations) were not important in explaining educational inequality at the school level, they may still play a role at the classroom level. I investigated the overall (latent) classroom environment instead of specific aspects of the classroom, but it can be expected that classroom differences reflect not only differences in teacher quality but also differences in classroom climate (Byrne et al., 2010). I do not expect that classroom differences reflect differences in resources, given that it is more likely that resources vary between schools rather than within schools.

The finding that school quality is only weakly related to average performance levels and not related to educational inequality, may be surprising. In the last years, the importance of school quality has been emphasized, among others in the debate on decreasing educational performance and increasing inequality (Inspectie van het Onderwijs, 2017, 2023a). However, the findings of this dissertation are not necessarily incompatible with this. First of all, it could be that the quality of education decreases in all schools. In that case, the (limited) variation in school quality would remain the same, while it still affects the average performance levels of pupils. Second, it could be that educational quality differences play a different role than thought. Inspectorates notice a wide variety in quality when they visit schools (Inspectie van het Onderwijs, 2017). During such visits, inspectors inform their overall assessment of the school quality in practice based on observations in some school classes and conversations with teachers (Inspectie van het Onderwijs, n.d.).[7] Given that not all classes and teachers are observed, within-school variation may be overlooked while the results of this dissertation suggest that it is not so much the quality between schools but within schools that matter.

7 In addition to school visits, school quality is also assessed by conversations with school leaders among others, and by analyzing data about the school and schoolboard (e.g., the school plan, social safety monitor).

1.6.4 Family SES plays a role but differently than often thought

It is commonly argued that a substantial portion of the disparities in children's educational performance stems from the different rearing environments associated with family SES (Sirin, 2005). The influence of family SES on educational performance is often thought to be uniform where family SES affects children in the same family in the same way (cf. Freese, 2008). This dissertation shows that this is only a small part of the story. First, I find that shared environmental variance in educational performance is relatively small, implying that family SES is less of an exogenous environment than thought. Second, the results show that genetic influences were smaller in higher-SES families (Chapters 2-4), which can be interpreted as high-SES parents compensating for children's genetic risks for lower performance. Non-shared environmental influences on children's performance were also smaller in high-SES families in the Netherlands (Chapters 2 and 3), but larger in Denmark (Chapter 4). This suggests that part of the influence of family SES originates from different responses to child dissimilarities, which dissimilarities are genetically and environmentally caused. Lastly, there is a lot of variation in educational performance among children with a lower-SES background. This within-group inequality is often overlooked if the focus is only on differences in performance between SES groups.

Family SES often accounts for all the shared environmental variance in educational performance. However, the shared environment accounts for only a small part of the individual differences in educational performance. That is not to say that family SES has little influence on children's performance and that classical sociological mechanisms related to parental resource transmission, investments, and expectations do not apply. Instead, it suggests that family influence is less likely to be exogenous and more likely to operate in ways that correlate with children's genetic endowment (Freese, 2008). This includes, for example, evocative gene-environment correlations where children's educational potential is recognized by parents, who provide an enriched environment to maximize this potential (Diewald et al., 2015; Plomin et al., 1977). Especially high-SES parents can be expected to recognize children's talents and/or provide more enriched environments. Compared to exogenous family SES characteristics, family SES experiences that operate via gene-environment correlations tend to be more persistent and recurring. Therefore, they are argued to have more systematic and lasting effects on development (Tucker-Drob & Harden, 2012a). The gene-family SES interplay thus forms an important way of how family SES shapes educational performance.

The results show less genetic variance in higher-SES families, which is consistent with the idea that the expression of genetic risks (or lower genetic potential) is compensated in higher-SES environments (Bernardi, 2014; Rende & Plomin, 1992; Shanahan & Hofer, 2005). In lower-SES families, on the other hand, there is more dispersion resulting from more genetic variance. This could be because in low-SES families there may be more environmental risk factors or stressors (e.g., family chaos, poverty, divorce) that could lead to heightened expression of genetic risks for lower performance (Asbury et al., 2005; Shanahan & Hofer, 2005). Genetic variance may also be larger in these families because of lower incentives to compensate and lower levels of resources available to invest in compensatory strategies. For example, low-SES parents may be less likely to invest in supplementary education (e.g., private tutoring or test preparation) that could otherwise mitigate the negative consequences of genetic risks (Bernardi & Grätz, 2015; Stienstra et al., 2021).

Family SES is also related to non-shared environmental variance, meaning that the extent to which environmental factors make children within the same family more different varies by SES. For Dutch pupils' educational performance, non-shared environmental variance is lower in higher-SES families (Chapters 2 and 3). However, in Denmark it is higher, but only among boys. The interaction between non-shared environmental variance and SES could reflect differential parental treatment that is stratified by SES (e.g., Conley, 2008; Grätz & Torche, 2016). According to this literature, parents could have a compensatory response and invest most in the less-endowed child, among others to avoid unequal outcomes among their children. Alternatively, parents could also reinforce differences and invest more in the more-able child. These responses may vary with family SES, due to differences in parenting practices and access to economic, cultural, and social resources (Grätz & Torche, 2016). If stratified parental responses indeed account for the changes in non-shared environmental variance by SES, the findings suggest that high-SES parents have compensatory responses in the Netherlands and reinforcing responses for boys in Denmark.

These differences could result from variations in educational performance measurements. In the Netherlands, educational performance is measured by a test at age 12, while in Denmark it is based on averaged exam grades at age 16. Age has been suggested as a potential moderator of the *ACE*-SES interaction (Gottschling et al., 2019; Tucker-Drob & Bates, 2016). However, theoretically it is unclear how the *ACE*-SES interaction (that can be positive, negative, or absent) would vary with age (Mönkediek et al., 2023). Therefore, it is hard to derive expectations on whether there would be reinforcement or compensation with increasing age. Another explanation for the different findings could be that exam grades represent a more selective educational outcome than the *Cito* test score (cf. Ghirardi & Bernardi, 2023). For obtaining a high *Cito* score, it may be sufficient if one has either a high ability or ample support of high-SES parents. Obtaining high-exam scores (averaged over a variety of subjects) may be more difficult and maybe both a high ability and high SES are needed to get a high GPA. Nevertheless, different parental investment is only one of the potential explanations. Providing more insight into the underlying mechanisms of non-shared environmental variance, and its dependency on the family environment, is therefore an interesting direction for future research (see Section 1.7.4)

1.6.5 Important to study schools, families, and genes together

It is important to study the influences of schools, families, and genes together to prevent that influences are mistaken for one another and to gain a better understanding of how they play a role in educational inequality. First, this is because these influences are correlated and the influence of one factor may be picked up by another factor. For example, family SES and the school environment are correlated due to the socioeconomic selection into schools. Hence, their influences are (partly) mistaken by each other if only one of them is studied (Chapter 3). It is also shown in the interpretation of the components of the twin design, where genetic variance captures positive active and evocative gene-environment correlations if these are present but unmodelled (which is the case in this dissertation). Second, it is important to study them together because there is not one general, independent effect of schools, families, or genes. Instead, the influence of one depends on one of the other factors. This becomes apparent in the several observed gene-environment interactions (Chapters 2-4).

Prior studies may have overestimated and misunderstood the influences of schools, families, and genes because these influences were not studied simultaneously. For example, the classroom environment captures influences of pupils' socioeconomic background or genetic predisposition if those influences are not sufficiently considered. Even studies with measures and methods that are thought to exclude such influences, such as added-value measures, do not completely succeed in this (Haworth et al., 2011; Morris et al., 2018). Prior studies may therefore have overestimated the influence of classrooms (see Chapter 2).

Also, gene-environment interactions would be misunderstood if the family and school environment are not studied simultaneously (see Chapters 3 and 4). For example, the compensating role of school quality and school SES decreases substantially when family SES is taken into account (Chapter 3). Previous studies investigated gene-school environment interactions without explicitly taking the family environment into account (e.g., Hart et al., 2013; Haughbrook et al., 2017; Taylor et al., 2010, 2020). Given the socioeconomic selection into schools, the school influences may actually reflect family influences. It would therefore be interesting to replicate these prior twin studies while including the family SES. The prior twin studies that investigate gene-school environment interactions are conducted using U.S. samples, where socioeconomic selection into school likely plays a more important role than in the countries that I studied (i.e., The Netherlands and Denmark). Hence, especially in the U.S., it can be expected that the school effects actually reflect family-based mechanisms if family SES is not considered. However, studying multiple interactions simultaneously may be more challenging due to the smaller sample sizes employed in twin studies. The advantage of administrative data is that the educational performance of almost the whole population is known. This provides many observations and thereby increases the power to investigate the interaction between families, schools, and genes.

Lastly, considering that schools, families, and genes interact leads to a better understanding of how they play a role in explaining educational performance and inequality. For example, the classroom environment only plays a very small role in explaining performance differences in general. But by investigating how the classroom contribution depends on family SES, I show that the role of classrooms is of importance for a subgroup, namely low-SES children (see Chapter 3).

1.7 Limitations and future research

While this dissertation provides valuable knowledge on the interplay between schools, families, and genes in shaping educational inequality, there are still challenges and questions for future research.

1.7.1 Measurement of school quality

The results of this dissertation point towards a limited role of school quality differences. School quality was at most weakly associated with average educational performance and learning growth and did not moderate genetic and environmental influences once SES was considered. This could be partly related to limitations of the school quality measure. School quality is measured

by relying on indicators that are used by the Dutch Inspectorate of Education. The inspectorate primarily assesses whether schools are of (in)sufficient quality based on whether schools meet the legal standards (Inspectie van het Onderwijs, 2018). Schools that meet the legal standards and are thus of 'sufficient quality' still vary in their quality. I did not use the overall school quality judgment of the inspectorate but relied on the underlying items, which increased the variance in the school quality measure. Nevertheless, the measure likely does not reflect all the quality differences that exist between schools.

Another limitation of the school quality measure is that it does not sufficiently cover what can be regarded as the closest and most immediate influence on children's learning, that is, teachers and the learning environment they provide (Raudenbush, 2008). Classroom differences play a role, as shown in Chapter 2. Also, prior research shows the importance of the classroom environment, especially the role of teachers. For example, pupils who are taught by teachers who use more effective teaching practices perform better (Burgess et al., 2022). While the inspectorate data cover classroom environment indicators, they are measured at the school level and thus do not provide enough information on differences within schools.

To gain more insight into the role of schools, future studies could use different measures and methods. For example, by using more specific measures on the classroom level, such as detailed observations of instructional activities, and teachers' skills and effort in carrying out specific teaching tasks (Burgess et al., 2022). Alternatively, more insight into the role of schools, families, and genes can be obtained by comparing what happens if there is no schooling. During the school year learning is influenced by both school and non-school factors, while in the absence of schooling, such as during the summer break, learning is solely shaped by non-school factors (e.g., Downey & Condron, 2016). Comparing the learning gains when school is in versus out of session, as is done in seasonal comparison studies, shows to what extent inequality is shaped by the school or non-school environment. Combining twin data with such a seasonal comparison design, or with a design incorporating the school closure due to the COVID-19 pandemic, would provide a novel way to gain more insight into the extent to which schools affect inequality as dispersion and inequality of opportunity. Although the focus of such a design (i.e., what is the effect of schooling compared to a no-school counterfactual) is different from comparing schools of different quality as I do in this dissertation, it can still be informative. In the U.S., such studies suggest that schooling reduces social inequality in average performance and inequality as dispersion (e.g., Condron et al., 2021; Von Hippel et al., 2018). If this also holds true for countries like the Netherlands where school quality differences are smaller, then it may be more effective to focus on increasing school *quantity* (e.g., via summer schools or a longer school year) rather than on decreasing between-school differences in school *quality* (Downey et al., 2004).

1.7.2 Ceiling effect

An alternative explanation for decreasing inequality as dispersion with increasing parental SES and school SES (see Chapters 2, 3, and 4) are ceiling effects. The *Cito* score, used as the operationalization of educational performance in Chapters 2 and 3, is scaled in a particular way to make scores comparable over the years. The number of correctly answered questions on the test is transformed into a scale with a range from 501 to 550, a mean of 535, and a standard

deviation of around 10. Scores that fall below 501 or above 550 are rounded (Van Boxtel et al., 2010). More pupils are scoring at the upper end of the scale, and these are more often pupils from higher-SES backgrounds. High-SES pupils more often obtain the maximum score of 550 (De Zeeuw et al., 2019). This could lead to a spurious interaction between SES and (genetic and environmental) variance in educational performance (see Rohrer & Adams, 2021). The variance in performance among children from higher-SES families is then not lower because genetic and environmental influences are smaller (e.g., because they are compensated). Instead, it is lower because there is variability beyond the range of the scale which becomes censored by the scale.

Although a ceiling effect is possible, I expect that the decreasing variance with increasing SES reflects (at least partly) substantive compensation mechanisms. The study by De Zeeuw et al. (2019) investigated the interaction between family SES and the underlying sources of variance in *Cito* scores and corrected for the effect of censoring at the high end of the scale, which led to the same results. Moreover, in Chapter 4 educational performance is measured by GPA which is more normally distributed than *Cito* scores. Also, here I find decreasing variance with increasing SES. Hence, these results could be substantively interpreted. More research into the mechanisms underlying the decreasing (genetic) dispersion with increasing SES could show whether the substantive interpretation of compensation indeed underlies the observed pattern. Future research could include potential mediators of the gene-SES interaction in the twin design (see Ruks, 2022; Tucker-Drob & Harden, 2012b, 2012c). Relevant mediators for compensation patterns could include characteristics that have a negative effect on performance, such as ADHD and dyslexia.

1.7.3 Capturing complexities

Another consideration is that the interplay between schools, families, and genes is more complex than modeled in this dissertation. I argued that it is important to study schools, families, and genes together because (i) these factors are correlated and may be mistaken for each other and (ii) these factors interact. The correlations among schools, families, and genes are a problem that I only partly solve in this dissertation. I isolate the role of the classroom environment from the alternative genetic and environmental influences it may be correlated with (Chapter 2). Also, I investigate the role of schools while taking into account that the school environment is correlated with the family environment (Chapters 3, 4, and 5).[8] Moreover, in all chapters I use a twin design to separate genetic and environmental influences. However, the reality is more complex and the applications of the twin design in this dissertation only partly capture this complexity. More specifically, correlations between genes and the family and school environment are not sufficiently modeled. Depending on the type of correlation (passive, active, or evocative) and their direction they end up in different *ACE* components (see Section 1.3).

For example, I find positive associations between family SES and average educational performance and learning growth (Chapters 2-5), as well as moderating effects of family SES

8 This is common practice in sociological studies, but these studies commonly do not investigate the role of genes. Behavioral genetics studies that study the role of genes and schools often do not sufficiently control for the family environment (Hart, Soden, et al., 2013; Haughbrook et al., 2017; Taylor et al., 2010, 2020).

(Chapters 2-4). However, these associations do not provide evidence for a causal environmental influence of SES (e.g., related to economic, cultural, and social resources), as they could be genetically confounded due to the presence of passive gene-environment correlation (Hart et al., 2021). This is not problematic if one is interested in the stratification of educational outcomes by family background in a broader sense, capturing all types of (dis)advantages that are associated with it. However, it becomes more problematic if only socially caused differences are considered to be unfair and should be compensated. In that case, one would need to disentangle social and genetic transmission, which requires different designs such as a children-of-twins model (McAdams et al., 2014).

Another implication is that genetic variance likely captures unmodelled gene-environment correlations, and that I was unable to disentangle gene-environment interactions from gene-environment correlations. I find that the contribution of genetic differences to the total amount of dispersion in performance is relatively large, and this is already the case at the start of formal education. While this shows that genes matter, it should not be overinterpreted as it is unclear how the environment mediates these effects. The genetic component likely includes active and evocative gene-environment correlations (see Section 1.3). Hence, the results suggest that gene-environment interplay before the start of schooling is important. This most probably reflects gene-environment correlations in the family environment in early childhood but could also include similar processes in early childcare and education. Extensions of the CTD could provide more insight into this. For example, if twin-specific information on the environment is available (e.g., child-based reports of parenting behavior), one can estimate if there is genetic variance in this environmental measure. If so, this indicates the presence of active or evocative gene-environment correlation (Kendler & Baker, 2007). Combining this with information on educational performance in a bivariate model would allow for estimating the gene-environment correlation (i.e., genetic influences that are common to the environmental measure and educational performance). It would also allow for investigating gene-environment correlation and gene-environment interaction simultaneously (Johnson, 2007).

1.7.4 Importance of the non-shared environment

In this dissertation, I focused on the role of genetic and shared environmental variance and did not formulate expectations on the role of the non-shared environment. Non-shared environmental influences are seldom discussed in the literature on genetic differences in socioeconomic outcomes (Erola et al., 2021). However, the non-shared environment turned out to play a role in driving dispersion in educational performance, often even a more important one than the shared environment. Moreover, the non-shared environment appeared to be context dependent. The extent to which the non-shared environment plays a role in educational performance differs depending on parental SES, school SES, and school career stage. Providing more insight into non-shared environmental influences could enhance the understanding of schools and families in educational performance and inequality. As a first step, substantive non-shared environmental influences could be separated from measurement error. For example, if there are repeated measures for educational performance, measurement error can be assessed and corrected

for. When the corrected variance in educational performance is then decomposed into the different variance components, the residual non-shared environmental variance would reflect more substantive influences. These substantive influences could be systematic (e.g., differential treatment and experiences) and non-systematic (e.g., accidents, illness, luck) (cf. Plomin et al., 2001).

Second, the specific non-shared environmental influences could be identified by incorporating more individual-specific information in the twin design. Possible relevant (systematic) influences accounting for non-shared environmental differences in educational performance include differential treatment of parents and/or teachers, differential experiences of the family and school environment, and personal factors (e.g., differences in personality, effort, motivation, interest, and enjoyment) (Asbury et al., 2008; Larsen et al., 2019). This would not only provide more insight into how educational performance differences come about but could also provide more insight into social inequality in education. Non-shared environmental influences are a source of within-family inequality. However, non-shared environmental influences, such as parental treatment or academic motivation and interest, may differ by family socioeconomic background. In that case, within-family inequality affects between-family inequality (Grätz, 2018).

1.8 Reflections for policy

Several policy reflections arise from this dissertation. First, in developing and evaluating policy it should be clearer how educational inequality is defined and which sources of inequality should be reduced. Second, there should be more awareness of the presence of genetic differences between pupils and the interplay with the environment. Finally, policies should focus more on SES differences between schools and differences within schools, and less on school quality differences.

1.8.1 Define educational inequality and consider different sources of dispersion

Politicians and policymakers should more clearly define what is meant by educational inequality, and design and evaluate policy based on this definition. Inequality in educational performance could refer to the total dispersion in performance, dispersion in performance related to family background differences, and dispersion related to genetic differences. Not all kinds of inequality are always seen as problematic and unfair. Providing more clarity in this regard facilitates the implementation of more targeted policies, and thereby a more efficient use of resources.

The answer to how much dispersion in educational performance in general, and which sources of dispersion in specific, are justified likely differs between different educational outcomes at different stages in the educational career. For example, everyone should leave compulsory education with the reading and mathematics skills that are needed to function independently in society (Asbury & Plomin, 2014; Inspectie van het Onderwijs, 2023a). This would imply that dispersion, including genetic dispersion, should be minimal for basic skill levels and the school

system has a responsibility to contribute to this. Once the fundamental skills are established for all pupils, education could aspire to tap into children's inherent potential and to provide support to nurture that potential (Asbury & Plomin, 2014). This can be expected to increase genetic differences between pupils, and thereby an increased total dispersion in educational performance. People's view on the desirability of this outcome would likely vary more.

When developing educational policies and interventions, careful consideration should be given to the specific type of educational inequality these policies aim to address and the potential consequences they may have on other sources of inequality. Policies could (unintendedly) increase or decrease genetic and family background differences. As an example, consider differentiation or ability grouping, which is a popular educational practice where pupils are assigned to groups and receive instructions and guidance based on their performance level.[9] Ability grouping could have several consequences (see, e.g., Hallinan, 1988; Knigge et al., 2022). First, it could allow low-ability pupils to perform better, for example, because they get more tailored support for the areas in which they face difficulties. This would decrease genetic dispersion. Second, it could allow high-ability pupils to perform better because they are provided with more challenging educational materials, for instance. As a result, genetic dispersion would increase. In theory, both low-ability and high-ability pupils benefit from ability grouping, which would reproduce genetic dispersion and increase average performance. However, in practice low-ability pupils do not benefit as much. Instead, there is evidence that they are negatively impacted by ability grouping. This has been related to lower teacher expectations and fewer interactions with high-ability pupils, among others (Condron, 2007; Hanushek & Wößmann, 2006). If these negative consequences for low-ability children outweigh the positive effects for high-ability children, genetic dispersion can be expected to increase. Lastly, ability grouping potentially increases dispersion related to family background influences. This is because low-SES pupils may suffer from lower teacher expectations and may therefore be disproportionately placed into lower-ability groups (Condron, 2007). These different consequences show that a policy measure can affect different sources of inequality in multiple ways, and that there may be a trade-off between different sources of inequality. Hence, it is important to clearly identify the source of inequality that should be addressed by the policy, and to consider whether addressing it may have adverse effects on other sources of inequality or on average educational performance.

Policies could also intend to reduce educational inequality but may be ineffective in doing so because they do not specifically address inequality as dispersion and the underlying genetic and social sources of dispersion. Examples include earlier access to universal childcare (Inspectie der Rijksfinanciën, 2020) and providing a 'rich' school day (Ministerie van Onderwijs, Cultuur

9 In the Dutch educational system, pupils with different performance levels are grouped within classes in primary education (i.e., working with subgroups within a class) and grouped to different classes or schools in secondary education (i.e., tracking or streaming). While the term ability grouping suggest differentiation based on ability or (genetic) potential, this is in practice not the case. Instead, previously measured performance is used. This is also why some use alternative terms such as 'skill grouping' or 'attainment grouping' (Condron, 2007).

en Wetenschap, 2022) to all pupils.[10] Such 'one size fits all' policies, could increase average performance, but their effectiveness in reducing the variance is expected to be limited (Asbury & Wai, 2020). That is, it shifts the entire distribution upwards but would do little to reduce inequality as dispersion and inequality of opportunity. If the policy goal is to reduce the dispersion, more individually targeted policies would be a better use of resources. For example, providing earlier childcare and rich school days only to low-achieving and/or low-SES pupils.

More individually targeted policies would also be better given that this dissertation shows heterogeneity (or within-group inequality) for low-SES families and low-SES schools. There is more variance, especially more genetic variance, in low-SES families, low-SES schools, and lower-quality schools. Some children in low-SES families have more genetic risks towards lower educational performance and are taught by less effective teachers, while other low-SES children have less genetic risks and high-quality teachers, for instance. This suggest that it may be worthwhile to focus on reducing within-group inequality next to between-group inequality. Although there are positive associations between educational performance and family SES, school SES, and school quality, these are small compared to the range of individual differences within families and schools (Scarr & McCartney, 1983). This implies that interventions should focus on subgroups within these broader groups. For example, rich school days could be focused on those pupils at school who are most at risk, such as low-achieving pupils, low-SES pupils and/or pupils who are taught by less effective teachers.

The way schools are funded could also address this. In the Netherlands, there is an educational disadvantage policy (in Dutch: *onderwijsachterstandenbeleid*) which includes that some schools receive additional funding to reduce disadvantages in pupils' educational performance related to social, economic, or cultural causes (Walhout & Scholtus, 2022). One could ask the question to what extent schools should receive additional funding to compensate for genetic differences. A case could be made that if social background differences should be compensated, genetic differences should be compensated too. One's genetic endowment is just as undeserved as one's social background (Meyer, 2016). All children with different talents and potentials should be able to leave school with at least a fundamental level of reading, writing, mathematics, and digital skills (excluding pupils with severe disabilities) (Asbury & Plomin, 2014). Children who enter school with more genetic risks for lower performance (e.g., as partly captured by lower measured cognitive ability), require more educational support to achieve this. Hence, schools could be provided with more financial resources to ensure that they could provide this additional support. This would decrease genetic dispersion, and thus reduce the biggest source of within-group inequality among low-SES families and schools.

10 The rich school day includes extending the school day with additional activities that are focused on pupils' development in at least two of the following areas: sport, culture, cognitive development, social development, and orientation towards yourself or the world (Ministerie van Onderwijs Cultuur en Wetenschap, 2022).

1.8.2 Realize that genetic differences play a role

In education policies, genetics is seldom considered (Asbury & Wai, 2020). This is despite the robust evidence provided by prior studies and this dissertation that genetic and environmental factors – and the interplay between them – play a role in educational outcomes. Although the precise mechanisms are still to be scrutinized, recognizing the role of genetic differences in educational performance can shift perspectives on educational inequality and the corresponding policy approaches to reduce it. It may temper the expectations and responsibilities that are put on teachers and schools to reduce educational inequality. Moreover, it could bring the focus to interventions that are closer connected to the main sources of educational inequality, which are likely more effective interventions.

The discussion and development of educational policies often rely on the (implicit) assumption that differences in educational outcomes are environmental in origin and on research findings that do not consider the role of genes (Asbury & Wai, 2020; Krapohl et al., 2014). The dominance of environmental explanations in policy and the public debate lead to the situation where parents and teachers are held responsible for children's lower school performance (Asbury & Plomin, 2014). There is considerable pressure on teachers and schools to reduce educational inequality. This becomes apparent in the ideal of education as 'the great equalizer' and as 'the fundament of inequality of opportunity' (Sociaal-Economische Raad, 2021). Teachers and schools are indeed important for educational performance, and more so for children whose parents had fewer years of schooling, which is also shown by the results of this dissertation. But the results also indicate that the idea that the primary school environment could act as a great equalizer seems over-optimistic. Most of the genetic differences are already present at the start of formal education, and the extent to which genetic differences account for performance differences does not differ between schools. The realization that genetic differences play a role may lead to more realistic expectations of what schools can and cannot do and would prevent unnecessary blaming of teachers and schools.

That genetic differences play a large role, does not mean that they are fixed and immutable. Instead, they reflect the realization of genetic endowments via environmentally mediated processes (Freese, 2008). This could include processes in the school environment too. However, it should be realized that most of the genetic differences are already present at the start of formal education. Therefore, there is most to gain by intervening before formal education starts. This includes the family environment, but also kindergarten and childcare environments, for example. While genes cannot be changed, the extent to which genes drive experiences can. Policymakers could support the process of gene-environment correlations (Asbury & Wai, 2020). In high-SES families genetic potential is more likely to be expressed because there are more resources and experiences available, and parents may be more likely to adapt the rearing environment to their children's potential (Baier & Lang, 2019). In low-SES families this occurs less. However, gene-environment correlations could also take place in other institutions than the family, such as in early childhood education and care. Therefore, if low-SES children receive more (higher-quality) early childhood education and care than high-SES children, this could substitute for the aspects in their family environment that hamper effective gene-environment correlations and thereby decrease educational inequality. Policy could foster this by ensuring that (i) low-SES children

have more access to childcare than high-SES children, (ii) childcare provides a wide range of developmental opportunities, including educational experiences and materials that fit children's genetic potential and further nurture their potential (Kovas et al., 2013), and (iii) there are childcare workers who have the sensitivity and time to notice the talents and challenges of a child, and respond accordingly (Asbury & Plomin, 2014).

1.8.3 Look beyond school quality differences

Policy should be directed to the aspects of schools that appear to have the most important impact. This dissertation shows that this is not school quality, but other aspects that differ between and within schools. Instead of reducing the quality between schools to address educational inequality, the focus should be on reducing school segregation and differences between classrooms. Even though the effects of school SES and classrooms are not strong, the school environment may still be an appealing option for policy interventions directed at increasing educational performance and decreasing educational inequality (Downey & Condron, 2016). Via schools, a diverse range of children can be reached, while it is more difficult to intervene in the family environment.

Based on the results of this dissertation, the focus should be less on reducing quality differences between schools and more on reducing school segregation. Often, emphasis is placed on quality differences between schools. This becomes apparent in the importance that parents, especially high-SES parents, place on choosing the best school for their children (Borghans et al., 2015a). Also, the Inspectorate of Education monitors school quality and aims to promote quality for better education (Inspectie van het Onderwijs, 2021b). However, quality differences between schools – as measured by various indicators encompassing school resources and climate – seemed to have a limited impact on average educational performance, inequality as dispersion, and inequality of opportunity. School SES, on the other hand, does play a role. Therefore, policies directed to decreasing school segregation have the potential to reduce inequality as dispersion. These include, for example, changing the school application process or influencing the school choice behavior of parents (Bulder et al., 2020).

When addressing school segregation, it is important to keep in mind two considerations. First, as argued before, the definition of educational inequality should be clear. Desegregation of schools can be expected to decrease dispersion. Based on this dissertation, this decrease in dispersion is expected to be mainly related to a decrease in genetic dispersion. Decreased genetic dispersion could be caused by an increase of performance among children with lower abilities while not changing the performance of children with a high ability. It is a political choice whether this outcome is desirable or not. Second, one should be aware that segregation does not only occur between schools, but also within schools. Pupils and their parents are more likely to form connections with same-SES others in class (Zwier & Geven, 2023). Ideally, this form of segregation should be addressed too, especially given the finding of this dissertation that classroom differences play a role. Interventions that would address such within-class segregation could include the stimulation of cross-SES interactions via seating arrangements or group work (Gremmen et al., 2018).

The finding that differences between classes play a role could also be considered in the way school quality is evaluated. It may be more important to focus on assuring and promoting the

quality of education within schools instead of between schools. A more in-depth evaluation of the quality of teaching, as is recently done in the 'Monitor Teaching Quality' of the Inspectorate of Education, is a valuable step in this direction (Inspectie van het Onderwijs, 2023b). The variation in teaching quality is likely the key factor contributing to differences between classes (cf. Byrne et al., 2010). Therefore, promoting teaching quality is important for better education and improved educational outcomes, especially for disadvantaged pupils. To what extent this is realized depends on the equal access to high-quality teachers. Given the shortage of teachers, which is more prominent in disadvantaged schools, this requires ongoing attention in the coming years.

CHAPTER 2

Are classrooms equalizers or amplifiers of inequality?

A genetically informative investigation of educational performance

A slightly different version of this chapter has been published as:
Stienstra, K., Knigge, A., Maas, I., de Zeeuw, E. L. & Boomsma, D. I. (2023). Are classrooms equalizers or amplifiers of inequality? A genetically informative investigation of educational performance. *European Sociological Review, 39*(5), 708-723. doi: 10.1093/esr/jcac054.

Kim Stienstra: Conceptualization, Methodology, Formal analysis, Writing – Original Draft, Visualization, Funding acquisition. **Antonie Knigge:** Conceptualization, Writing – Review & Editing, Funding acquisition. **Ineke Maas:** Conceptualization, Writing – Review & Editing, Funding acquisition. **Eveline L. de Zeeuw:** Investigation, Resources, Data curation, Writing – Review & Editing, Funding acquisition. **Dorret I. Boomsma:** Investigation, Resources, Data curation, Writing – Review & Editing, Funding acquisition.

2 Abstract

We investigate the influence of the classroom environment on educational performance and whether classroom effects are dependent on parental socioeconomic status (SES). We consider that the classroom environment can have a compensatory effect and decrease educational inequality, in which case the classroom context is more important for children originating from lower-SES families. Alternatively, there can be an amplifying effect, in which case the classroom environment is more important for high-SES children, which would increase educational inequality. We investigate the two alternatives by applying a twin design to data from 4,216 twin pairs from the Netherlands Twin Register (birth cohorts 1991-2002). Some twin pairs share a classroom and other twins from the same pair are in different classrooms. We use this fact to decompose the variance in educational performance at the end of primary school into four components: genetic variance, classroom variance, shared environmental variance not related to the classroom, and non-shared environmental variance. We find that of the total variance in educational performance only a small share (around 2%) can be attributed to differences between classrooms within schools. The influence of the classroom was larger when the level of parental SES was lower (up to 7.7%) indicating a compensatory effect.

This work was supported by the Dutch Research Council (NWO) research talent grant for the project 'Quality and inequality: The compensatory and multiplicative effects of school quality' (NWO: 406-18-557, awarded to Stienstra, Knigge, and Maas) and an NWO Veni grant for the project 'Towards equal educational opportunities: The complex interaction between genes, families, and schools' (NWO: 451-17-030, awarded to Knigge). We gratefully acknowledge research program 'Consortium on Individual Development' which is funded through the Gravitation program of the Dutch Ministry of Education, Culture and Science and the NWO (NWO: 0240-001-003); 'Decoding the gene–environment interplay of reading ability' (NWO: 451-15-017); 'Netherlands Twin Registry Repository: Researching the interplay between genome and environment' (NWO: 480-15-001/674); 'Twin-family study of individual differences in school achievement' (NWO: 056-32-010); 'Longitudinal data collection from teachers of Dutch twins and their siblings' (NWO: 481-08-011); 'KNAW Academy Professor Award' (PAH/6635).

We are grateful to the twin families and the teachers for their participation. Earlier versions of this work were presented at the ISA RC28 spring meeting (Frankfurt, 2019), Dag van de Sociologie (Amsterdam, 2019), Social Stratification Research Seminar (Amsterdam, 2019), Migration and Social Stratification Seminar (Utrecht, 2019), and ICS spring day (Nijmegen, 2019). We thank the participants at these meetings for their valuable comments.

2.1 Introduction

Do schools reduce social inequality in educational performance and serve as 'the Great Equalizer'? Or do they reproduce or even amplify inequalities? Over the past century, these questions have been a core focus of social science scholars and central to the debate on educational policies and reforms (Coleman et al., 1966; Downey et al., 2004; Downey & Condron, 2016). There are diverging ideas on how children from different socioeconomic status (SES) backgrounds benefit from their school environment. On the one hand, schools could amplify inequality. Even if low-SES and high-SES children attend the same school, various opportunities and practices in the classroom may favor the educational performance of high-SES children. For example, high-SES students could profit more from the materials and lessons in effective classrooms because they enter the school with better academic preparation (Hanselman, 2018; Stanovich, 1986). On the other hand, schools could reduce inequality (i.e., a compensatory effect). Learning opportunities in schools and families overlap and could substitute for each other. An advantageous classroom environment would then benefit students from less resourceful family environments more (Chiu & Khoo, 2005; Hanselman, 2018).

Given the concerns about (increasing) educational inequality in many countries (OECD, 2019), it is important to understand to what extent the classroom context works as an equalizer and is part of the solution, or contributes to inequality and is thus part of the problem. Therefore, we ask: 'To what extent are classroom effects contingent on SES background?' Empirical studies did not provide a conclusive answer to this question yet (e.g., Alexander et al., 2007; Downey et al., 2004, 2019; Downey & Condron, 2016; Von Hippel et al., 2018). The reasons for the mixed findings remain unclear. They do not seem to be dependent on students' age or country, for example.

Investigating classroom effects on educational performance is complex. The quality of teachers and the classroom environment is difficult to observe (Hanushek & Rivkin, 2006). For example, measures for teacher quality (e.g., teachers' experience and educational level) often have a small effect on educational performance (Rivkin et al., 2005). While it could be that the importance attributed to teacher quality is overstated, it could also be that measurable characteristics cover only a small part of the true variation in teacher quality (Rivkin et al., 2005). Another challenge is that classroom effects can be biased when alternative influences on educational performance are not sufficiently controlled for. For example, high-SES parents more often choose high-quality schools for their children (Borghans et al., 2015a). Also, children's educational performance is genetically influenced (De Zeeuw et al., 2019), and the composition of classrooms concerning genetic predisposition may not be independent of the classroom environment. For example, children with a higher polygenic score for learning could be concentrated in higher-quality classrooms (cf. Smith-Woolley et al., 2018). The classroom environment may capture influences of children's socioeconomic background or genetic predisposition if those influences are not sufficiently considered.

Prior research tries to deal with such issues by considering so-called value-added models. These models use students' prior achievement to estimate the contribution ('added value') of the teacher or school to students' progress in achievement over time (OECD, 2018). Such value-added effects can be adjusted by student- and school-level covariates to further control for

pre-existing differences between students, although it is debated to what extent such covariates should be included (Marks, 2021). Value-added models are often argued to provide a suitable way to estimate the contribution of classrooms or schools separate from alternative influences (see, e.g., Wayne & Youngs, 2003), but they also come with disadvantages. One of these, which has only been limitedly acknowledged, is that value-added effects can be genetically influenced and may therefore provide biased estimates of the contribution of teachers and schools (Morris et al., 2018).

In this study, we take genetic influences into account by using a suitable method that is rarely applied in this research area. We estimate overall classroom effects and their dependency on parental SES by comparing the educational performance of twins who attend the same or different classrooms. With this method, classroom influences are not directly measured but rather a latent factor that can be estimated from the data, in addition to other latent factors capturing the contribution of genes, the shared environment (i.e., influences making twins raised in the same family more similar), and non-shared environment (i.e., influences making twins raised in the same family dissimilar) (e.g., Plomin, DeFries, McClearn, & McGuffin, 2008). With this twin design, we can detect classroom effects that suffer less from insufficient measures and bias due to unobserved influences. An additional advantage of the twin design is that we do not only isolate classroom effects but also pinpoint the sources of alternative influences, i.e., genetic, shared environmental, or non-shared environmental influences.

Two prior studies used this approach to estimate classroom influences in Australia and the United States (Byrne et al., 2010; Grasby et al., 2020).[11] They found that 0-9% of the variance in literacy and numeracy could be accounted for by the classroom environment. We take the next step by investigating if and how classroom effects depend on parental SES. In this way, we can look at whether the classroom environment affects educational performance in an SES-neutral, equalizing, or amplifying way.

We do so by analyzing data on 4,216 twin pairs from the Netherlands Twin Register (NTR). We use the score on the nationwide standardized achievement test (i.e., *Cito* test) to measure educational performance. Children take this test at the end of primary school around age 12. The test score, combined with the teacher's recommendation, determines the enrolment in a secondary educational track. The *Cito* score is thus important for children's educational careers. A key aspect of primary education in the Netherlands is that within-school tracking or academic streaming – a practice where classroom allocation is based on prior educational performance – is uncommon. This means that being in the same classroom does not depend on similarity in educational performance between twins, making the Dutch context well suited for investigating classroom effects with the twin method.

11 There are a few other twin studies that also provide information on twins in the same and different classes, but that do not explicitly test for classroom variance (Eifler et al., 2019; Kovas et al., 2007).

2.2 Theory

2.2.1 Classroom influences

A classroom variance component derived from twin models is an omnibus measure for several interrelated ways through which the classroom environment could affect educational performance. Although we do not test the specific influences, we provide an overview of ways through which the classroom environment could contribute to educational performance.

The first route is via the *quality of teachers*. Teachers play a key role in the educational performance of students, either directly (e.g., via instructions) or indirectly (e.g., contributing to a safe and orderly classroom climate). How well teachers do their job and thereby facilitate student achievement, is difficult to observe (Hanushek & Rivkin, 2006). Hence, much of the literature and empirical evidence is based on observable teacher characteristics that are thought to affect teachers' performance, such as knowledge and experience (Hanushek & Rivkin, 2006; Parcel & Dufur, 2001). These studies show that teacher experience matters, but generally only in the first years of their employment. After three to five years of experience, additional experience does not explain additional variation in students' educational performance (see Coenen et al., 2018). Teachers' cognitive skills have been found to be positively associated with student performance (Hanushek et al., 2019), although there are indications that it matters for students' math performance but not for reading (Coenen et al., 2018). Lastly, teaching certificates and advanced degrees (i.e., having a Master's or PhD degree relative to a Bachelor's degree) are in general not associated with students' performance (Coenen et al., 2018; Hanushek & Rivkin, 2006; Wayne & Youngs, 2003).

A second classroom feature that affects educational performance involves the available *resources* in the classroom. These are, for example, equipment, smaller class sizes, and a lower student-teacher ratio. Smaller class sizes and a better student-teacher ratio are thought to reduce resource dilution effects: students have fewer classmates to "compete" for teachers' time and attention (Parcel & Dufur, 2001). Furthermore, it has been argued that in smaller classes there are fewer disciplinary problems and therefore there is more instructional time and greater opportunity to learn (e.g., by having more time for individualized instructions) (see Blatchford & Russell, 2020). In empirical studies, indicators such as per-pupil expenditure, student-teacher ratio, and class and school size have been used. Per-pupil expenditure shows a consistent positive association with educational performance (Greenwald et al., 1996), whereas for student-teacher ratio, class size and school size positive but also no or negative relations with educational performance are found (Blatchford & Russell, 2020; Greenwald et al., 1996).

A third aspect is the *climate* in the classroom. When there is an academically oriented environment with high expectations, this signals certain standards about schoolwork and ideal students' performance, which is thought to affect educational performance (Anderson, 1982). Also, when there is a cohesive community with dense social ties this could positively benefit students' performance (Dronkers & Robert, 2008; Parcel & Dufur, 2001). Such a cohesive community could affect students' sense of belonging, which has been shown to relate to academic performance (Ma, 2003). Empirical studies show that the climate in classes and schools contributes to educational

outcomes, with some evidence that it does not only influence short-term performance but that its effect seems to persist for years (Thapa et al., 2013).

Lastly, *peers* in the classroom can play a role. Students' educational performance can be influenced by their peers in different manners. One way is via direct interactions in learning. This does not only relate to, for example, figuring out a solution to an exercise, but also to processes such as developing vocabulary and obtaining knowledge that other students gained from museum visits, traveling, and so on (Hanushek et al., 2003; Kahlenberg, 2001). Additionally, peers influence each other's motivation, aspirations, and attitudes toward education (Hanushek et al., 2003; Kahlenberg, 2001; Raabe & Wölfer, 2019). Furthermore, peers affect the extent to which an environment is conducive to learning (Kahlenberg, 2001). For example, if other students show disruptive behavior, this affects the classroom processes and lowers the learning opportunities for all students in the class (Lazear, 2002). Students who are highly motivated and skilled, on the other hand, contribute to a learning-oriented peer culture (Rumberger & Palardy, 2005). Numerous empirical studies show that peer influences on students' educational outcomes (e.g., grades, standardized test scores) exists, as well as peer influences on other outcomes that contribute to educational performance including motivational outcomes (e.g., goals, efficacy, interest) and positive behavioral styles (e.g., cooperative behavior, self-confidence) (see Wentzel & Muenks, 2016). Given these mechanisms and prior empirical studies, we expect that the classroom context influences children's educational performance.

2.2.2 The moderating role of parental SES

Children's educational performance is strongly affected by their family background. Children from a high-SES background perform on average better in school. This may be because of multiple reasons, including genetics, and the financial, cultural, and social resources high-SES parents can invest in children's educational success (Blau & Duncan, 1967; Bourdieu & Passeron, 1977).[12] We examine whether parental SES also affects the extent to which the classroom environment matters for educational performance.

One possibility is that there is a *compensatory effect* that leads to reduced educational inequality. In this case, the educational performance of children from low-SES backgrounds will depend more on their school environment than the performance of children from high-SES backgrounds. A reason to expect a compensatory effect is that learning opportunities within schools overlap with those within socioeconomically advantaged families and can substitute for each other (Chiu & Khoo, 2005; Hanselman, 2018). A low-quality classroom environment may be less harmful to high-SES students because the fewer learning opportunities in such classes can be substituted by parental resources (e.g., providing tutoring), while low-SES parents cannot provide such compensation (Hanselman, 2018). In a similar vein, a higher-quality classroom environment may be more beneficial for students from less advantaged family backgrounds. Low-SES students may be more susceptible to the supportive environment in advantageous classrooms because they are less likely to find supportive influences at home (Coleman et al.,

12 Because parental genes can influence both parental SES and children's educational performance (via children's genes), the association between parental SES and children's educational performance may be partly spurious. We come back to this issue and its implications in the Conclusion and Discussion section.

1966; Jencks & Mayer, 1990; Rumberger & Palardy, 2005). For example, children in low-SES families receive less language stimulation at home as low-SES parents, in general, tend to expose their children to less vocabulary and grammar, read to them less, and purchase fewer reading materials for their children (Bradley & Corwyn, 2002). These children may benefit more from the available books in the classroom and interactions with peers that contribute to developing vocabulary.

There are also reasons to expect the opposite pattern: the classroom environment may be especially beneficial for high-SES students. This refers to an *amplification effect*, which has also been referred to as multiplication and multiplicative accumulation (Erola & Kilpi-Jakonen, 2017a) or cumulative advantage (DiPrete & Eirich, 2006). High-SES students may take greater advantage of the classroom environment because they enter this environment with better academic preparation (Hanselman, 2018). Taking language stimulation as an example, high-SES students enter school with a stronger vocabulary. They may therefore benefit more from the learning opportunities in school than low-SES students, because they may understand instructional materials better and learn more from particular lessons than low-SES students (Hanselman, 2018; Stanovich, 1986).

Another consideration relates to the cultural correspondence between the classroom and the home environment. From a cultural reproduction perspective (Bourdieu, 1986; Bourdieu & Passeron, 1977), it can be expected that the high-status cultural signals of high-SES children (e.g., behaviors, tastes, and attitudes) are positively evaluated by teachers and that these children may experience a greater sense of belonging in their class, leading to better educational performance (De Graaf et al., 2000). This could be especially the case in high-quality classrooms. In such classrooms, the climate is more ambitious and academically oriented, which coincides with high-SES parents' expectations and ambitions for educational success. This may further increase high-SES students' educational performance. For low-SES students, on the other hand, such a classroom environment means a mismatch between their family background and classroom experiences which may lead to negative self-perceptions and emotional distress, negatively affecting educational outcomes (Crosnoe, 2009). Additionally, it can be expected that there is a stronger influence of the transmission of norms if multiple actors socialize students in the same way (Centola & Macy, 2007). Therefore, if the culture at home and in the classroom matches and both parents and teachers emphasize, for example, the importance of educational performance, students may be more likely to internalize this norm and behave according to it.

Empirical tests of compensation and amplification effects based on a twin design have not yet been performed. Prior empirical studies investigate compensation and amplification effects through an interaction between parental SES and measured classroom (or school) characteristics, or by investigating the 'overall' contribution of classrooms or schools (e.g., with a value-added approach), all with mixed results. Studies with measured characteristics focused on many different aspects of the classroom and found no interaction with SES (e.g., concerning teacher quality: Akiba, LeTendre, & Scribner, 2007; Borman & Kimball, 2005), compensatory effects (e.g., concerning student composition and climate, see: Berkowitz et al., 2015; Gustafsson et al., 2018; Opdenakker & Damme, 2001, and amplification effects (e.g., concerning resources, climate, and teacher quality, see: (Atlay et al., 2019; Gustafsson et al., 2018; Scholten & Wolbers, 2018).

The results of studies that investigated the overall classroom effect as an omnibus measure capturing all the influences are also inconclusive. For example, the value-added study by Kyriakides et al. (2019) shows that more effective classrooms and schools compensate for the socioeconomic achievement gap. Additionally, it has been reported that the absence of schooling, for instance during the summer break (e.g., Alexander et al., 2007; Downey et al., 2004) or the COVID-19 pandemic (e.g., Engzell et al., 2021), increases the SES gap in educational outcomes. The finding that the SES gap is smaller when schools are in session than when school is not in session has led to conclude that schools have a compensatory effect. Yet, there are also studies that challenge this conclusion. The findings of seasonal comparison studies do not always replicate (Von Hippel et al., 2018, 2023; Von Hippel & Hamrock, 2019). Additionally, the study by Hanselman (2018) used a value-added approach and found no interplay with economic background, which suggests that classrooms and schools neither amplify nor compensate for socioeconomic inequality, but rather reflect existing inequality.

Since there are neither clear theoretical nor empirical arguments to favor either compensation or amplification, we consider two alternative hypotheses: *the classroom environment has a larger influence on educational performance if students' SES background is lower (H1 - compensation)* and *the classroom environment has a larger influence on educational performance if students' SES background is higher (H2 - amplification).*

2.3 Data and methods

2.3.1 Data

We analyze twin data collected by the Netherlands Twin Register (NTR), which was established in 1986 by recruiting twins and multiples a few weeks or months after birth. The NTR is still ongoing and registers around half of all newborn multiples in The Netherlands (Boomsma et al., 2002). For young twins, twins' parents receive a survey at registration and when twins are 2, 3, 4/5, 7, 9/10, and 12 years old. After obtaining parental consent, twins' teachers receive surveys when the twins are 7, 9/10, and 12 years old. More details on the NTR are reported elsewhere (see Ligthart et al., 2019; van Beijsterveldt et al., 2013).

We include twins from birth cohorts 1991–2002. For these cohorts, *Cito* data are available for at least one of the twins for 5,672 twin pairs (1,943 monozygotic [MZ] pairs and 3,729 dizygotic [DZ] pairs). The NTR determines the zygosity of twins based on questionnaire items and on DNA or blood group polymorphism. The questionnaire items allow for determination of zygosity with an accuracy of 97% (Ligthart et al., 2019).

To determine whether twins were in the same classroom or not, we rely on the information from the mother's, father's, and teacher's reports when twins were 12 years old. Changing classrooms is not very common in the Netherlands. The large majority of pupils share their classroom with the same children throughout their primary school career (Polderman et al., 2010). We initially rely on mother's report. If data were missing, we took father's report, and if both parents' reports included missing data we used the teacher's report. Parents were asked which school situation is or was most applicable: (1) same class, (2) same school, parallel class, (3) same

school, different class, (4) different school, (5) partly same, partly different. The class information from the teacher report was also measured when twins were around 12 years old, but in a less specific way than in the parents' report as only a distinction between 'in the same class' or 'not in the same class' was made. We excluded those twin pairs for whom it was unclear whether they were in the same class or not, either due to missing data (N_{pairs} = 897, 15.8%) or because they were partly in the same and partly in different classes (N_{pairs} = 333, 5.9%), leading to a sample of 4,442 twin pairs. Additionally, we repeated the analyses after also excluding twins who went to different schools (N_{pairs} = 175, 3.1%) and those who were in the same school but in different grades (N_{pairs} = 657, 11.6%). The conclusions remain the same (see Appendix B2 for the results).

Lastly, we excluded twin pairs with missing information on parental SES (N_{pairs} = 226) as our model cannot deal with missingness on the moderator. We used Full-Information Maximum Likelihood (FIML) estimation (Arbuckle, 1996) to handle missing values for all other variables. Our final sample consists of 4,216 twin pairs (880 MZ twin pairs in the same class, 596 MZ twin pairs in different classes, 1,444 DZ twin pairs in the same class, and 1,296 DZ twin pairs in different classes).

2.3.2 Measurements

To measure our dependent variable, *educational performance*, we use students' scores on the *Cito* test. These scores were initially obtained via teacher reports and later also via the parents' surveys and via children's self-report at later ages. The *Cito* scores from the different sources are highly correlated: the correlation between the *Cito* scores reported by parents and twins is .98 and between teachers and twins is .93 (Van Beijsterveldt et al., 2013). The *Cito* test is a nationwide standardized educational achievement test that is taken at the end of primary education (around age 12). It consists of multiple-choice items on Dutch language, mathematics, study skills, and world orientation (e.g., geography, biology, and history). The domains are combined into a total score using Item Response Theory, and this score is standardized on a scale from 501 to 550. Because the subdomain world orientation is not mandatory, this is not included in the calculation of the total score. The national average is a score of 535, with a standard deviation (SD) of 10. Our sample has a somewhat higher average and lower SD (see Table 1). Means and variances of the *Cito* scores for twin 1 (typically the firstborn twin) and twin 2 are not statistically significantly different.

We measure parental SES by *parental education,* which is the most stable and most important indicator of parental SES when predicting children's educational performance (Sirin, 2005). We use the information on the mother's and father's highest educational level from the parents' survey when the twins were around 10 years old. When the mother's or father's education at age 10 was missing, we used information from the survey for younger twins (ages 7, 3, and 1). The original variable from the parent survey for 10- and 7-year-old twins consisted of 13 categories ranging from elementary school to post-graduate degree or PhD. The variable measured at twins ages 3 and 1 consisted of five categories ranging from elementary school to scientific education. We converted these categories into scores on the International Standard Level of Education (ISLED) scale (Schröder & Ganzeboom, 2014). The ISLED is a well-validated continuous measure of education with a range from 0 to 100, which allows for comparison across surveys and countries. We coded the categories into the highest level of finished education using the online appendix

(see ISLED, 2014). We used the average ISLED score when multiple values were applicable (e.g., for the category combining higher general secondary education 'havo' and pre-university education 'VWO'). An overview of the coding is presented in Table B1.1 (Appendix B1).

In all analyses, we control for *sex* (girls = 0, boys = 1) and *age* (in years) when the *Cito* test was taken. While most children take the test when they are 12 years old, there is variation in age (see Table 2.1) which we want to correct for.

Table 2.1 Descriptive statistics for MZ twins in the same classroom, DZ twins in the same classroom, MZ twins in different classrooms, and DZ twins in different classrooms.

	Same classroom					Different classrooms				
Variable	**Mean**	**SD**	**Min.**	**Max.**	***N***	**Mean**	**SD**	**Min.**	**Max.**	***N***
MZ twins										
Cito twin 1	538.07	8.27	508	550	872	538.42	8.33	509	550	562
Cito twin 2	537.57	8.49	505	550	866	538.36	8.20	510	550	552
Boy twin 1	0.44	-	0	1	880	0.49	-	0	1	595
Boy twin 2	0.44	-	0	1	880	0.49	-	0	1	595
Education parents	63.95	17.64	22.98	92.63	880	68.72	17.33	22.98	92.63	596
Age	12.26	0.49	11	14	740	12.27	0.52	11	14	472
$r_{cito1,cito2}$	.80					.77				
DZ twins										
Cito twin 1	537.82	8.83	503	550	1,415	537.78	8.46	510	550	1,132
Cito twin 2	537.74	8.60	501	550	1,416	537.73	8.34	509	550	1,110
Boy twin 1	0.49	-	0	1	1,444	0.52	-	0	1	1,294
Boy twin 2	0.47	-	0	1	1,443	0.53	-	0	1	1,296
Education parents	64.80	17.29	22.98	92.63	1,445	67.77	17.55	22.98	92.63	1,296
Age	12.25	0.47	10	14	1,181	12.28	0.51	11	14	936
$r_{cito1,cito2}$	.47					.44				

Note: For twin-specific variables, the *N* refers to the number of individuals. For the twin pair variables (education parents and age), the *N* refers to the number of pairs.

2.3.3 Twin model

The classical twin design, as shown in Figure 2.1, decomposes the variance in educational performance into variance due to three latent factors: additive genetic influences (*A*), common or shared environmental influences (*C*), and non-shared or unique environmental influences including measurement error (*E*) (Franić et al., 2012; Knopik et al., 2016). Twin data enable disentangling the different variance components, as twins differ in genetic relatedness (MZ twins share all of their genes at conception, DZ twin pairs share on average half of their segregating genes) and MZ and

DZ twin pairs are assumed to share their environment to the same extent. Hence, the covariance in educational performance between twin 1 and twin 2 is $Cov_{MZ} = a^2 + c^2$ for MZ twins and $Cov_{DZ} = 0.5a^2 + c^2$ for DZ twins. A larger similarity in performance for MZ twins than DZ twins is consistent with a hypothesis that genetic influences play a role. If MZ twins are not alike, the source of this dissimilarity is by definition the non-shared environment.

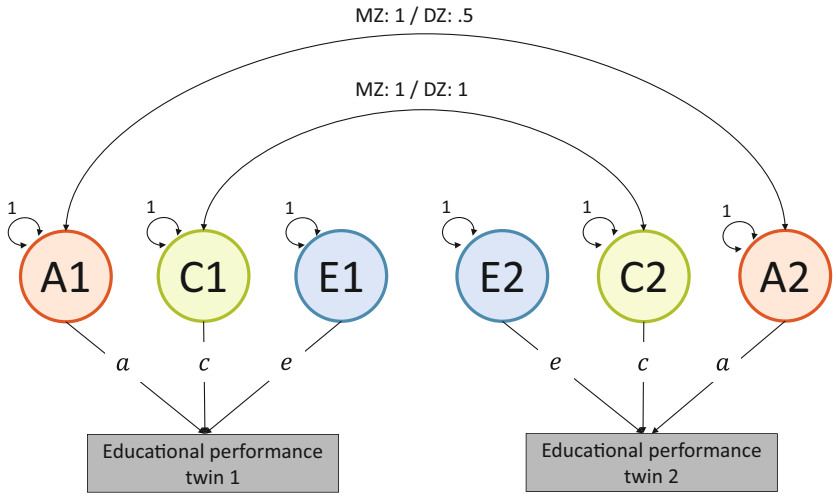


Figure 2.1 The Classical Twin Design (CTD).

Note: The CTD is fitted to data from monozygotic (MZ) and dizygotic (DZ) twins. The latent variables (circles) represent genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) factors. Their influence on educational performance (observed variables: squares) is given by path coefficients *a*, *c*, and *e*.

We can estimate a fourth variance component capturing classroom influences (*CL*) because we have measured information on whether twins attend the same classroom in primary school or not. When twins share their classroom, classroom influences may make twins more similar in their educational performance. Classroom influences make twins dissimilar if they are in different classrooms. The difference between twin correlations of twins in the same and different classrooms is the basis to estimate classroom influences (Byrne et al., 2010). As can be seen in Figure 2.2, the covariances in educational performance between twins for the different groups are:

$$Cov_{MZSC} = a^2 + c^2 + cl^2 \text{ for MZ twins in the same classroom} \quad (2.1)$$

$$Cov_{MZDC} = a^2 + c^2 \text{ for MZ twins in different classrooms,} \quad (2.2)$$

$$Cov_{DZSC} = 0.5a^2 + c^2 + cl^2 \text{ for DZ twins in the same classroom,} \quad (2.3)$$

$$Cov_{DZDC} = 0.5a^2 + c^2 \text{ for DZ twins in different classrooms.} \quad (2.4)$$

Estimating classroom variance changes the interpretation of *C* and *E*. These components now refer to the shared environmental and non-shared environmental influences not related to the classroom. The classroom component mainly captures within-school differences, because most twins attend the same school. Since between-school differences in the classroom environment are very limited in our case, this may lead to an underestimation of the classroom effect compared to regular multilevel studies of children in classes. This affects the interpretation of the results but does not prohibit testing our hypotheses.

We use multigroup Structural Equation Modelling (SEM) to decompose the variance in educational performance into *A*, *C*, *CL*, and E. The model includes the measured educational outcome for twin 1 and twin 2 for the groups: MZ twins in the same classroom (MZSC), DZ twins in the same classroom (DZSC), MZ twins in different classrooms (MZDC), and DZ twins in different classrooms (DZDC). The latent factors *A*, *C*, *CL*, and *E* are set to a variance of one to identify the model. The path coefficients *a, c, cl,* and *e* represent the effects of the standardized latent factors on the observed outcome. The total variance of educational performance is given by summing all squared path estimates:

$$V_{total} = V_A + V_C + V_{Cl} + V_E = a^2 + c^2 + cl^2 + e^2 \quad (2.5)$$

The variance components can be standardized. For example, the proportion of the genetic variance component to the total variance in educational performance, which is called heritability, is given by:

$$\frac{a^2}{a^2 + c^2 + cl^2 + e^2} = \frac{V_A}{V_{total}} \quad (2.6)$$

The contribution of genetic and environmental factors can depend on a moderator, in our case parental SES (see Figure 2). For example, *cl* becomes $cl + b_{cl}M$, where *M* is the level of parental SES. The total variance in this moderation model becomes:

$$V_{educ|M} = \left(a + b_aM\right)^2 + \left(c + b_cM\right)^2 + \left(cl + b_{cl}M\right)^2 + \left(e + b_eM\right)^2 \quad (2.7)$$

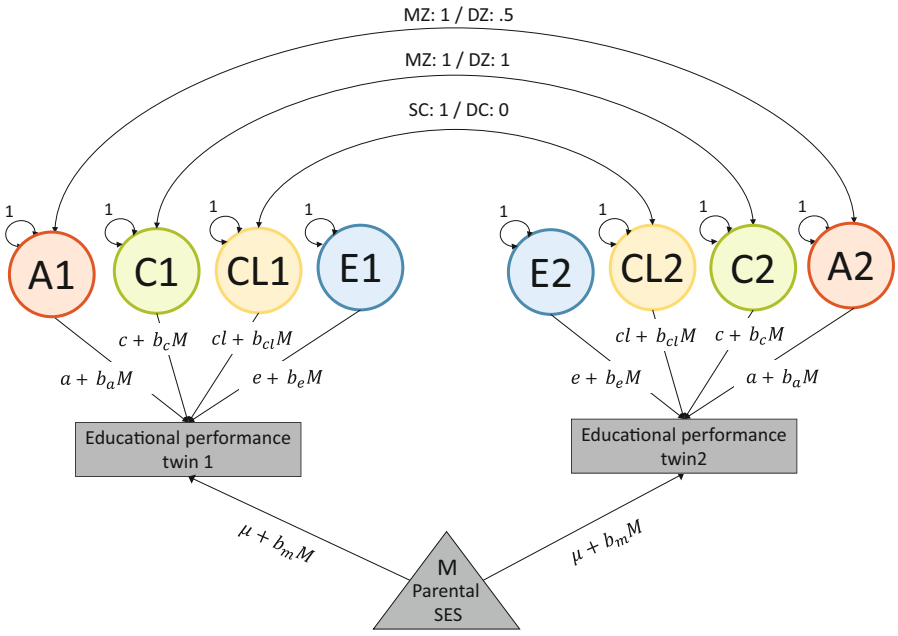


Figure 2.2 *ACE* moderation model extended with a classroom factor (*CL*).

Note: The latent variables represent genetic (*A*), shared environmental (*C*), classroom (*CL*), and non-shared environmental (*E*) factors, with the corresponding path coefficients *a*, *c*, *cl*, and *e*. These are estimated from data in monozygotic (MZ) and dizygotic (DZ) twins who are in the same classroom (SC) or different classrooms (DC). The correlation between CL1 and CL2 equals 1 if twins share a classroom, otherwise it is zero. *M* stands for Moderator and is a measure of parental SES, which is also included in the model as a fixed effect (triangle). The model also includes fixed effects of age and sex (not shown to avoid clutter).

2.3.4 Analytical strategy

We fit a series of models in Mplus using FIML estimation. In all models, we control for age and sex. We do not take nesting into classes and/or schools into account, partly because of the practical reason that we do not have a school identifier for 60% of the twins, and partly because the number of twin pairs in the dataset per school is low (on average 1.5 for those with a school identifier).

We decompose the variance in educational performance with and without a classroom component. This shows how the classroom component is captured by the *C* and *E* components if

it remains unmodelled. Then, we include our measure for parental SES, which is expected to have a main effect on educational performance and explain part of the total variance. More specifically, because parental SES is measured at the family level and always shared between twins, it can only explain shared environmental variance (Turkheimer et al., 2005). In our final model, we allow the paths to be moderated by parental SES in a continuous, linear gene-environment moderation model (Purcell, 2002). With this model, we test whether there is a compensatory (*H1*) or amplifying (*H2*) effect. If compensation takes place, classroom level variance would be smaller the higher the parental SES (i.e., b_{cl} is negative), whereas in the case of amplification classroom level variance would be larger with higher levels of parental SES (i.e., b_{cl} is positive). Given that compensation and multiplication could occur simultaneously, a negative interaction effect means that compensation is stronger than possible amplification effects (the reversed is true when we find a positive interaction). Lastly, we elaborate on how selection into school classes may bias our results.

Before fitting the twin models, we check whether the means and variances of educational performance are equal for twin 1 and twin 2 (first and second born), MZ and DZ twins, and twins in the same and different classrooms. Therefore, we fit a saturated model that describes the data with no free parameters left (i.e., no constraints on means and variances) and compare this with models that included the constraints using likelihood-ratio tests. The assumptions of equal means and variances are met.

2.4 Results

2.4.1 Decomposition of educational performance

The decomposition of educational performance into the *A*, *C*, *CL*, and *E* components is presented in Model 2 of Table 2.2. The total variance in educational performance is 72.00 (= 6.99^2 + 2.70^2 + 1.24^2 + 3.79^2). The classroom variance is $V_{CL} = 1.24^2 = 1.54$. While for classroom variance the path loading of 1.24 is statistically significant ($p = .022$), the variance of 1.54 is not ($p = .252$). This may be because the power for estimating a variance component is smaller than for a path loading. Standardizing this variance component shows that individual differences in educational performance are for 2.1% related to differences in classroom environment (mainly) within schools (V_{CL}/V_{total}). Additionally, there is statistically significant shared environmental variance ($V_C = 2.70^2 = 7.29$) and non-shared environmental variance ($V_E = 3.79^2 = 14.33$), making up 10.1% and 21.0% of the total variance in educational performance, respectively. The largest source of variance in educational performance is genetic ($V_A = 6.99^2 = 48.85$). This leads to an estimated heritability of 67.8%.

If classroom variance is unmodelled, it is captured in the *C* and *E* components, because part of the twin pairs share the same classroom while the other part does not. Model 1 (Table 2) shows that if the classroom variance is not modeled, the shared environmental variance is $V_C = 2.88^2 = 8.29$ and the non-shared environmental variance is $V_E = 3.86^2 = 14.913$. Comparing these shared and non-shared environmental variances with those in Model 2 ($V_C = 7.29$ and $V_E = 14.33$, respectively) shows that the classroom variance is to a larger extent captured in the *C*

component. This is as expected given that the twins in our sample were more often in the same than in different classrooms.

We can use the difference in the *C* variance components from Models 1 and 2 to calculate the extent to which shared environmental influences can be attributed to shared classroom experiences. Around 11.5% of the differences in educational performance can be attributed to shared environmental influences if classroom effects are not considered and this is 1.4 percentage points lower (10.1%) if we take classroom effects into account. This shows that 1.4/11.5 = 12% of the shared environmental influences can be attributed to the same classroom experiences. One could do the same for the non-shared environment (*E*) and different classroom experiences, but this is not very informative given that *E* does not only include different environmental influences and experiences but also measurement errors.

2.4.2 The role of parental SES

Model 3 (Table 2.2) includes the main effect of parental education, our measure for parental SES, which has a positive significant effect on educational performance, where each unit increase in parental education is associated with a 0.15-point increase in *Cito* score (b = 0.15, p <.001). The standardized effect (not shown in Table 2.2) is .32, meaning that an SD increase in parental education is associated with almost a third of an SD increase in educational performance (i.e., a medium effect). Parental education entirely explains the (non-classroom) shared environmental variance.[13]

In Model 4 (Table 2.2), we allow the variance components to be moderated by parental education. These interaction effects are indicated by b_a, b_c, b_{cl}, and b_e. If there is a compensatory effect (*H1*), we would expect classroom influences to become smaller with increasing parental education (i.e., b_{cl} is negative). In case of an amplification effect (*H2*), we expect to observe the opposite, namely larger classroom influences with increasing parental education (i.e., b_{cl} is positive). Our findings point towards a compensatory effect, as there is a negative significant moderation of the classroom effect (b_{cl} = -0.03, p = .006). The classroom variance in this moderation model is computed by $V_{CL}|M = \left(cl + b_{cl}M\right)^2$. For children with the lowest educated parents (i.e., primary education – ISLED 22.98), the estimated classroom variance is V_{CL} = (3.43 + (-0.03 × 22.98))^2 = 7.63. As can be seen in Figure 2.3, the classroom level variance decreases and almost approaches zero (V_{CL} = 0.55) for the highest educated parents (i.e., postdoctoral education – ISLED 92.63).

13 Including parental education also leads to decreased model fit. This would be less the case if we would drop the *C* component. We did not do so because the *C* variance may still turn out to be of importance for part of the parental education scale when we estimate the moderation model in the next step. Moreover, it has been argued that presenting more parsimonious models where *C* is dropped lead to oversimplified models with overestimation of the *A* component (Verhulst et al., 2019).

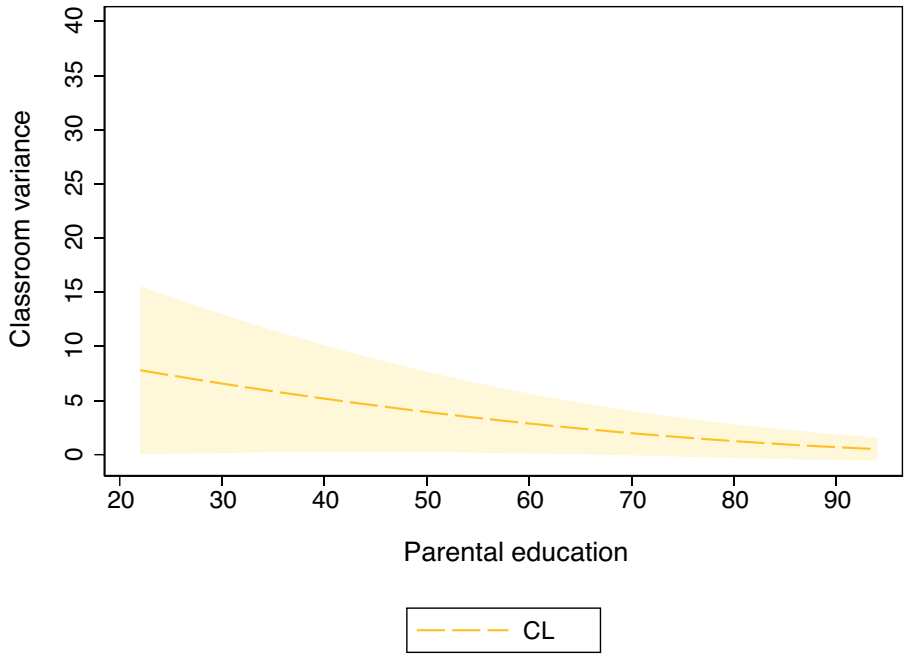


Figure 2.3 Unstandardized classroom variance in educational performance moderated by parental education (ISLED) including 95% confidence interval.

Note: Results based on Table 2.2 Model 4.

We find that classroom variance reduces with increasing parental education, but also the total variance in educational performance decreases with higher levels of parental education (see Figure 2.4a). To investigate whether the compensation pattern appears because of the decreasing total amount of variance, we also look at the standardized results, where this is taken into account (see Figure 2.4b). The standardized results also show a pattern of compensation. We previously showed that for children with the lowest educated parents the estimated classroom variance is 7.63. Dividing this by the total variance in educational performance among these children (98.89), shows that 7.7% of the variance in educational performance can be attributed to the classroom context. For the average parental education (higher and pre-scientific secondary school – ISLED 67.11), 3.5% can be attributed to classroom effects, and for children with the highest educated parents (postdoctoral education – ISLED 92.63) 1.2%.

Although we did not hypothesize on the moderation of other variance components, the results show that there are statistically significant negative moderations of unstandardized genetic influences (b_a = -0.03, p < .001) and non-shared environmental influences (b_e = -0.02, p = .002). Yet, these largely disappear when we look at the standardized components (see Figure 2.4b). With increasing parental education, the relative genetic variance of educational performance and non-shared environmental variance increase a little bit but are largely stable around .75 and .20,

respectively. Shared environmental variance is not dependent on parental education: it is entirely explained by parental education for all levels of parental education.

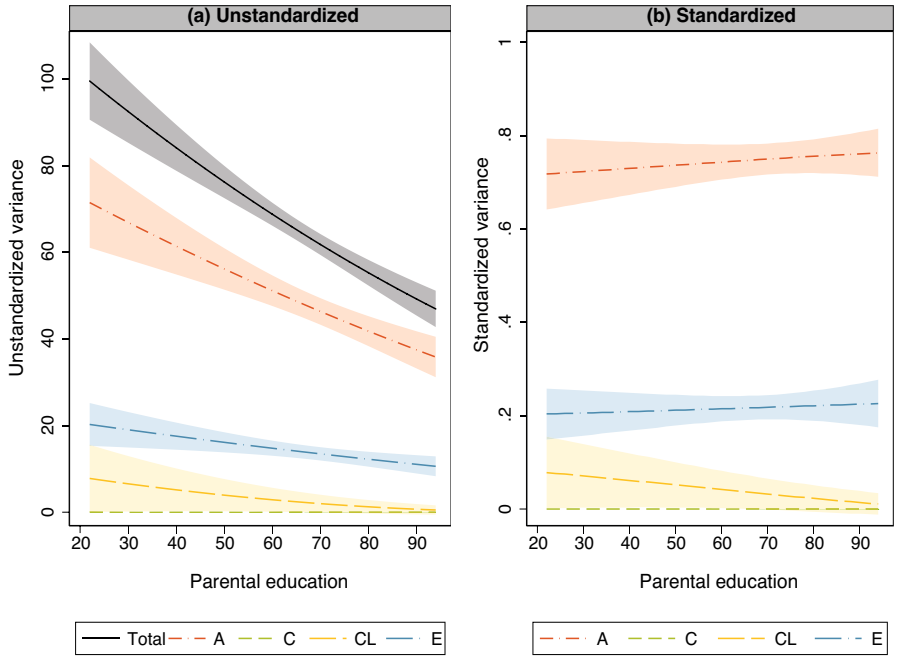


Figure 2.4. Decomposition of the (a) unstandardized and (b) standardized variance in educational performance moderated by parental education (ISLED).

Note: Sources of variance include genetic (*A*), shared environmental (*C*), classroom (*CL*), and non-shared environmental (*E*) variance. Results based on Table 2.2 Model 4.

Table 2.2 Results of twin models for educational performance for MZ twins in the same classroom (N_{pairs} = 880), DZ twins in the same classroom (N_{pairs} = 1,444), MZ twins in different classrooms (N_{pairs} = 596), and DZ twins in different classrooms (N_{pairs} = 1,296).

	Model 1		Model 2		Model 3		Model 4	
	Estimate	se	Estimate	se	Estimate	se	Estimate	se
Intercept	551.70 ***	(3.39)	551.77 ***	(3.39)	535.20 ***	(3.26)	535.72 ***	(3.21)
Parental education					0.15 ***	(0.01)	0.16 ***	(0.01)
a	6.99 ***	(0.21)	6.99 ***	(0.21)	7.00 ***	(0.12)	9.21 ***	(0.45)
c	2.88 ***	(0.44)	2.70 ***	(0.50)	0.00 [a]	(0.00)	0.00 [a]	(0.00)
cl			1.24 *	(0.54)	1.25 *	(0.51)	3.43 ***	(0.92)
e	3.86 ***	(0.10)	3.79 ***	(0.12)	3.78 ***	(0.11)	4.88 ***	(0.40)
b_a							-0.03 ***	(0.01)
b_c							0.00 [a]	(0.00)
b_{cl}							-0.03 **	(0.01)
b_e							-0.02 **	(0.01)
V_A	48.81 ***	(2.88)	48.85 ***	(2.88)	48.93 ***	(1.63)	[b]	
V_C	8.29 **	(2.54)	7.29 **	(2.68)	0.00 [a]	(0.00)	[b]	
V_{CL}			1.54	(1.35)	1.57	(1.27)	[b]	
V_E	14.91 ***	(0.77)	14.33 ***	(0.91)	14.32 ***	(0.83)	[b]	
V_{Total}	72.02 ***	(1.37)	72.00 ***	(1.37)	64.82 ***	(1.24)	[b]	
Freely estimated parameters	30		31		39		44	
LL	-35692.26		-35691.18		-53444.70		-53380.95	
AIC	71444.52		71444.35		106967.39		106849.89	

Note: Estimates are unstandardized, controlled for age and sex in all models. *se* = standard error, V_A = genetic variance, V_C = shared environmental variance, V_{CL} = classroom variance, V_E = non-shared environmental variance, V_{total} = total variance, *LL* = loglikelihood, *AIC* = Akaike Information Criterion. The variances are based on the squared path estimates *a, c, cl, e* (e.g., $a^2 = 6.99^2 = 48.81$).

[a] Constrained to be zero, because it was fully explained by parental SES and not constraining it gives model convergence problems.

[b] Not applicable, because in the moderation model the size of a variance component depends on the level of parental education.

*p<.05, **p< .01, ***p< .001 (two-tailed test).

2.4.3 Selection into classrooms

The Dutch Association of Parents of Multiples (NVOM) advises that the decision to allocate twins to the same versus different classrooms when they start school should be made on an individual basis by teachers and parents together (NVOM, 2019). To what extent schools follow this advice or have a different policy is not entirely clear. International studies show that an official school policy to separate twins is exceptional. For example, a survey in the United Kingdom showed that only 1% of the schools had an official policy on the education of twins (Preedy, 1999). Precise information on the Dutch context is lacking, although the NVOM (2019) reports that 9% of the parents that participated in their survey indicate that only the school decides on the classroom allocation.

There may be, however, other selection effects. One of them relates to zygosity. An assumption underlying the classical twin design is that environmental influences are shared to the same extent by MZ and DZ twins. Under this assumption, greater similarity in educational performance among MZ twins compared to DZ twins can be attributed to MZ twins' greater genetic similarity. However, if MZ twins are more often in the same classroom than DZ twins, greater similarity in MZ twins' performance is due to both greater genetic similarity and greater classroom similarity. In our sample, MZ twins are somewhat more often in the same classroom than DZ twins (respectively 59.6% vs. 52.7%, $\chi^2 = 18.57$, $df = 1$, $p < .001$). If classroom sharing is unmodelled, this would lead to an overestimation of genetic influences (Grasby et al., 2020). However, since we explicitly model classroom sharing, a bias is avoided in our study. Greater similarity among MZ twins is no longer an unobserved mixture of greater genetic similarity and being more often in the same classroom because incorporating classroom information allows for separating these sources.

While selection based on zygosity is not problematic for estimating classroom influences, selection based on other characteristics may be. Twins (dis)similarity in certain characteristics may be a reason to allocate twins to the same class or different classes. This is only problematic if such characteristics also affect our dependent variable: educational performance at age 12. Studies suggest that there are mainly two characteristics that affect the decision to place twins in separate classrooms: ability and behavioral problems (Jones & De Gioia, 2010). As these characteristics may also affect educational performance, the estimated classroom component could be inflated. This is because greater twin similarity in educational performance for twins in the same class would then reflect similarity in prior ability or (the lack of) behavioral problems, rather than being solely a consequence of exposure to the same classroom context. We do not expect such selection processes to play a major role in Dutch primary schools. In Dutch primary education, there is no ability tracking. Neither parents nor schools can choose a high-performing classroom for one twin and a low-performing classroom for the other. Moreover, prior twin studies investigating classroom effects in countries where within-school tracking is common indicate that class allocation based on prior performance did not affect the conclusions on classroom effects (Byrne et al., 2010; Grasby et al., 2020). Concerning behavioral problems, studies using Dutch twin samples show that the association between problem behavior and classroom allocation, if any, is small and that it does not affect educational performance in the long run (i.e., performance at the age of 12) (Polderman et al., 2010; Van Leeuwen et al., 2005). Thus, the Dutch educational context and prior empirical evidence lead us to conclude that it is unlikely that originally dissimilar twins

end up in different classes. Even if it would occur, the effect on the (over)estimation of classroom effects will likely be minimal.

Lastly, if the decision to place twins in different classrooms is selective (because of differences in performance or behavioral problems), it would influence the interaction that we are interested in if the bias depends on parental SES. The literature suggests that low-SES parents tend to be more 'twinship oriented' (i.e., treating twins more similar), while high-SES parents tend to be more 'differentiation oriented' (i.e., treating twins more dissimilar, emphasizing individuality), which may be especially pronounced for MZ twins (Robin & Casati, 1994; Tourrette et al., 1989). If so, our reported negative interaction effect between parental SES and the influence of the classroom would be an underestimation, because the classroom effect would be overestimated for high-SES parents but not so much for low-SES parents. We find that twins of lower-SES parents are somewhat more often in the same class ($\chi^2 = 78.41$, $df = 9$, $p < .001$) (see Figure 2.5). We do not expect the underestimation of the interaction to be large, because the SES difference in class allocation is small and our results indicate that high-SES parents are not more differentiation-oriented than low-SES parents. Environmental influences that make twins more dissimilar, such as parents treating twins differently, are captured in the non-shared environment component (*E*). If high-SES parents would indeed be much more differentiation-oriented, one would expect *E* to be larger in high-SES families, which is not the case (see Figure 2.4a).

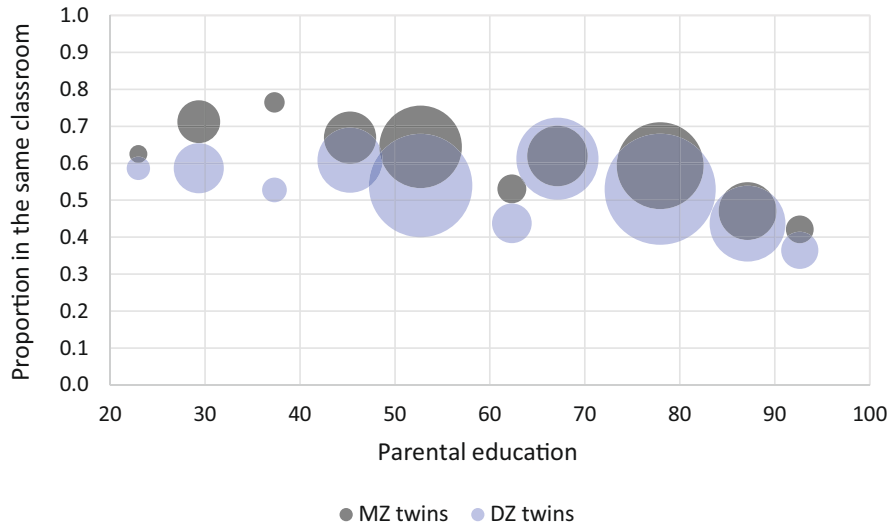


Figure 2.5 Proportion of MZ twins and DZ twins in the same class by parental education (ISLED).

Note: The size reflects the number of observations.

2.5 Conclusion and discussion

In many countries, educational inequality based on socioeconomic background is of great concern, leading researchers, policymakers, and educational professionals to question how to counter this. Classrooms are important contexts in children's lives, but it is unclear whether these are part of the problem by amplifying educational inequality or part of the solution by reducing educational inequality. We used a novel approach based on Dutch twin data and showed that the role of classrooms in explaining educational performance of primary school pupils is relatively minor when alternative sources (including genetic and non-shared environmental influences) are considered. However, classroom influences are stronger for children from lower-SES families. This suggests that classrooms compensate for influences of family background and thus serve as equalizers rather than amplifiers of educational inequality.

We find that on average 2.1% of the variance in educational performance of Dutch primary school pupils is attributable to classroom influences. Our classroom estimate falls within the range (i.e., 0-9%) that has been previously found in twin studies (Byrne et al., 2010; Grasby et al., 2020). There are different ways to interpret this result. On the one hand, finding an explained variance of 2.1% indicates that only a small part of *all* differences in educational performance can be accounted for by the classroom environment. Given the general view that teachers and other classroom aspects are important for children's educational outcomes, this may be somewhat surprising. It is important to be aware that our classroom effect shows the extent to which the classroom environment explains differences in performance, not the effect on the average performance. For example, a good teacher may increase average student performance, but if most classes have good teachers it would not explain much of the individual differences in performance. Yet, one would not conclude that good teachers are not important.

On the other hand, the classroom effect may also be interpreted as being more substantial if we relate it to the environmental part of the variance instead of the overall variance. The reason is that genetic variance is a major source of differences in educational performance. In our case, genetic differences account for 68% of the total variance, which implies that only 32% of the differences in educational performance result from environmental factors. Compared to these environmental differences, the share of the variance that is explained by classrooms is more sizeable. For example, we showed that of all the environmental influences that children from the same family share (e.g., parents, neighborhood, school, etc.) 12% can be attributed to the same classroom experiences. Moreover, classrooms are more important for some children than for others, as we discuss next.

Classroom effects are not the same for all children but are dependent on family background. The classroom influences are larger the lower-educated the parents are, explaining up to 7.7% of the total variance in educational performance for children from the lowest educated families. This suggests a compensatory effect and is in line with the argument that children from low-SES families could substitute school resources for family resources. Because these children are less likely to find supportive influences at home, for example, they may be more susceptible to a supportive school environment (Coleman et al., 1966; Jencks & Mayer, 1990; Rumberger & Palardy, 2005). However, finding a compensatory effect does not mean that amplification effects do not

take place at all. They could occur simultaneously: favorable classroom environments could have a compensating effect for low-SES children while at the same having an amplifying effect for high-SES children. If this would occur to the same extent, we would not find any differences in the classroom component by SES. Since we find larger classroom effects for low-SES children, we can conclude that compensatory effects of classrooms are stronger than possible amplification effects.

While we did not hypothesize on how other sources of variance depend on SES background, we find less genetic and non-shared environmental variance in higher SES families. Gene-environment interaction in educational outcomes is often studied from a bioecological framework, where genetic potential for high educational performance is thought to be actualized in more advantaged environments such as high-SES families (Bronfenbrenner & Ceci, 1994; Rowe et al., 1999; Scarr-Salapatek, 1971). However, we find that genetic variance is higher in low-SES families. This can be interpreted as genetic influences associated with lower educational performance having more detrimental effects on children from more disadvantaged socioeconomic backgrounds (see de Zeeuw et al., 2019). When we take into account that the total amount of variance in educational performance decreases by standardizing the results, we do not find a moderation of genetic variance (nor non-shared environmental variance) by parental SES.

Although the analysis of twin data allows us to contribute to prior studies by investigating classroom effects and their dependency on parental SES in a novel way, it does not solve all issues. In particular, the (moderating) effect of measured parental SES should not be interpreted as causal. Parents' genetic potential influences parents' SES, and they transmit part of this genetic potential to their children, which in turn influences the children's school performance. If these genes that parents and their children share are unmodelled, the association between parental SES and children's education is likely genetically confounded and thus does not only reflect environmental effects (e.g., effects of resources and behaviors shaped by parental SES) (Hart et al., 2021). This issue is not specific to our twin analyses but applies to most analyses involving parental SES and children's outcomes (except those with a causal design such as an instrumental variable or difference-in-difference approach). For our conclusions, this means that the observed compensation of the impact of family background by classrooms does not pertain necessarily only to environmental (dis)advantages passed on by families, but possibly also to the transmission of genetic (dis)advantages. This is not problematic if one is interested in the stratification of educational outcomes by family background in a broader sense, capturing all types of (dis) advantages that are associated with it. If one is interested in separating environmental and genetic intergenerational transmission, one would need other designs such as the children-of-twin design or measured genotype design.

The analysis of twin data also comes with some complexities of its own. A concern may be that twins form a special group and that conclusions based on twin data are not generalizable to the general population. Twins, especially identical ones, may be raised in more similar environments and may be treated more similarly than fraternal twins or non-twins. This could affect estimations if differential treatment is related to the outcome (i.e., educational performance). We do not expect this to be the case. Prior studies show only limited evidence for a violation of the equal environment assumption and if a violation occurred it did not affect the genetic and environmental

effects on educational outcomes (Conley et al., 2013; Mönkediek et al., 2020). Additionally, in a population-based study on educational performance in the Netherlands, twin-based estimates of *Cito* scores were not biased (Schwabe et al., 2017). These findings suggest that our results are generalizable to the general population of Dutch primary school pupils.

Nevertheless, relying on twin data may lead to conservative estimates of classroom effects. Twins almost always attend the same school, meaning that we capture only within-school differences between classrooms. Some classroom characteristics (e.g., SES composition, school resources) may cluster within schools, implying that the differences between classrooms in our study are smaller than the differences between classrooms in general. This is something we cannot capture with our design. The classroom effects that we find, therefore, probably reflect aspects that vary between classrooms within a school (e.g., teacher quality, climate) rather than aspects that vary mostly between schools (e.g., resources, student composition). In future research, measured classroom characteristics could be included in the twin model to investigate which specific characteristics explain the classroom variance. Another reason why our estimate of classroom effects is conservative is related to non-shared environmental influences. In our study, the non-shared environment after distinguishing the classroom component makes up around 20% of the total variance in educational performance. Part of this non-shared environmental variance may still be related to what happens in the classroom. Even if twins are in the same classroom, they may perceive their classroom environment differently or teachers could treat them differently, for example. Thus, sharing a classroom could also lead to differences between twins, which is captured in the non-shared environment component. Altogether, this means that our classroom estimate should be seen as a lower bound. Future work using complementary approaches could provide more insight into the interrelated influences of genes, families, and classrooms.

In conclusion, we find that classroom influences depend on family SES in Dutch primary schools. There is more classroom variance with lower levels of parental education, suggesting that children from lower-SES families benefit more from a high-quality classroom than children from high-SES families. However, we cannot conclude that classroom environments are great equalizers because the role of classrooms in explaining differences in educational performance is relatively small. Moreover, it can be expected that especially children from high-SES families are in high-quality classes given socioeconomic selection in schools (Borghans et al., 2015a). High-SES children are thus more likely to be in high-quality classes while the added value of such an environment is relatively little for them in terms of educational performance. Contrarily, low-SES children, for whom a high-quality class environment could make more of a difference, are likely to be less often exposed to this. Our results are thus indicative of a compensatory effect and show the modest potential of the classroom environment to reduce educational inequalities. Whether this potential is actually realized depends on the quality of the classroom environment that low-SES children are exposed to.

CHAPTER 3

Does school quality decrease educational inequality?

Evidence from gene-environment interaction analysis

A slightly different version of this chapter has been submitted to an international journal as:
Stienstra, K., Knigge, A., Maas, I. Does school quality decrease educational inequality? Evidence from gene-environment interaction analysis.

Kim Stienstra: Conceptualization, Methodology, Formal analysis, Writing – Original Draft, Visualization, Funding acquisition. **Antonie Knigge:** Conceptualization, Writing – Review & Editing, Funding acquisition. **Ineke Maas:** Conceptualization, Writing – Review & Editing, Funding acquisition.

3 Abstract

We study to what extent schools increase or decrease environmental and genetic influences on educational performance. Building on behavioral genetics literature on gene-environment interactions and sociological literature on the compensating and amplifying effects of schools on inequality, we investigate whether the role of genes and the shared environment is larger or smaller in higher-quality school environments. We apply twin models to Dutch administrative data on the educational performance of 18,384 same-sex and 11,050 opposite-sex twin pairs, enriched with data on the quality of primary schools. Our results show that school quality does not moderate genetic and shared-environmental influences on educational performance once the moderation by SES is considered. We find a gene-environment interplay for school SES: genetic variance decreases with increasing school SES. This school SES effect partly reflects parental SES influences. Yet, parental SES does not account for all the school SES moderation, suggesting that school-based processes play a role too.

This work was supported by the NWO research talent grant for the project 'Quality and inequality: The compensatory and multiplicative effects of school quality'(NWO: 406-18-557, awarded to Stienstra, Knigge, and Maas), an NWO Veni grant for the project 'Towards equal educational opportunities: The complex interaction between genes, families, and schools' (NWO: 451-17-030, awarded to Knigge), and the CBS ODISSEI Microdata Access Grant for the project 'Gene-environment interplay in education: The impact of school quality and tracking' (awarded to Knigge, Stienstra, and Maas).

We thank Esmee Bosma for her help with the data cleaning of the school quality indicators. Prior versions of this work were presented online at the ICS forum day (2020), the Migration and Social Stratification seminar (2020), the RC28 spring meeting (2021), Dag van de Sociologie (2021), and the Genes, Social Mobility, and the Inequalities across the Life-Course conference of the ifo Institute (2021). We thank all the participants for their valuable comments.

3.1 Introduction

Children who perform well in school are more likely to continue their education and obtain higher degrees, which in turn has numerous economic and social benefits including higher income, higher occupational status, and better health (UNESCO, 2018). Inequalities in educational performance thus translate into inequalities in other domains. It is therefore important to know why some pupils perform better than others and how inequality can be reduced. An often-studied source of differences in children's performance levels are social differences among them, most notably, the influence of their family socioeconomic status (SES) background (Sirin, 2005). Another important source of differences are genetic differences between children (De Zeeuw et al., 2015). The extent to which these family background influences and genetic influences play a role could be dependent on children's school environment. Higher-quality schools could strengthen the influence of families and genes and thereby increase performance differences between pupils. Conversely, higher school quality could also decrease differences in performance because the family background and/or genetic influences are reduced in these schools.

There are opposing arguments on whether family background and genetic influences are multiplied or compensated in higher-quality schools. Social science literature suggests that family background influences could be stronger in high-quality schools, for example, because high-SES children benefit more from good learning opportunities as they enter school with better academic preparation (Hanselman, 2018). Alternatively, the stable, stimulating, and resourceful learning environment in school could compensate for a less favorable home environment (Coleman et al., 1966; Rumberger & Palardy, 2005). In that case, especially low-SES children may benefit from higher-quality schools. Similarly, the behavioral genetics literature provides opposing models on whether advantageous environments such as those provided by high-quality schools increase or decrease genetic influences. From the bioecological model, stronger genetic influences in higher-quality schools can be expected because the more resourceful and stable environment in such schools could promote the realization of genetic potential for greater achievement (Bronfenbrenner, 1992; Bronfenbrenner & Ceci, 1994). However, weaker genetic influences can also be expected. Following the diathesis-stress model (e.g., Rende & Plomin, 1992) and the idea of compensation interaction (Shanahan & Hofer, 2005), the absence of stressors and the more supportive learning environment in higher-quality schools could compensate for the realization of genetic risks for lower performance.

To what extent schools increase or decrease environmental and genetic influences on educational performance is important to know for reducing educational inequality. Whether schools reduce educational inequality does not only depend on the multiplicative or compensatory effect of schools. It also depends on which sources of differences in performance are seen as part of educational inequality. Family background differences in educational performance are commonly problematized and labeled as inequality of educational opportunity, or social inequality in education (Strietholt, 2014; Van de Werfhorst, 2014; Van de Werfhorst & Mijs, 2010). Concerning genetic differences, there are different standpoints on whether genetic differences (or genetic inequality) in educational performance are problematic and should therefore be reduced. On the one hand, genetic differences in educational performance can be seen as an indicator of

opportunity for achievement (Guo & Stearns, 2002; Nielsen, 2006). If the realization of children's innate talent is not restricted by social barriers, differences in educational performance would be to a larger extent explained by genes (Nielsen, 2006). This (implicitly) assumes that differences due to good or bad luck in the genetic lottery are justified (Dias Pereira, 2021; Diewald et al., 2015). On the other hand, one cannot control their genetic endowment any more than their family background. Genetic differences could therefore, just as social differences, be interpreted as an unjust source of inequality (Diewald et al., 2015; Harden, 2021; Tannock, 2008). Thus, family background differences are generally seen as unfair and as a form of educational inequality, and this source of inequality can be dependent on the school environment. The school environment may also affect the role of genetic differences in performance, but there are different perspectives on whether genetic differences are seen as an unfair and problematic source of inequality as well.

We investigate how the school environment contributes to educational inequality and ask: 'To what extent does the school environment increase or decrease genetic and family background influences on educational performance?' We study this by using a twin design, which provides latent overall measures capturing genetic, shared (i.e., common, between-family) environmental, and non-shared (i.e., unique, within-family) environmental variance in educational performance (Knopik et al. 2016). We investigate how genetic and shared environmental variance varies across schools of different quality, also known as gene-environment interaction analysis. While gene-environment interactions are predominantly focused on the family environment, more recently, interactions with the school environment have been studied. These studies yielded mixed results concerning whether more advantageous school environments increase or decrease genetic and shared environmental differences (Hart, Soden, et al., 2013; Haughbrook et al., 2017; Taylor et al., 2010, 2020). Additionally, it is unclear whether these gene-environment interactions can indeed be attributed to the school environment or reflects processes in the family environment instead. Children from high-SES parents more often attend high-quality schools (Borghans et al., 2015a; Robert, 2010). School effects may therefore be confounded with family effects. Therefore, in this study, we do not only investigate whether the school environment moderates genetic and environmental influences but also to what extent the moderation by school quality is actually a moderation related to SES.

We study 29,434 twin pairs (birth cohorts 1994-2007), that we identified in administrative data from Statistics Netherlands (*CBS*). These data cover the whole population and therefore do not suffer from the self-selection bias that twin samples may do (Figlio et al., 2017; Schwabe et al., 2017). Additionally, the number of observations provided by administrative data lead to ample power to detect genetic and shared environmental influences and their interactions with school quality. The administrative data contain children's scores on a national standardized achievement test (*Cito* test) administered at the end of primary school around age 12. We enriched these data with many school quality indicators, as derived from the Dutch Inspectorate of Education.

The Netherlands provides an interesting context for investigating the role of school quality on educational inequality. The *Cito* test is a high-stakes test with major importance for children's educational careers. Together with a recommendation of the teacher, the test result determines which secondary school track children will attend. Once enrolled in a particular track, opportunities to switch to a higher track are limited (Naayer et al., 2016). Educational inequality in

this test score, whether it is related to social background differences and/or genetic differences, has thus large implications for future educational and career opportunities.

3.2 Theory

3.2.1 Influence of school quality on genetic differences

Although there is no clear definition of school quality, there seems to be a consensus that it entails at least two aspects: school resources and school climate. School resources refer to aspects that can (potentially) be bought either directly (e.g., educational materials) or more indirectly (e.g., educational time, teacher attention) (Grubb, 2009; Hofflinger & Von Hippel, 2020). School climate characteristics cannot readily be bought and are more difficult to change. These include norms, values, and expectations (e.g., academically oriented culture, high expectations), relationships (e.g., teacher-pupil relationships, cohesion), teaching and learning practices (e.g., structured instruction, differentiation), and larger organizational structures (e.g., educational leadership) (Cohen et al., 2009; Grubb, 2009).

High-quality schools can be expected to both increase and decrease genetic influences on educational performance. An increase can be expected from the bioecological model (Bronfenbrenner, 1992; Bronfenbrenner & Ceci, 1994), according to which genetic potential for developmental outcomes such as greater educational achievement is more actualized with increased levels of proximal processes (i.e., enduring forms of interaction in the immediate environment, e.g., parent-child interactions). This model has generally been applied to the role of family environment in explaining cognitive ability, which has become known as the Scarr-Rowe hypothesis (Rowe et al., 1999; Scarr-Salapatek, 1971). This hypothesis claims that in high-SES families, genetic potential is more fully expressed. Environments such as those provided by high-SES families can be seen as more advantaged and stable. They more often comprise different resources (e.g., material resources, cultural capital) and proximal processes that are more aligned with children's genetic potential and are therefore expected to enhance genetic expression (Baier & Lang, 2019; Bradley & Corwyn, 2002; Bronfenbrenner & Ceci, 1994). While the focus was originally on the realization of genetic potential for cognitive abilities, this has been extended to educational outcomes. For both outcomes, support for the Scarr-Rowe interaction has been mixed (Baier et al., 2022; Baier & Lang, 2019; De Zeeuw et al., 2019; Tucker-Drob & Bates, 2016).

The bioecological model and Scarr-Rowe hypothesis could be applied to the impact of schools on genetic influences on educational performance. Similar to high-SES families, high-quality schools are more stable and resourceful environments. In these schools, higher levels of positive proximal processes (e.g., teacher-child interactions) can be expected, which implies that teachers' behavioral patterns are more responsive to children's characteristics and actions. This is seen as the principal mechanism through which genetic potential for effective developmental functioning is actualized (Bronfenbrenner & Ceci, 1994). Available school resources (e.g., more experienced teachers, more teacher attention) and school climate (e.g., monitoring students' progress, differentiation) may make it easier to discover children's specific talents. Additionally, a high-quality school environment is characterized by aspects that could lead to children developing

their talents further, such as the availability of challenging materials and the presence of high-achievement norms.

From the diathesis-stress model, on the other hand, can be derived that with increasing school quality genetic influences decrease. According to this model, the realization of genetic risk for lower performance (e.g., learning or behavioral problems) is more likely when there are more environmental risks and stressors (Rende & Plomin, 1992; Shanahan & Hofer, 2005). The school environment in low-quality schools can be expected to have more environmental stressors, such as higher levels of classroom disorder and negative peer influences. Hence, the expression of genetic risks is more likely in such environments but decreases when such environmental stressors are less present as in high-quality schools. Moreover, high-quality schools have other positive features which may compensate for the realization of genetic risk (Shanahan & Hofer, 2005). For example, in high-quality schools, teachers may be more likely to notice specific risk factors for lower performance and these schools may also be better able to provide adequate support (e.g., remedial teaching).

3.2.2 Influence of school quality on family background differences

The influence of the family environment can also be expected to become either more or less important depending on the quality of the school. In the sociological literature, several arguments are provided for why the influence of family background might be stronger in higher-quality schools. One is that children from high-SES families have a cumulative advantage and benefit more from a high-quality school environment because they enter school better academically prepared (Hanselman, 2018). Based on the idea that 'skills beget skills', children's skills gained early in life increase children's capacity to benefit from later instruction in school (Heckman, 2000; Sørenson & Hallinan, 2016). For example, in high-SES families, children may develop more language skills because their parents tend to engage children more in conversations and use a richer vocabulary (Bradley & Corwyn, 2002). They may therefore understand instructional material better and reach higher performance levels in school. Another argument is that there is greater cultural correspondence between the home and school environment for high-SES children. The more ambitious and academically oriented culture in high-quality schools coincides with high-SES parents' expectations and ambitions. For low-SES students, such a culture means a mismatch between their family and classroom experiences which may lead to negative self-perceptions and emotional distress, negatively affecting educational outcomes (Crosnoe, 2009).

On the other hand, family background influences can also become less important in high-quality schools. According to the bioecological model, proximal processes do not only increase the realization of genetic potential, but also reduce, or buffer against, (shared) environmental differences in developmental outcomes (Bronfenbrenner & Ceci, 1994). Sociological literature provides more insight into how a higher-quality school environment may reduce family background differences in performance. Children from disadvantaged families tend to grow up in a more unstable environment outside school, receive less parental support, and have access to fewer parental resources. The environment in high-quality schools may be especially important for them (Coleman et al., 1966; Jencks & Mayer, 1990; Rumberger & Palardy, 2005). High-quality schools provide access to learning opportunities that overlap with those in socioeconomically

advantaged families. If learning opportunities in families and schools substitute for each other, school resources typically benefit students from a less resourceful family environment more than they benefit high-SES students (Chiu & Khoo, 2005; Hanselman, 2018). Also, school climate aspects (e.g., academic climate, good student-teacher relationships) improve student achievement, especially for children from more disadvantaged families (Gustafsson et al., 2018). These children could have a 'differential sensitivity' to such aspects in their school environment because they experience them less in their families (Coleman et al., 1966). Differences in educational performance attributable to family background may thus become less pronounced in high-quality schools.

3.2.3 Socioeconomic selection into schools

Children from high-SES parents more often attend high-quality schools (Borghans et al., 2015a; Robert, 2010). Consequently, higher-quality schools are not only characterized by their more advantageous resources and climate but also a high-SES composition. This composition may also affect educational performance. Moreover, the school's SES composition may affect the influences of genes and the environment on educational performance. Partly, school SES influences overlap with those of school quality because they are correlated. Schools' SES composition is associated with school characteristics such as teaching and instruction practices, and school organization and management processes (Reardon & Owens, 2014). For example, high-SES schools may more easily attract good and experienced teachers and have more rigorous curricula (Armor et al., 2018; Sykes & Kuyper, 2013). Since we explicitly measure school quality and rely on many indicators, we likely capture school-based influences related to school quality characteristics such as teaching and instruction practices, and school organization. However, school SES may also reflect other school-based mechanisms that are less captured by school quality, such as peer interactions (Armor et al., 2018). For instance, high-SES students with higher aspirations, better study habits, and less disruptive classroom behavior may have a positive influence on the performance of other students (Gutiérrez, 2023). Given such school-based mechanisms, some previous studies proposed that a large proportion of students from high-SES backgrounds is an indicator of school quality (Jonsson & Treuter, 2019; Van Hek et al., 2017). For this reason, it is worthwhile to study school SES in addition to school quality.

Another maybe even more important consequence of socioeconomic selection into school is that the school environment may capture family SES influences if these are not considered. High-SES parents tend to provide more stable and resourceful environments, just as high-quality schools do. Hence, parental SES may moderate genetic and shared environmental influences similar to school quality. Not considering this would overestimate the moderation by school quality.

Altogether, we explore whether school quality increases or decreases genetic and family background differences in educational performance. To provide more insight into the possible moderation of the school environment we additionally study if school SES increases or decreases genetic and family background differences in educational performance and explain part of the moderation effect of school quality. Lastly, we investigate if the moderation effect of schools is explained by parental SES.

3.3 Data and methods

3.3.1 Data

We use linked microdata from Statistics Netherlands (*CBS*) covering the whole population.[14] We construct a dataset including twin families with information on children's educational performance, school environment, and family SES. To construct twin families, we rely on basic demographic information on children and their legal parents, using linked parent-child data (CBS, 1995) combined with the municipal personal records database (CBS, 2021). We identify families based on children who share the same legal parents. Only knowing the legal parents and not the biological parents is not very problematic for our purposes. This is because genetic influences are derived from genetic relatedness between children (i.e., twins), not between parents and children.[15] After constructing families, we identify twin pairs. Since the birth day is not available because of privacy reasons, we base this on children who have the same birth month and year. Based on the sex composition, we identify same-sex (SS) and opposite-sex (OS) twin pairs. Multiple twin pairs in one family are analytically complex. Therefore, we select one random twin pair in these cases.

We use the *Cito* database (CBS, 2018) to obtain information on educational performance. These data are available for 2006-2019 (birth cohorts 1994-2007) at the time of this study. Primary schools can permit *Cito* to share the data with the *CBS*, who anonymized the data and assigned identification numbers to link the data at the individual and school level.

Data sources on parental SES that we use to construct school SES are the highest education database (CBS, 2019) for the year 2018 and personal income for the period 2005-2018 (CBS, 2011). Data on educational attainment are largely based on diverse registrations of individuals who completed their education at an educational institution funded by the government. There is no (reliable) register data available for privately funded education (which is relatively rare in the Netherlands), education abroad, and long-term corporate training. To add information on this, the *CBS* used data from the Labor Force Survey which is collected on a sampling basis. Income data is based on administrative information, mostly provided by the tax authorities.

We supplement children and parent data from *CBS* with official information on the school environment obtained from the Dutch Inspectorate of Education and the Dutch Education Executive Agency. Inspectorate of Education data include many indicators that are used to assess the quality of schools by the inspectorate and are generally available from 1999 onward. Education Executive Agency data include information on general school characteristics such as the number of students and teachers. These data are (mostly) available from 2011 onward. School data can be linked to the *CBS* data via a school identifier (*BRIN*).

14 All results are based on own calculations using non-public microdata from Statistics Netherlands. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information: microdata@cbs.nl.

15 Yet, there is still a small chance that the identified twins are not full biological siblings. There are a few cases where siblings share the same parent ID, but parental date of birth differs. This might indicate that siblings do not share the same parent. Correcting for this did not change the results.

3.3.2 Selections and selectivity

Figure 3.1 shows the sample selection. We only study twins from birth cohorts 1994-2007, due to data availability of our dependent variable. Only twin pairs for whom at least one twin has information on educational performance are included. One reason for missingness is that some schools did not permit to share the results with *CBS*. Another reason is that schools can choose to administer another test. Most schools use the *Cito* test. Until recently, around 80% of the schools administered this test and these schools did not differ from the total school population regarding region, school size, urbanization, and percentage of students from low-educated families (Van Boxtel et al., 2010). For the most recent years, the percentage of schools administering the *Cito* test decreased (63.8% in 2017/2018, 55.9% in 2018/2019) and became a bit more selective. Schools in more urbanized areas and larger schools more often administered the *Cito* test (Inspectie van het Onderwijs, 2019; Van Boxtel et al., 2010). Excluding these years did not substantially change our results.

We only include twins who went to the same primary schools. Most twins (in our data 98%) go to the same school. Those who attend different schools form a selective group (e.g., one twin attends a school for special needs). For 6,415 twin pairs, at least one of the twins had missing information on the school identifier. In most of these pairs (6,234 pairs), there was a co-twin with non-missing information. In these cases, we assume that both twins attended the same school. We exclude the twin pairs where both twins had missing information on school data. We also exclude twin pairs with missing information on parental SES. This leads to our analytical sample of N_{pairs} = 29,434.

3.3.3 Measurements

We measure our dependent variable, *educational performance*, by students' scores on the *Cito* test. The *Cito* test is a nationwide standardized educational achievement test taken at the end of primary education around age 12. It consists of multiple-choice items on Dutch language, mathematics, study skills, and world orientation (e.g., geography, biology, and history). The domains are combined into a standard score with a formula taking into account the difficulty of the test for that year, to make the scores comparable over the years. Because the subdomain world orientation is not mandatory, this is not included in the standard score. The score is on a scale from 501 to 550, with a national average of 535 and a standard deviation of 10 (Van Boxtel et al., 2010). The *Cito* score is available for multiple years if children repeated the final grade. In these cases, we use the most recent year.

For our moderator *school quality*, we use data from the Inspectorate of Education and Education Executive Agency to construct a factor score. The Inspectorate data consist of many official indicators that are used to assess the quality of schools. These are generally available for the period 2000-2011, and sometimes up to 2019. Data from the Education Executive Agency includes information on the financing of schools and the number of students and staff, among others. These data are available for the period 2011-2019, where each school has a measure for each year. However, the indicators derived from these data do not end up in our measurement model in the end (see Appendix C1). The inspectorate data are not collected for scientific research purposes, but to assess whether schools meet a certain quality standard. The inspectorate

usually visited schools once every four years, and the set of indicators that were used differed over the years. Although this provides a rich source of information, these data are not directly suited for research and require extensive data handling. The structure of the Inspectorate data with the resulting missing data makes it impossible to measure school quality per year or even a couple of years. Therefore, for each indicator, we take the average of all available years. Items are mostly measured on a three-point scale (insufficient, sufficient, good) or a four-point scale (bad, insufficient, sufficient, good). Sometimes, also a two-point scale is used (insufficient, sufficient; no, yes).

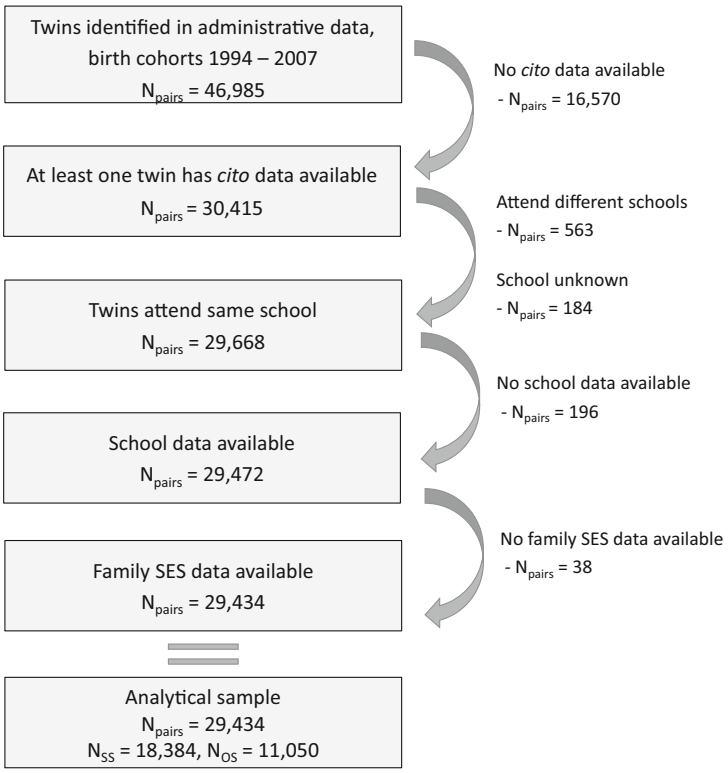


Figure 3.1 Selection of analytical sample.

Note. SS=same-sex twin pairs, OS=opposite-sex twin pairs.

We construct factor scores based on the standardized items using factor analyses with Full Information Maximum Likelihood (FIML) estimation in Mplus. We conduct two Exploratory Factor Analysis (EFA) analyses: one for all the items related to school resources leading to two dimensions, and one for school climate leading to seven dimensions. Altogether, this leads to

nine dimensions of school quality: (1) range of educational activities, (2) (implementation of) school curriculum, (3) guidance of educational needs, (4) parental involvement, (5) monitoring and evaluating (special needs) students, (6) learning climate, (7) social climate, (8) safety, (9) quality assurance. Based on these dimensions, we construct one overall school quality factor in a third (confirmatory) factor analysis. The dimensions of social climate, parental involvement, and safety have a low loading on this overall factor (Table C1.3, Appendix C1) and/or a low correlation with schools' average *Cito* score (Table C1.4, Appendix C1). As an alternative operationalization, we exclude these dimensions and construct a factor score based on the remaining six dimensions. This does not lead to substantially different results and therefore we keep all the dimensions in. Given the numerous latent variables and items, it is not possible to integrate the full measurement model with our analytical model. Therefore, we save the factor scores and include these in our analytical model as a single variable while imposing a measurement error correction. More information on this correction – as well as further details on the procedure, items, and factor analyses – are provided in Appendix C1.

Parental SES is measured by a factor score based on parental education and income. For parental education, we use the father's and mother's highest attained level of education, which is coded according to the International Standard Classification of Education (ISCED) 2011. We rely on the most recent data file from 2018. For income, we use the father's and mother's percentile scores of personal yearly income of the year before the *Cito* test. Personal income includes the gross income from labor, own company, income insurance benefits, and social security benefits (excluding child benefits and child-related budget). Premiums for income insurance have been deducted. For the percentile score, personal income is divided into 100 equal groups of people with income in private households. We construct a factor score for SES based on standardized items using CFA in Mplus (Appendix C2). FIML is used to handle missing data, which is especially present for parental education (Table C2.1 in Appendix C2).

School SES is an aggregate of parental SES. We use the average parental SES of all children in the school who took the *Cito* test in the year that the twins took this test.

We control for *year of birth* and *sex* (0 = female, 1 = male) in all models. Descriptive statistics of all variables are presented in Table 3.1.

Table 3.1 Descriptive statistics for same-sex (SS) and opposite-sex (OS) twins.

	SS twins			OS twins		
Variable	***N***	**Mean**	**S.D.**	***N***	**Mean**	**S.D.**
Twin specific						
Cito twin 1	16,530	535.17	9.77	9,909	535.08	9.82
Cito twin 2	16,456	535.32	9.68	9,909	535.13	9.73
Male twin 1	18,384	0.49	-	11,050	0.50	-
Male twin 2	18,384	0.49	-	11,050	0.50	-
Twin pair						
School quality	18,384	0.10	0.46	11,050	0.10	0.45
School SES	18,384	0.03	0.34	11,050	0.03	0.34
Parental SES	18,384	0.06	0.77	11,050	0.04	0.78
Year of birth	18,384	2000.12	3.73	11,050	2000.19	3.68

Note: All continuous independent variables are *z*-standardized prior to the analyses. Minimum and maximum values are not provided because of the confidentiality guidelines of Statistics Netherlands.

3.3.4 Twin model

In the classical twin design, structural equation modeling (SEM) is used to decompose the variance in a characteristic into three latent components. First, a component capturing additive genetic variance (*A*). Second, common or shared environmental variance (*C*), which includes all environmental aspects making twins more alike such as influences of family resources, parenting practices, educational expectations, and the broader environmental context that differs between families (Baier & Lang, 2019). We use the *C* component as a comprehensive measure of family background influences, reflecting social inequality in educational performance. Last, unique, non-shared, environmental variance (*E*), capturing aspects making twins dissimilar. These include, for example, subjective experiences, differential treatments, luck, and measurement error (Knopik et al., 2016). Latent components *A*, *C*, and *E* are set to a variance of 1. Path coefficients *a*, *c*, and *e* represent the effects of the latent factors on educational performance. The variance is equal to the square of the path coefficient; hence the *ACE* model can be written mathematically as:

$$V_{educ} = a^2 + c^2 + e^2 = V_A + V_C + V_E \quad (3.1)$$

where V_{educ} is the total variance of our phenotype educational performance. We extend this model by including a continuous moderator (*M*) (see Purcell 2002). In our case, our moderator school quality affects the average educational performance as shown by $\mu + b_m M$. It could also moderate, for example, *a* to become $a + b_a M$ (see Figure 3.2). The total variance in this moderation model changes to:

$$V_{educ|M} = \left(a + b_a M\right)^2 + \left(c + b_c M\right)^2 + \left(e + b_e M\right)^2 \quad (3.2)$$

Parameters in this model are usually identified because zygosity is available. Identical (i.e., monozygotic; MZ) and fraternal (i.e., dizygotic; DZ) twins differ in genetic resemblance: where MZ twins are genetically identical at conception, DZ twins share on average half of their segregating genes. Hence, the genetic correlation (*A1-A2*) can be constrained to 1 for MZ twins and 0.50 for DZ twins. It is assumed that MZ and DZ twins share their environment to the same extent, meaning that shared environmental correlation (*C1-C2*) can be constrained to 1 for both MZ and DZ twins. Accordingly, the MZ covariance is $Cov_{mz} = V_A + V_C$ and for DZ twins this is $Cov_{Dz} = 0.5V_A + V_C$.

Constraining the shared environmental variance to be equal reflects the Equal Environment Assumption (EEA).[16] Additional assumptions are no assortative mating, generalizability of twins to the general population, minimal gene-environment correlation, and absence of non-additive genetic effects. Violations of these assumptions could bias *A* and *C* but do so in different directions (see Table 1.1 in Chapter 1 of this dissertation)(Rijsdijk & Sham, 2002). The consequences of violations (upward or downward bias of *A* and *C*) are reflected in the difference in genetic relatedness between twin types. We perform the analyses with different genetic correlations to test to what extent our results are sensitive to the assumptions.

3.3.5 Twin model with unknown zygosity

We do not have information on zygosity but instead rely on data from 18,384 SS twin pairs and 11,050 OS twin pairs. OS twins are always DZ, hence their genetic correlation (*A1-A2*) is equal to 0.50. SS twins are a mixture of MZ and DZ twins. The true value for the average genetic correlation of SS twins (i.e., rSS_G) is unknown, and there are different ways in the literature to deal with this (Figlio et al. 2017; Pokropek and Sikora 2015; Rodgers, Rowe, and May 1994). A common way is using Weinberg's differential rule to estimate the proportion of MZ and DZ twins within SS twin pairs (Weinberg, 1901). According to this rule, the probability of male births equals the probability of female births, and therefore among DZ twins the number of SS twins equals the number of OS twins. The total number of DZ twins is thus twice the number of OS twins. The proportion of MZ twins within SS pairs in our data can be estimated by $p_{MZSS} = (N_{SS} - N_{OS}) / N_{SS} = (18{,}384 - 11{,}050) / 18{,}384 = .40$. For DZ twins within SS pairs this is $p_{DZSS} = 1 - .40 = .60$. This leads to an average genetic relatedness among SS twins of $rSS_G = (1*.40) + (.5*.60) = .70$. Another common approach, which is based on the assumption that among SS twins half will be MZ and half will be DZ, is to use $rSS_G = .75$ (i.e., the average genetic relatedness of MZ and DZ twins) (e.g., Rodgers et al. 1994).

While the MZ twin rate is relatively stable over time and MZ twinning is thought to be the result of a random event, this is not the case for DZ twins. DZ twin births are related to individual characteristics (DZ twin pregnancies are more common when the mother is older, taller, has a higher BMI, and smokes, among others) and the usage of assisted reproductive technology (ART) such as in vitro fertilization (IVF) (Glasner et al., 2013). There are also indications that usage of ART is related to MZ twin pregnancies, but the underlying causes are unknown (Glasner et al., 2013;

16 Violation of the EEA could overestimate A and underestimate C, but only if differential treatment is related to the outcome under study. Several studies showed that the EEA is unproblematic for a wide range of outcomes (for an overview, see, Felson 2014), including school performance specifically (Conley et al., 2013; Mönkediek, 2021).

Vitthala et al., 2009). In the Netherlands, the average maternal age and use of ART increased over the past decades, although the IVF policy has become more conservative (increasingly only one embryo is being transferred) (Glasner et al., 2013). Assuming a fifty-fifty mixture of MZ and DZ twins among the SS pairs is likely not realistic for the population that we study. Relying on the estimated genetic relatedness using Weinberg's differential rule overcomes this problem. Indeed, when we calculate the MZ/DZ ratio among SS twins we find a larger share of DZ twins among SS twins (i.e., a ratio of 40/60, leading to the estimated genetic relatedness of rSS_G = .70).

Relying on a twin model with unknown zygosity has been criticized (e.g., Eaves & Jinks, 1972). One concern is that the design is less powerful than using information on zygosity. This is less applicable to our study given the use of population data. Another concern is that the method relies on the assumption that the correlation of SS twin pairs only differs from that of OS twin pairs because SS twins are on average genetically more similar, not for other non-genetic reasons (Figlio et al., 2017). This assumption is violated if SS DZ twins are more similar to one another than OS DZ twins because the first are from the same sex and the latter are not. One could test this by comparing intraclass correlation coefficients (ICCs) of SS and OS pairs with known zygosity. This has been done for reading and mathematics achievement using data from the Netherlands Twin Register (De Zeeuw & Boomsma, 2017). Comparing the similarity in educational achievement for SS and OS pairs and MZ and DZ pairs with different sex compositions demonstrated that the assumption holds (De Zeeuw & Boomsma, 2017). Another way to test this is by comparing the ICCs of SS and OS non-twin sibling pairs. This shows that SS siblings are slightly more similar (average ICC for males and females = .44) than OS siblings (ICC = .42), suggesting small sex influences (Appendix C3).[17] Although the difference is only .02, it could still lead to a non-negligible upward bias in estimates of genetic variance and a downward bias in shared environmental variance.[18] One can correct this bias by using a larger value for rSS_G.[19] We, therefore, perform our analyses using three values of rSS_G (.70, .75, .80). As we will show, our conclusions are robust to the different values.

17 Additionally, comparing ICCs of OS twins and OS siblings shows that OS twins are more similar than OS siblings, suggesting that twin status influences pair similarity. The difference is relatively small, especially considering that OS twins have the same age while OS siblings differ in age.

18 To give an intuition for the possible size of the bias, if one uses the ICC of OS DZ twins (.45), the descriptive estimate of heritability would be .80, which can be calculated with the formula $(ICC_{SS} - ICC_{OS}) / (rSS_G - rOS_G) = (.61 - .45) / (.70 - .50)$. If we assume that the sex-effect for twins is the same as for non-twin siblings, the ICC for SS DZ twins would be .45 + .02 = .47. Based on this ICC, heritability would be .70.

19 Theoretically, one would adjust the shared environmental correlation of OS twins downwards to consider that their environments are less similar because they are of different sexes. However, genetic and shared environmental relatedness account for the same pattern in the data (cf. Spinath et al. 2004), so it makes sense to only adjust one at a time. Practically, increasing rSS_G is similar to decreasing the shared environmental correlation of OS twins.

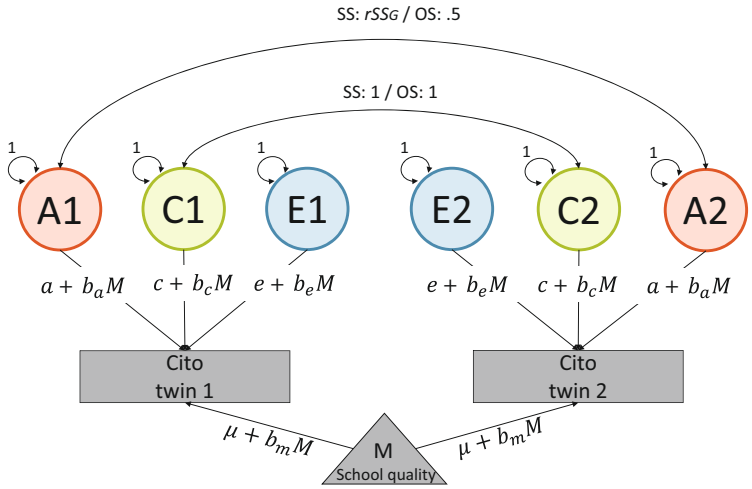


Figure 3.2 *ACE* moderation model.

Note. Latent variables represent genetic (*A*), shared-environmental (*C*), and non-shared environmental (*E*) components of educational achievement, with corresponding path coefficients *a*, *c*, and *e*. Measured variable *M* refers to the moderator. Genetic covariance for same-sex (SS) twin pairs is estimated by $rSS_G = 1\frac{N_{SS} - N_{OS}}{N_{SS}} + .5\frac{N_{OS}}{N_{SS}} = .70$ and is .5 for opposite-sex (OS) twin pairs. We also use .75 and .80 as alternative values for rSS_G.

3.3.6 Analytical strategy

We fit a series of *ACE* models in Mplus. We have data on 29,434 twin pairs nested in 5,843 schools. To account for this nested structure, we adjust the standard errors for clustering at the school level. In all models, the influences of sex and birth year are controlled for by including them as covariates. We *z*-standardize all continuous independent variables prior to the analyses. Before fitting the twin models, we test for equal means and variances between SS and OS twins. The difference in the mean of educational performance (Wald test = 0.83, *df* = 1, *p* = .362) and variance in educational performance (Wald test = 0.12, *df* = 1, *p* = .733) are not statistically significant, indicating that equality of means and variances can be assumed.

We first examine the *ACE* model and include the main effects of school quality, school SES, and parental SES on educational performance in a stepwise fashion. It should be noted that these school and family measures are always shared between twins and thus can only explain shared environmental variance in the *ACE* model, even though these measures include genetic and non-shared environmental variability (Turkheimer et al., 2005). Hence, their associations with educational performance should not be interpreted as causal, as they can be genetically confounded (see, e.g., Hart et al., 2021). Next, we allow the *ACE* components to be moderated by school quality and school SES to test whether genetic and shared environmental variance in educational performance increases or decreases with increasing school quality and school SES. Subsequently, we test the moderation by school quality and school SES simultaneously, to see whether school SES explains part of the moderation effect of school quality. Lastly, we control

for the moderation by parental SES to further scrutinize whether moderation by the school environment measures reflects school-based processes or are instead driven by what happens in the family environment. We perform several robustness checks to assess to what extent our results are dependent on our model assumptions and operationalization of school quality.

The few behavioral genetics studies examining whether the school environment moderates genetic and shared environmental variance generally looked at absolute variance components (Hart et al. 2013; Taylor et al. 2020, 2010), although standardized components were also used (Haughbrook et al. 2017). For standardized components, each variance component is made proportional to the total variance. For example, relative genetic contribution (i.e., heritability) is obtained by $sV_A = \frac{V_A}{V_{educ}}$. An advantage of standardized components is that it considers that the total variance may differ between contexts while the effect of genes and the environment do not differ. For example, in certain schools children may be more genetically similar or more similar concerning their (non-)shared environmental background than in other schools (Knigge et al., 2022). An advantage of using unstandardized, absolute variance components is that genetic and shared environmental variances can be contingent on school quality independent of each other. Solely focusing on standardized components will conceal underlying processes. As both standardized and unstandardized variance components have pros and cons, we report both.

3.4 Results

3.4.1 Genetic and environmental influences on educational performance

Before examining the *ACE* moderation model, we first investigate unmoderated genetic and environmental influences on educational performance by decomposing the variance in educational performance in genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variance. We do so for different values of rSS_G. The total variance in educational performance is $V_{educ} = 95.34$. While the total amount of variance does not depend on rSS_G, the variance decomposition differs. In our lower bound scenario ($rSS_G = .70$, Model 1, Table 3.2), genetic differences explain 90.9% of the variance in educational performance ($SV_A = 86.65 / 95.34 = .909$) and we find no shared environmental variance. When we use $rSS_G = .75$ (Model 2, Table 3.2), we find that the variance in educational performance is to a lesser extent attributable to genetic differences (73.0%) and more to shared environmental variance (8.9%). When we further increase rSS_G (Model 3, Table 3.2), genetic variance becomes smaller and (non-)shared environmental variance larger. Altogether, 61-91% of the variance in educational performance can be attributed to genetic variance, 0-15% to shared environmental variance, and 9-24% to non-shared environmental variance.

Table 3.2 *ACE* model for cito for different values of rSS_G ($N_{SSpairs}$=18,384, $N_{OSpairs}$=11,050).

	Model 1 rSS_G = .70		Model 2 rSS_G = .75		Model 3 rSS_G = .80	
Parameter	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**
Intercept	534.63 ***	(0.08)	534.63 ***	(0.08)	534.63 ***	(0.08)
a	9.31 ***	(0.06)	8.34 ***	(0.24)	7.62 ***	(0.22)
c	0.00	(0.04)	2.91 ***	(0.47)	3.78 ***	(0.32)
e	2.95 ***	(0.12)	4.16 ***	(0.16)	4.80 ***	(0.11)
V_A	86.65 ***	(1.14)	69.61 ***	(3.96)	58.00 ***	(3.30)
V_C	0.00	(0.00)	8.46 **	(2.74)	14.26 ***	(2.42)
V_E	8.69 ***	(0.71)	17.28 ***	(1.32)	23.08 ***	(1.02)
Freely estimated parameters	6		6		6	
Loglikelihood	-190307.99		-190307.98		-190307.98	
Scaling correction factor	1.08		1.30		1.30	
AIC	380627.98		380627.97		380627.97	

Note: ***p<.001 (two-tailed test). Controlled for sex and year of birth. Robust standard errors accounting for clustering at the school level are shown in parentheses. Parameters *a*, *c*, and *e* refer to the unmoderated path coefficients capturing genetic, shared-environmental, and non-shared environmental influences, respectively. Squaring these path coefficients gives the variance components V_A, V_C, and V_E.

Next, we sequentially include the main effects of school quality, school SES, and parental SES on educational performance (Table C4.1, Appendix C4). The estimated sizes of these associations do not depend on rSS_G. School quality is statistically significantly associated with educational performance (b = .61, β = .06, p < .001). This is not substantial; each standard deviation (S.D.) increase in school quality is associated with a 0.61 point (0.06 S.D.) increase in *Cito* score. Additionally, sequentially including school SES and parental SES shows that parental SES has a much stronger association with educational performance (b = 2.89, β = .30, p < .001) and explains part of the association between educational performance and school quality and school SES. The school quality association with performance reduces to .24 (β = .02, p < .001). Controlled for parental SES (and school quality), shows that there is a relatively weak positive association between school SES and educational performance (b = 0.92, β = .09, p < .001). If there is shared environmental variance present, as is the case for rSS_G = .75 and rSS_G = .80, this is (almost) entirely accounted for by school quality, school SES, and parental SES.

3.4.2 Genetic and environmental influences moderated by school quality and school SES

Next, we test how school quality moderates genetic and environmental variance while using a genetic correlation of rSS_G = .70 (Model 1, Table 3.3). We could have used .75 and .80 as well, which we do in the next section as robustness checks. We find that with increasing school quality, genetic influence decreases statistically significantly, b_aSQ = -.019, p = .009 (Figure 3.3a and Table

3.3, Model 1). Shared environmental variance is absent once school quality is taken into account. Because of the decreasing genetic variance, the total variance in educational performance also decreases with increasing school quality. When this is taken into account by standardizing the *ACE* components, we see that the relative genetic influences barely decrease with increasing school quality (Figure 3.3b).

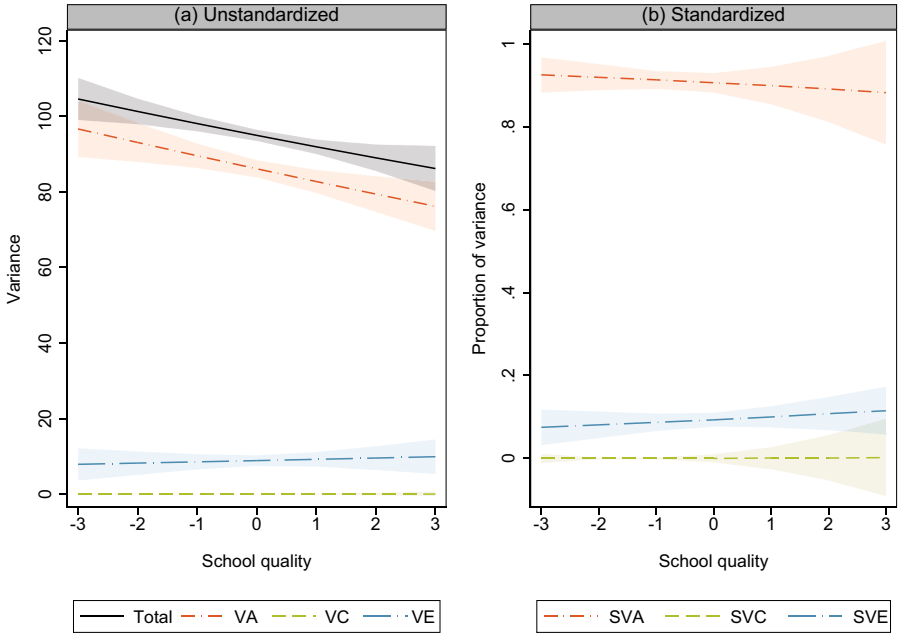


Figure 3.3 Unstandardized and standardized genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variances of educational performance moderated by school quality, including 95% CI.

Note: Based on Model 1 of Table 3.3 using a genetic correlation of rSS_G=.70.

Similar to our findings for school quality, our results show that school SES decreases genetic influence and does not affect shared environmental influence on educational performance (Figure 3.4 and Table 3.3, Model 2). Although we did not have expectations on the moderation of non-shared environmental variance, we find that school SES statistically significantly decreases non-shared environmental variance (Model 2, Table 3.3). When the decreasing total variance with increasing school SES is taken into account, we do not find any moderations in the relative contribution of genetic and environmental variances (Figure 3.4).

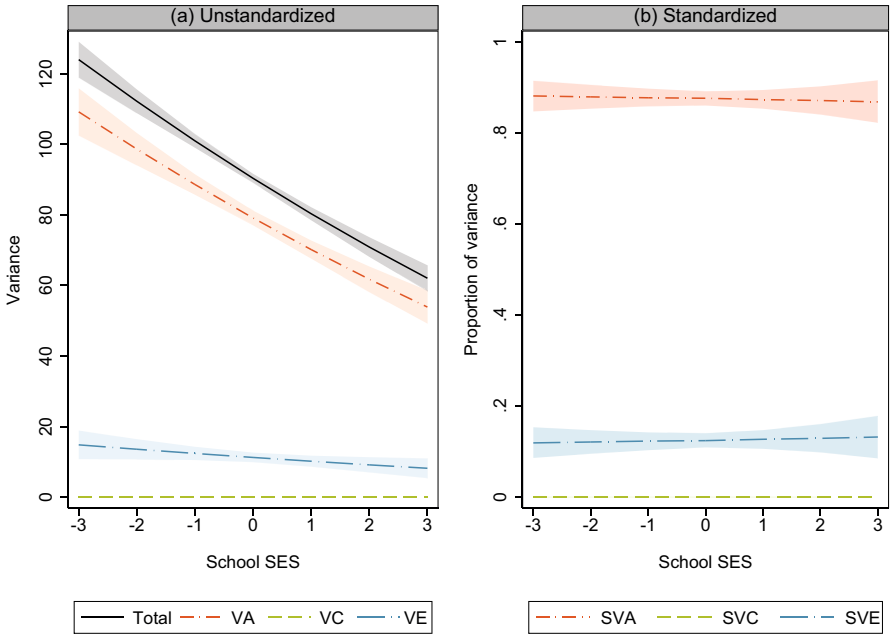


Figure 3.4 Unstandardized and standardized genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variances of educational performance moderated by school SES, including 95% CI.

Note: Based on Model 2 of Table 3.3 using a genetic correlation of rSS_G=.70.

Table 3.3 *ACE* model for cito with interaction effects of school quality, school SES, and parental SES ($N_{SSpairs}$=18,384, $N_{OSpairs}$=11,050).

	Model 1		Model 2		Model 3		Model 4	
Parameter	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**
Intercept	534.63 ***	(0.08)	534.65 ***	(0.07)	534.65 ***	(0.07)	534.67 ***	(0.06)
a	9.28 ***	(0.08)	8.90 ***	(0.06)	8.88 ***	(0.07)	8.22 ***	(0.07)
b_a*SQ*	-0.19 **	(0.07)			-0.05	(0.06)	0.00	(0.06)
b_a*SchoolSES*			-0.52 ***	(0.05)			-0.31 ***	(0.06)
b_a*ParentalSES*					-0.51 ***	(0.05)	-0.62 ***	(0.06)
c	0.16	(1.31)	0.00	(0.00)	0.13	(0.25)	0.26	(0.19)
b_c*SQ*	0.10	(1.05)			0.39	(0.50)	-0.32	(0.24)
b_c*SchoolSES*			0.00	(0.00)			-0.84 ***	(0.23)
b_c*ParentalSES*					-0.12	(0.16)	0.71 ***	(0.19)
e	2.98 ***	(0.13)	3.35 ***	(0.10)	3.37 ***	(0.11)	3.91 ***	(0.09)
b_e*SQ*	0.06	(0.13)			0.01	(0.10)	-0.04	(0.09)
b_e*SchoolSES*			-0.17 *	(0.08)			-0.05	(0.09)
b_e*ParentalSES*					-0.17 *	(0.08)	-0.22 **	(0.08)
School quality	0.61	(0.07)			0.24 ***	(0.06)	0.22 ***	(0.06)
School SES			2.23	(0.06)	2.20 ***	(0.06)	0.91 ***	(0.06)
Parental SES							2.94 ***	(0.05)
Freely estimated parameters	13		10		17		21	
Loglikelihood	-217429.91		-174629.60		-216384.47		-214779.29	
Scaling correction factor	2.23		0.98		1.90		1.77	
AIC	434885.82		349279.20		432802.93		429600.58	

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). A genetic correlation of rSS_G=.70 is used. Controlled for sex and year of birth. All continuous independent variables are z-standardized prior to the analyses. Robust standard errors accounting for clustering at the school level are shown in parentheses. Parameters *a*, *c*, and *e* refer to unmoderated path coefficients capturing genetic, shared-environmental, and non-shared environmental influences, respectively. The *b* coefficients refer to the moderation effects of *a*, *c*, and *e*, by school quality (SQ), school SES, and parental SES.

3.4.3 Simultaneous test of the moderating role of school quality and SES

The average SES of children in a school may explain part of the moderation effect of school quality. When we test the moderation by school quality and school SES simultaneously in Model 3 (Table 3.3), the model fits the data better than the model that only includes the moderation by school quality. When the moderation effects of school SES are taken into account, the genetic moderation by school quality is indeed reduced and no longer statistically significant. School SES could capture both school effects (e.g., peer group processes) and family effects. Therefore, we additionally include parental SES as a moderator in Model 4, which fits the data better than Model 3. This final model shows that the moderating role of school SES is for a part attributable to parental SES. The previously found moderation of genetic variance by school quality is thus partly attributable to selection of high-SES children in high-SES schools. When this is considered, we find no evidence for moderation effects of school quality anymore (Figure 3.5).

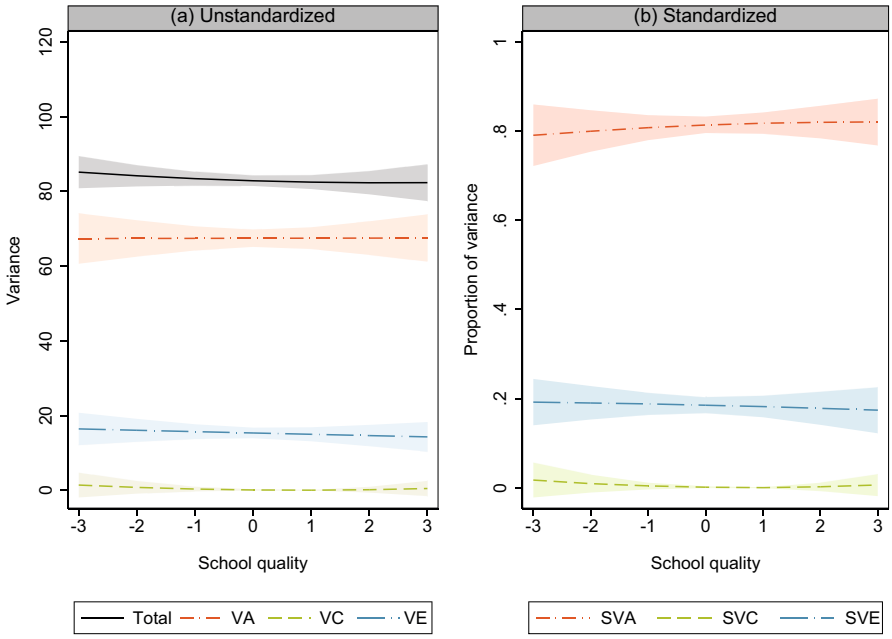


Figure 3.5 Unstandardized and standardized genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variances of educational performance moderated by school quality controlled for the moderation effect of school SES and parental SES, including 95% CI.

Note: Based on Model 4 of Table 3.3 using a genetic correlation of rSS_G=.70.

The final model shows that the moderating role of school SES is for a part attributable to parental SES, but not entirely. School-based processes likely play a role too, as the genetic moderation effect by school SES reduces, but more than half the size of the genetic moderation

coefficient remains and is statistically significant (see Table 3.3, Model 4). We also find a statistically significant moderation of shared environmental variance by school SES once parental SES is controlled for. However, we find the evidence for decreasing shared environmental variance rather weak given the small amount of shared environment variance that is present to begin with. Lastly, the decreasing non-shared environmental variance by school SES that we found in Model 3 (Table 3.3), appears to be attributable to parental SES (see Model 4, Table 3.3).

3.4.4 Robustness checks

First, we performed auxiliary analyses to test whether our findings are robust against using different values of genetic relatedness of SS twins. Conclusions based on our estimated genetic relatedness of rSS_G = .70 still hold when the alternative values .75 and .80 are used (Figure 3.6; Appendix C4). We still find no moderation of genetic and shared environmental variance by school quality. In our main analyses, we found a significant moderation of genetic variance by school quality when we did not control for school SES and parental SES. In our robustness check, this moderation effect by school quality is not statistically significant. As can be seen in Figure 3.6 (and more detailed in Appendix C4), the moderation effects, as well as the variance components in general, are estimated with less precision. Concerning the moderation by school SES, we still find a decrease of genetic variance with increasing school SES. When controlling for parental SES this negative moderation remains statistically significant and substantial when rSS_G = .75 is used (as was the case for our main results using rSS_G = .70), but not if rSS_G = .80 is used. Lastly, concerning our expectation that SES confounds the gene-school quality moderation effect, we found empirical support for this when using rSS_G = .70. For rSS_G = .75 and .80, we do not find statistically significant moderation effects of school quality in the first place, meaning there is no moderation effect that can be confounded by school SES and parental SES (Tables C4.2 and C4.3, Appendix C4).

Second, we investigated whether our gene-environment interaction is driven by SES differences in the estimated MZ/DZ ratio among SS twins. We relied on the average estimated genetic relatedness of SS twins of rSS_G = .70 (and, alternatively, values of .75 and .80). However, given the factors that affect DZ twin pregnancies (e.g., IVF usage, maternal age, BMI, smoking), rSS_G may differ between SES groups. If there are relatively more DZ twins in higher-SES families than in lower-SES families, rSS_G will be lower in high-SES families than the assumed average of .70. Our observed gene-SES interaction could then be the result of an underestimation of genetic variance among twins from high-SES families (and an overestimation of genetic variance among lower-SES families). Our gene-SES interaction does not appear to be driven by SES differences in the MZ/DZ ratio. If anything, the estimated proportion of DZ twins among SS twins is larger for low-SES than high-SES families. Our gene-SES interaction may therefore even be slightly underestimated.

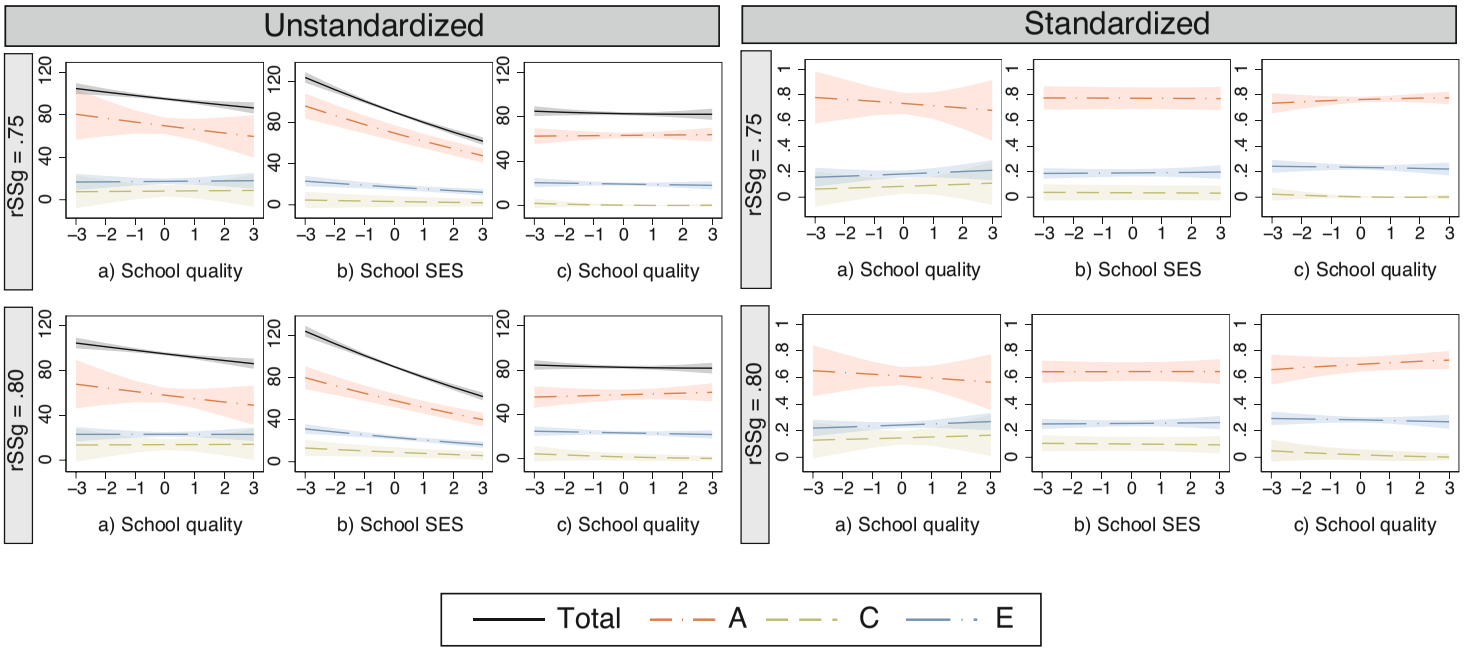


Figure 3.6 Unstandardized and standardized genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variances of educational performance moderated by a) school quality, b) school SES, and c) school quality controlled for the moderation effect of school SES and parental SES, including 95% CI.

Note: Based on a model using a genetic correlation of rSS_G=.75 (top row) and rSS_G=.80 (bottom row).

Third, we investigated non-parametric gene-environment interactions for school quality and school SES. The *ACE* moderation model including a continuous moderator assumes linear moderating effects on the *ACE* components, while there may be threshold effects (Purcell 2002). For example, it could be that only the most disadvantaged schools show increased genetic differences. Therefore, we performed a multigroup (i.e., non-parametric) gene-environment interaction for quantiles of school quality and school SES (see Appendix C5). The results largely mirror the main analyses with continuous linear moderations. The only difference is that for school quality (not controlled for school SES and parental SES) the genetic variance is not declining linearly with increasing school quality. There is more genetic variance in the lowest quality schools and less in the other four quantiles. Only the difference in genetic variance between the first and fourth quantile is statistically significant (Wald test: $\chi^2 = 5.78$, $df = 1$, $p = .016$). For school SES, there is a clearer linear decline in genetic variance although not all group differences are statistically significant (see also the large confidence intervals in Figure C5.3, Appendix C5).

Lastly, we used different operationalizations of our school quality variable. The influence of certain more specific school quality aspects might be masked by using one overall school quality measure. We looked more specifically into the moderating role of school quality by separating it into the school resources and school climate dimensions. We reach the same conclusion if we use these dimensions instead of one overall school quality factor. School resources and school climate are both positively associated with average educational performance, with a similar strength as the overall school quality factor (Figure C6.1, Appendix C6). Similar to the main results, we find that school resources and climate negatively moderate genetic variance, but not when school SES and parental SES is controlled for (Table C6.1, Appendix C6). We also investigated all nine underlying school quality dimensions separately. When school SES is controlled for, only three dimensions remain statistically significantly associated with average performance (see Figure C6.2, Appendix C6). These are guidance of educational needs, monitoring and evaluating (special needs) students, and learning climate. For these three dimensions, we performed moderation analyses. None of the dimensions moderates genetic variance. However, once we control for both school SES and parental SES, the remaining shared environmental variance turned out to be moderated by learning climate (see Model 2 of Table C6.2, Appendix C6). We do not interpret this effect, because it is the only statistically significant school quality moderation we found out of many tests, and we did not correct p-values for multiple testing. Moreover, the effect is not substantial.

3.5 Conclusions and discussion

Inequality of educational opportunity is seen as a problematic phenomenon in many societies, making researchers, policymakers, and educational practitioners question how to reduce it. Especially high-quality schools may have the potential to reduce educational inequality. We investigated this using gene-environment interaction analyses applied to administrative data on twins. Smaller shared environmental variance in higher-quality schools is indicative of less inequality of opportunity in these schools. Family background would then be less decisive for educational performance. Whether smaller genetic variance is also indicative of less inequality

depends on one's perspective what educational inequality entails. If genetic variance is seen as unfair because children have no control over the genes they inherit, smaller genetic variance would indicate less inequality. Alternatively, genetic variance can be considered to capture innate talent, and performance differences due to differences in talent could be seen as fair. In that case, smaller genetic variance gives an indication of more inequality of opportunity.

We do not find evidence that school quality decreases educational inequality, neither concerning social inequality nor genetic inequality. Initially, it seems to be the case that genetic variance is smaller in higher-quality schools. However, the lower genetic variance in these schools appears not to be related to school quality but to school SES and parental SES instead. If genetic differences in performance are seen as unfair and part of educational inequality, there is less inequality in high-SES schools. We would have misattributed the decreasing genetic variance in educational performance to higher-quality school environments instead of higher-SES family and school environments (and overestimated the influence of school SES) if school quality, school SES, and parental SES were not studied simultaneously.

These findings thus suggest that not school quality but instead SES plays a role. The results suggest that it is both parental SES and school SES that matter, and that the underlying mechanisms thus reflects processes in both the family and school context. The smaller genetic influence in higher-SES environments is consistent with the diathesis-stress model (Rende & Plomin, 1992). This model suggests that the fewer environmental risks and the more positive factors in higher-SES families and schools neutralizes or compensates for the expression of genetic risks towards poor educational performance (see also Shanahan and Hofer 2005). Low-SES environments are generally considered less favorable for educational performance and may thus enforce such influences of genetic risk. This implies that if children have genetic risks (e.g., related to learning or behavioral problems), this will have fewer negative consequences for their educational performance if they have high-SES parents and attend high-SES schools. For example, because high-SES parents are more likely to provide adequate support. This is consistent with the sociological compensatory advantage mechanism, according to which prior negative outcomes (e.g., health and cognitive endowments at birth, previous school results) are compensated by high-SES parents (Bernardi, 2014). Such negative prior outcomes are genetically influenced; hence, the compensatory advantage mechanism could be expanded to include the compensation of disadvantageous genetic dispositions. Similarly, a larger share of high-SES children in school may contribute to a more advantageous environment. For example, high-SES pupils may influence the performance, aspirations, and student habits of their peers and contribute to an environment that is more conducive to learning (Gutiérrez, 2023; Lazear, 2002). This may be especially beneficial for pupils with more genetic risks for lower performance.

The decreasing genetic variance is accompanied by a decrease in the total variance in performance with increasing SES (i.e., there is less dispersion). It could be that pupils differ less in their performance levels in higher-SES environments because genetic influences are compensated via the potential mechanisms that we just discussed. However, there could also be less variance in performance in high-SES environments because these environments are more homogenous in terms of children's genetic makeup and/or environmental characteristics. In that case, focusing on the standardized results would be more appropriate. The standardized results do not show a

gene-environment interaction. Hence, an alternative explanation for the lower genetic variance in high-SES schools is that selection into schools plays a role instead of a substantive interplay between genes and the school environment.

Although we had no expectations of the moderation of non-shared environmental variance in performance, we find that it decreases with increasing SES. One interpretation of this finding may be that similar to the compensation of genetic risks, also non-shared environmental risks may be compensated in high-SES environments. Non-shared environmental risks include child-specific influences that negatively affect educational performance (e.g., accidents, illness, negative peer influences). If one twin has the risk of lower educational performance due to such individual circumstances, high-SES parents may be more likely to compensate for this (Bernardi, 2014; Conley, 2008). Contrarily, low-SES parents may not have the opportunity to compensate (e.g., due to their lower levels of economic and cultural resources), hence, twins may end up performing more differently. This would then be reflected in the larger non-shared environmental differences with lower SES. Since the non-shared environment also includes measurement error, an alternative interpretation is that there is less measurement error in the educational performance of high-SES twins. More research is needed for conclusions on the potential differential impact of the non-shared environment.

We did not find evidence for genetic and environmental influences being dependent on school quality, and only small effects of school quality on average educational performance. This could mean that the school environment is not as important for (inequality in) educational performance as thought. It may also be that there is less (systematic) variation in school quality in the Netherlands than in other countries because of how the educational context is organized. Private schools are rare and both public schools and religious schools receive public financing, proportional to the number of pupils. Schools attended by pupils from more disadvantaged backgrounds receive additional funding (Ritzen et al., 1997). This could result in fewer quality differences between schools than in other countries where school funding is more unequal. In other contexts, the effects of school quality on average performance may be stronger and the gene-school quality interaction might work differently. More comparative research could provide insight into this. An alternative explanation is that school quality matters, but that we do not sufficiently capture it with our indicators. For example, maybe it is not so much between-school quality differences but within-school differences that play a role (e.g., teacher quality, classroom processes). When differences in educational performance between twins that are in the same versus different classrooms are investigated, the classroom environment indeed turned out to play a role in the Netherlands (Stienstra et al., 2023 / Chapter 2).

In addition to the operationalization of school quality, also the operationalization of educational performance deserves some further reflection. Our measure, pupils' *Cito* score, has the advantage that it is based on a test that is taken nationwide and is standardized. Moreover, it is a meaningful measure in the Dutch context given its importance for pupils' future educational career. However, if one thinks of this measure as capturing the full underlying educational performance distribution of pupils, the *Cito* test has some disadvantages. It has certain properties that results in a negatively skewed distribution with some censoring at the top (and to a lesser extent also the bottom) of the distribution (Van Boxtel et al., 2010). The negatively skewed

distribution results from the chosen difficulty level of the test, expressed by the p-value which indicates the proportion of pupils who answered the item correctly. For the *Cito* test, it is aimed to have questions of a difficulty between .40 and .90 and an average of around .70 (Van Boxtel et al., 2010). The censoring results from the scaling of the items. To make the scores comparable over the years, the number of correct answers on the test are transformed into a scale with a mean of 535, a standard deviation of 10, and a range from 501 to 550. If the transformed scores are below 501 or above 550, they are rounded to the minimum or maximum (Van Boxtel et al., 2010). More pupils score at the upper end of the scale. Also, this occurs more among children from higher-SES backgrounds; they more often obtain the maximum score of 550 (De Zeeuw et al., 2019). On the one hand, it could be argued that this does not matter that much. In the end, it is children's position on the *Cito* score scale that matters for their future educational career. Hence, it is worthwhile to study this reality, including the properties of the scale. On the other hand, when one is interested in the underlying sources of differences of the (latent) educational performance of pupils, it is important to study the full range of differences between pupils instead of a somewhat censored scale. In this study, it could be that the reducing (genetic) variance with increasing SES could be a ceiling effect. However, I expect that the decreasing variance with increasing SES reflects (at least partly) substantive compensation mechanisms. The study by De Zeeuw et al. (2019) investigated the interaction between family SES and the underlying sources of variance in *Cito* scores and corrected for the effect of censoring at the high end of the scale, which led to the same results.

Another potential limitation relates to the use of SS and OS twins. Although we use high-quality administrative data, a limitation is the absence of information on zygosity. This could also lead to biased estimates for genetic and shared environmental variance. Since the true genetic relatedness among SS twins is unknown, we had to rely on estimated genetic relatedness. To check how sensitive our results are to the model assumptions, we used different values for genetic relatedness among SS twins. We think our approach led to valid conclusions. Previous studies on educational performance for similar cohorts in the Netherlands, but based on a non-random twin sample with zygosity, found estimates for genetic, shared environmental, and non-shared environmental variance within the range of our estimates (i.e., 61-81% genetic, 0-15% shared environmental, and 9-24% non-shared environmental variance) (De Zeeuw et al., 2016; Knigge et al., 2022; Stienstra et al., 2023). Moreover, our conclusions relating to the gene-environment interactions remain the same irrespective of which value of genetic relatedness among SS twins is chosen.

The observed decrease in genetic differences and non-shared environmental differences in more advantageous environments is not entirely surprising, as a prior twin study on educational performance in the Netherlands found less unstandardized genetic variance (and less environmental variance) in educational performance with increasing family SES (De Zeeuw et al., 2019).[20] We show that this also holds for school SES. Also studies using polygenic indices (PGIs) provide evidence in line with this compensation pattern. PGIs are composite measures for each individual based on the correlation between genetic variants and an outcome, and

20 In Chapter 2, we replicate this finding using another operationalization of family SES.

therefore provide an estimate of an individual's genetic liability to this outcome (Dudbridge, 2013). Recent studies used this for investigating the interaction between schools and the influence of educational attainment PGI on educational attainment, college completion, and dropping out of math, amongst others, and likewise found evidence for compensation (Cheesman et al., 2022; Harden et al., 2020; Trejo et al., 2018).

These findings can be used as a starting point for future research to investigate the mechanisms underlying the negative gene-SES interaction. The (decreasing) genetic differences in higher-SES environments as provided by the twin model, but also the educational attainment PGI, can be seen as a black box. They do not provide enough information on whether the genetic influence on educational performance occurs via characteristics that have a positive effect (e.g., cognitive ability) or negative effect (e.g., deviant behavior, psychiatric disorders). Future studies could identify mediators of the gene-SES interaction by investigating whether the gene-SES interaction in educational performance can be explained by gene-SES interactions in (non-) cognitive characteristics (see Ruks, 2022). Empirically distinguishing between positive genetic potential and negative genetic risk would provide a more informative way to investigate whether the environment enhances genetic potential (i.e., bio-ecological model) or compensates for genetic vulnerability (i.e., diathesis-stress model, compensatory advantage). If more advantageous environments compensate for genetic risks, it can be expected that this would be especially pronounced for (the genetic component of) more specific learning problems such as dyslexia and ADHD than (the genetic component of) general educational performance or cognitive ability, for example. This could be investigated by including such specific characteristics that have a negative effect on educational performance in the twin design and/or using the PGIs of these characteristics.

Both the twin design and usage of PGIs have advantages and disadvantages (see, e.g., Mills & Tropf, 2020). Therefore, providing a definite conclusion on the interplay between the school environment and genetic and shared environmental influences requires combining different methods. For now, based on our twin analyses, we conclude that school quality does not decrease (and neither increase) educational inequality. We find evidence for a gene-environment interplay in educational performance in the Netherlands, where genetic differences are smaller in more advantageous environments consistent with the idea of compensation of genetic risks. This gene-environment interplay in education turns out to be an SES composition effect rather than a school quality effect. Therefore, the results of this study suggest that reducing quality differences between schools would likely not be sufficient to reduce educational inequality.

CHAPTER 4

The nature-nurture of educational performance at the intersection between gender, family background, and school context

A slightly different version of this chapter has been published as:
Stienstra, K. & Karlson, K. B. (2023). The nature-nurture of academic achievement at the intersection between gender, family background, and school context, *Social Science Research*, 111(102870), 1-14. doi: 10.1016/j.ssresearch.2023.102870.

Kim Stienstra: Conceptualization, Methodology, Formal analysis, Writing – Original Draft, Visualization, Funding acquisition. **Kristian B. Karlson:** Conceptualization, Formal analysis, Writing – Review & Editing.

4 Abstract

We investigate the role of gender, family SES, school SES, and their intersection in educational performance using a twin design. Drawing on theories of gene-environment interaction, we test whether high-SES environments compensate for genetic risks or enhance genetic potential, and its dependency on gender. Using data on 37,000 Danish twin and sibling pairs from population-wide administrative registers, we report three main findings. First, for family SES, but not for school SES, we find that genetic influences play a slightly smaller role in high-SES environments. Second, this relationship is moderated by gender: in high-SES families, the genetic influence is considerably lower for boys than for girls. Third, the moderating effect of family SES for boys is almost entirely driven by children attending low-SES schools. Our findings thus point to significant heterogeneity in gene-environment interactions, highlighting the importance of considering the multiplicity of social contexts.

This work was supported by the NWO research talent grant for the project 'Quality and inequality: The compensatory and multiplicative effects of school quality'(NWO: 406-18-557, awarded to Stienstra, Knigge, and Maas) and the ECSR Visitor Grant (awarded to Stienstra).

Prior versions of this work were presented at the Welfare, Inequality, and Mobility seminar (University of Copenhagen, 2021), Migration and Social Stratification seminar (Utrecht University, 2022), European Social Science Genomics Network conference (Bologna, 2022), European Consortium for Sociological Research (ECSR) conference (University of Amsterdam, 2022), and the ECSR Network Workshop 'Compensatory advantage and other inequality generating mechanisms' (European University Institute, 2022). We thank the participants for their valuable comments.

4.1 Introduction

Educational performance has consequences for economic, social, and health outcomes. Therefore, ample research has focused on the question of why some students perform better than others. Children differ from one another in many ways, including their gender, the family they are raised in, and the school they attend. An additional source of individual differences that received less attention in the sociological literature is genetic influences. Twin studies show that differences in children's educational performance are for a large part explained by genetic differences between them (e.g., de Zeeuw, de Geus, & Boomsma, 2015). Genetic influences do not operate in a vacuum but are influenced by the environment, including historical contexts, institutions, and local communities (Elder & Shanahan, 2006; Herd et al., 2019). The extent to which genes play a role in educational performance can thus be expected to depend on key socializing contexts in children's lives, including their family and school environment.

There are competing expectations on how genes and environmental contexts interact in shaping children's educational performance. On the one hand, genetic influences may be larger in more advantageous environments. According to the bioecological model, genetic potential is enhanced in more stable and resourceful environments, such as provided by high-SES families (Bronfenbrenner, 1992; Rowe et al., 1999; Scarr-Salapatek, 1971). On the other hand, genetic influences may be smaller in such advantageous environments. This pattern follows from the diathesis-stress model that states that the realization of genetic risks is less likely to be encouraged if levels of environmental risks and stressors are lower (Rende & Plomin, 1992), but also from the sociological idea that an advantageous environment may compensate for negative endowments (Bernardi, 2014; Shanahan & Hofer, 2005).

Twin studies provide a way to investigate gene-environment interactions in children's educational outcomes. Previous research has examined interactions in educational outcomes with family SES (Baier et al., 2022; Baier & Lang, 2019; Figlio et al., 2017; Tucker-Drob & Bates, 2016) and, to a lesser degree, the school environment (Hart, Soden, et al., 2013; Haughbrook et al., 2017; Taylor et al., 2010, 2020). Yet, so far family and school environments have been studied as independent contexts. Moreover, gender differences in these gene-environment interactions have not been examined. Research suggests that boys' academic outcomes are more strongly affected by environmental factors than girls', as the gender gap appears to be larger in disadvantaged families and schools (e.g., Autor et al., 2019; Legewie and DiPrete, 2012). For example, boys may be more sensitive to school contexts, whereas the educational performance of girls is less responsive to the extent to which the school environment is learning-oriented, disruptive, or disorganized (Legewie & DiPrete, 2012; Van Hek et al., 2017). Therefore, if the family and school environment moderate genetic and environmental influences, this moderation could be more pronounced among boys than girls. Providing insight into this is of interest to both researchers and policymakers in light of the reversal of the gender gap in education (Buchmann & DiPrete, 2006).

In this paper, we examine gene-environment interactions across family and school environments and investigate whether these interactions differ by gender. Thus, we effectively study the genetic influence on educational performance at the intersection between gender,

family background, and school context. To do so, we analyze high-quality data on about 37,000 Danish twin and sibling pairs from population-wide administrative registers. These registers comprise information on children's compulsory school grade-point-average (GPA), their parents' socioeconomic status, the schools they attend, and child's gender. One significant advantage of our register data is that they provide us with many cases, yielding sufficient statistical power to detect the different interactions, something that is not possible in relatively small twin samples. However, our data do not comprise information on zygosity. Therefore, we rely on the comparison of same-sex (SS) twins and siblings. We follow previous studies and estimate the genetic relatedness of SS twins (e.g., Figlio et al., 2017; Rodgers et al., 1994). By comparing SS twins with an assumed average genetic relatedness of 75%, with SS siblings who share 50% of their genetic makeup, we decompose the variance in GPA into genetic variance (*A*), shared environmental variance (*C*), and non-shared environmental variance (*E*). We subsequently break down these components by groups defined by family SES, school SES, and gender.

By considering the multiplicity of contexts, we advance existing research in this area. Prior research examines families and schools as separate environments, thus potentially neglecting that individuals are simultaneously embedded in both family and school environments.[21] Adding gender as an extra layer to the interaction allows us to examine whether these interactions play out differently for boys and girls.

4.2 Theory

4.2.1 Gene-environment interactions in educational performance

Today scholars realize that genetic and environmental influences are not additive but interdependent explanations for educational performance (Conley & Fletcher, 2017; Freese, 2008). This interdependency is apparent in gene-environment interactions, that is, when environmental characteristics moderate genetic influences (Shanahan & Hofer, 2005). Theories on how genes and the environment interact are dominated by two competing interaction patterns (Asbury et al., 2005). On the one hand, there could be a *multiplicative* interaction between genes and the environment. An advantaged social context would then enhance the actualization of genetic endowment for education (Shanahan & Hofer, 2005). Such an interaction pattern can be derived from the bioecological model (Bronfenbrenner, 1992; Bronfenbrenner & Ceci, 1994), which poses that enriched environments exacerbate the influence of genetic potential. Enriched environments have higher levels of positive proximal processes (i.e., enduring forms of interaction characterized by increasing complexity) that promote the actualization of genetic potential. Simultaneously, these environments are also thought to buffer against environmental differences in developmental outcomes. Under this model, more advantageous environments will therefore increase genetic variance and decrease shared environmental variance.

21 Chapter 3 of this dissertation forms an exception, as this study examines the family and school environment simultaneously.

On the other hand, there could be a *compensatory* interaction between genes and the environment. It has been argued that a social context could compensate for genetic vulnerabilities in educational performance (Shanahan & Hofer, 2005). The diathesis-stress model states that the realization of a diathesis (i.e., a genetic vulnerability) is more likely when the level of environmental risks and stressors is higher (Paris, 1999; Rende & Plomin, 1992). The absence of these stressors and/or the presence of positive features in the environment can neutralize or compensate for the realization of genetic vulnerability. This model implies an interaction pattern in opposite direction from the bioecological model, namely less genetic variance in advantageous environments. Although the diathesis-stress model has mostly been applied to psychopathological outcomes such as depression and ADHD, it can also be considered for "positive" developmental outcomes such as cognitive ability and educational performance. For example, if there is a favorable literacy environment, a child with genetic risks for poor reading may reach more similar reading levels as a child without genetic risks (Pennington et al., 2009).

4.2.2 Family background and school SES

Gene-SES interaction: multiplication

The family SES environment is often considered as an enhancing social context in studies investigating gene-environment interactions in cognitive and educational outcomes (Hart, Soden, et al., 2013). The bioecological model predicts that genetic potential for educational outcomes is more fully realized in high-SES families, while in low-SES families the importance of shared environmental influences is larger (Bronfenbrenner & Ceci, 1994; Scarr-Salapatek, 1971).[22] In high-SES families, the home environment is not only more resourced and stable but also thought to be more adapted to children's genetic potential. High-SES parents' parenting style is more active and focused on planned interactions and cognitively and emotionally stimulating activities with their children (Lareau, 2003). This implies that high-SES parents actively foster children's specific talents (Baier & Lang, 2019).

The type of school in terms of SES composition that children attend can also be expected to play a role in the gene-environment interplay. School SES could have an effect beyond the family, as it relates to peer group influences, the quantity and quality of resources and teachers, and organizational and management processes within the school, amongst others (Caldas & Bankston, 1997; Sykes & Kuyper, 2013). Low-SES schools are characterized by more instability and disorganization than high-SES schools. For example, students in low-SES schools may experience higher levels of classroom disorder (Van Hek et al., 2017) and teacher shortage and turnover (Ingersoll, 2001). If we follow the bioecological model, such unstable and less resourceful environments reduce the extent to which genetic potential is realized because in such environments there are lower levels of proximal processes (Bronfenbrenner & Ceci, 1994). High-SES schools, on the other hand, may provide a more stable and resourceful environment with teacher-child interactions that are more responsive to children's potential. A high SES

22 The gene-family SES interaction in cognitive ability has also become known as the Scarr-Rowe hypothesis. This hypothesis can be seen as a specific case of the more general gene-environment interaction that follows from the bioecological model.

composition is also positively associated with school characteristics that contribute to better learning outcomes such as teacher quality, school resources, and academic culture (Rumberger & Palardy, 2005; Sykes & Kuyper, 2013). Therefore, it can be expected that in high-SES schools children's specific talents are more easily discovered (e.g., via more experienced teachers) and that these schools are better equipped to further develop children's talents (via, e.g., the availability of challenging materials or high-achievement norms) (see also Stienstra et al., Chapter 3). This pattern may be less prevalent in low-SES schools, and thus genetic potential may more often be left unrealized in these schools.

Gene-SES interaction: compensation

In contrast to the multiplication of genetic influences, a compensation interaction implies that genetic influences are *lower* in high-SES contexts. Under the diathesis-stress model, genetic differences associated with negative outcomes are more likely to be realized when the level of environmental risks and stressors is high (Rende & Plomin, 1992). Such risks and stressors more often appear in low-SES families, which more often are single-parent families, have financial problems, experience stress, and are characterized by home disorganization and unpredictability (Bradley & Corwyn, 2002; Dumas et al., 2005). Many of the characteristics that pertain to low-SES families also pertain to low-SES schools in that they have more environmental risk factors. As mentioned earlier, the school environment in these schools is more often unstable and disorganized than high-SES schools. In high-SES schools, genetic variance in educational performance can be expected to be lower, simply because risk factors and stressors appear less often. If there are fewer risk factors and stressors in the environment, genetic risks are less often realized.

Moreover, in addition to fewer risk factors, high-SES environments are characterized by more protective factors. The protective factors in high-SES families and schools may neutralize or compensate for low genetic endowment (or negative genetic risk) for educational performance. For example, in supportive educational environments at home or in school, children with reading disabilities receive adequate support, and the negative consequences of genetic susceptibilities for reading disabilities can be avoided (Friend et al., 2008). This idea is consistent with the notion of compensatory advantage, a situation in which prior negative outcomes (e.g., health and cognitive endowments at birth, poor school performance) entail fewer negative consequences for children from higher-SES backgrounds (Bernardi, 2014). To the extent that the negative outcomes are genetically influenced, this notion could be expanded to include the compensation of genetic risks.

Although the bioecological model is the dominant framework in studies investigating gene-environment interplay in education, the empirical evidence for this is mixed. A meta-analysis of gene-environment interactions of family SES in educational performance and cognitive ability found support for the bioecological model in the United States (Tucker-Drob & Bates, 2016). Outside the United States, often no interaction or a compensation interaction pattern has been found (Asbury et al., 2005; Baier et al., 2022; De Zeeuw et al., 2019; De Zeeuw & Boomsma, 2017; Figlio et al., 2017; Ruks, 2022). Compared to the family environment, the school environment is much less studied. Also here the results are inconclusive. Genetic influences on reading performance and general educational performance were both found to be larger (Haughbrook

et al., 2017; Taylor et al., 2010, 2020) and smaller (Hart et al., 2013; Stienstra et al., Chapter 3) in more advantageous school environments, as measured by (proxies for) school quality, teacher quality, and school SES.

The interplay between family SES and school SES

The roles of family SES and school SES are not necessarily independent and could interact. To our knowledge, such interaction has not yet been considered in gene-environment studies of educational performance. However, sociologists pay substantial attention to the interplay between family and school environments. Family SES and school SES could – just as genes and SES – compensate or multiply for each other's influence. In case of *compensation*, family and school influences could act as substitutes in predicting educational outcomes. Higher levels of parental resources could substitute for poorer learning opportunities at low-SES schools by providing, e.g., supplemental educational investments such as private tutoring (Hanselman, 2018). Similarly, low-SES children may benefit more from high-SES schools because these schools provide environmental inputs (e.g., resources, academic climate, higher levels of motivation and aspiration) that these children are less likely to find at home (Coleman et al., 1966; Kahlenberg, 2001). In case of *multiplication*, children from high-SES families would have a cumulative advantage and reap greater rewards from the learning opportunities in high-SES schools. High-SES children enter school better academically prepared and may therefore benefit more from school opportunities (Hanselman, 2018). Moreover, the cultural overlap between family and school environments may play a role. From a cultural reproduction perspective, high-SES students are more positively evaluated by teachers, and they experience a greater sense of belonging at school, all leading to improved educational outcomes (Bourdieu & Passeron, 1977; De Graaf et al., 2000).

Applied to gene-environment interactions in education, a compensation effect of families and schools would imply that the gene-family SES interaction is stronger in low-SES schools than in high-SES schools. If families and schools multiply each other's influence in the enhancement of genetic potential or compensation of genetic risk, then the gene-family SES interaction would be stronger in high-SES schools (see Table 4.1).

4.2.3 Gender differences

Boys tend to perform better in math, whereas girls outperform boys in most other educational outcomes (Downey & Vogt Yuan, 2005). Boys and girls also differ in the dispersion of educational performance: boys show greater variance in educational performance irrespective of area (reading, mathematics, and science) and educational level (primary or secondary school) (Baye & Monseur, 2016; Gray et al., 2019). We expect that the moderating role of family and school SES on genetic and environmental influences differs between boys and girls. Studies suggest that environmental factors have a larger impact on boys' academic outcomes than girls, as the gender gap appears to be larger in disadvantaged families and schools (Autor et al., 2019; Legewie & DiPrete, 2012). Thus, insofar as the environment moderates genetic and environmental influences, this moderation would be more pronounced among boys than girls (see Table 4.1).

Two interrelated mechanisms may account for SES disparities in the gender gap: investments in boys relative to girls could depend on SES (i.e., differential investment) and boys could be more

affected by investments and circumstances associated with SES (i.e., differential sensitivity) (Autor et al., 2019). First, concerning differential investments, high-SES parents are argued to invest more in boys, and low-SES parents more in girls (Freese & Powell, 1999). When gender roles are more traditional, fathers are expected to spend less time with their children in general. Additionally, mothers might be more focused on girls and fathers more on boys. Given that low-SES families tend to have more traditional gender-role attitude and a higher level of single mothers, parents may spend less time monitoring and interacting with boys than girls in low-SES families (e.g., Bertrand & Pan, 2013; Buchmann & DiPrete, 2006). In contrast, high-SES parents may provide more compensatory investments in boys than girls (Autor et al., 2019). Parents may invest more resources to improve boys' education because boys are expected to have a lower likelihood of earning a college degree (Quadlin, 2019). Especially high-SES parents may compensate since they have a larger pool of resources (Conley, 2008). Still, empirical evidence for differential investments by child gender and its dependency on family SES is mixed (Autor et al., 2019; Buchmann et al., 2008). The notion of differential investment also applies to the school context. It could be that aspects of higher-quality school environments, such as teacher involvement and attention, are more beneficial for boys than girls simply because they are more exposed to such teacher interactions (Opdenakker, 2021). Teachers seek to give equal treatment to girls and boys and believe they do so (Younger et al., 1999). Yet, they may still (unconsciously) favor one gender over the other, although the direction is not entirely clear (Buchmann et al., 2008). For example, some studies suggest that boys receive more questions and feedback from the teacher, while others found that teachers' attitudes, expectations, and the broader learning environments favor girls (Meece, Glienke, & Burg, 2006).

Second, even if inputs from the family or school environment are qualitatively and quantitatively similar for boys and girls, they may be more influential for boys' educational performance given boys' greater sensitivity to their environment.[23] One reason for this greater sensitivity is related to boys' lower levels of intrinsic motivation (Opdenakker, 2021; Vantieghem & Van Houtte, 2018). If boys are more often externally motivated, they may also be more triggered by environmental factors such as the encouragement of the teacher (Opdenakker, 2021). Another reason is provided by the gender-role socialization perspective, according to which mothers are more important for the development of girls' educational aspirations and attainments, and fathers more for boys (Buchmann et al., 2008). Therefore, in single-mother families or in families with little father contact – which is more prevalent in low-SES families – boys may be more strongly impacted by the absence of a male role model and lower levels of paternal time and resources (Buchmann & DiPrete, 2006; Lei & Lundberg, 2020).

The third reason for boys' greater sensitivity is that boys may be particularly sensitive to how gendered identities are constructed in schools. Masculinity tends to be associated with disruptive behavior and opposition to school authority, not with educational effort and performance, which

23 Disentangling whether the underlying causes of boys' greater environmental sensitivity are genetic or environmental in origin is difficult. Environmental sensitivity is a characteristic that can be genetically influenced (Pluess & Belsky, 2013), but to our knowledge there is no research investigating gender differences in this relationship. Therefore, we focus on the (largely environmental) explanations that are available in the literature. Still, one should be aware that the actual mechanisms are probably more complex than what is suggested here (including genetic and environmental influences, and their interplay).

more often is labeled as feminine and thereby stigmatized (Legewie & DiPrete, 2012; Van de Gaer et al., 2006). For this reason, boys may be more likely to develop an anti-school subculture, whereas for girls femininity is not associated with disengagement from school. These gendered patterns are reinforced by peer pressure to conform to (gender-stereotypical) norms of the peer group, which is experienced stronger for boys than girls (Warrington et al., 2000). These processes may be especially prevalent in schools lacking academically oriented learning climates such as schools with lower school quality and less motivated and lower performing students (e.g., low-SES schools). In such schools, there may be a stronger oppositional culture in male than female peer groups. Disengagement from school could increase boys' peer group status in such schools (Legewie & DiPrete, 2012). Conversely, a school context that is more academically oriented promotes academic competition as an aspect of masculinity and thereby provides a different way of 'doing gender' (Legewie & DiPrete, 2012). Therefore, boys are argued to benefit more from aspects that contribute to this learning-oriented culture, such as high-quality teachers, and a larger proportion of high-SES students and girls in the classroom (Legewie & DiPrete, 2012; Van Hek et al., 2017; Van Houtte, 2004).

Table 4.1. Overview of the expected gene-environment interaction patterns.

Moderator	Multiplication		Compensation	
Family SES	(+)	genes × family SES	(-)	genes × family SES
School SES	(+)	genes × school SES	(-)	genes × school SES
Family SES × School SES				
Compensation				
Low-SES school	(+ +)	genes × family SES	(- -)	genes × family SES
High-SES school	(+)	genes × family SES	(-)	genes × family SES
Multiplication				
Low-SES school	(+)	genes × family SES	(-)	genes × family SES
High-SES school	(+ +)	genes × family SES	(- -)	genes × family SES
Gender	(+ +)	all of the above stronger for boys	(+ +)	all of the above stronger for boys

Note. (+) indicates a positive gene-environment interaction where A is higher and C is lower with increased levels of SES, (-) indicates a negative gene-environment interaction where A is lower and C is higher with increased levels of SES. A double sign (+ + or - -) indicates that the interaction is expected to be stronger.

4.3 Data and methods

4.3.1 Data

We analyze data from the Danish administrative registers, which are annually updated databases comprising a range of information on all members of the population. The register data are of exceptional quality, containing highly reliable variables and only very few missing observations

(Jensen & Rasmussen, 2011). We mainly rely on the education registers of Statistics Denmark, which are generated from the administrative records from the education institutions. We restrict our analyses to all children in the Danish registers who completed compulsory school between 2003 and 2014 (with the vast majority being born between 1986 and 1998). We choose these cohorts as 2003 is the first cohort for which GPA is observed in the administrative registers. Given our twin study design, we only consider (i) SS twin pairs and SS sibling pairs, (ii) siblings who are spaced by no more than three years, and (iii) twins and siblings that attend the same school type.[24] Our final sample comprises 5,010 twin pairs and 32,283 sibling pairs.

For the cohorts we analyze, children attend compulsory school from grades 1 through 9; that is, primary and lower secondary schools which children attend from approximately ages 6 through 16. These grades are completely untracked, and students typically follow their classmates through grade 9. As a result of how schools are funded in Denmark and the fact that all school-level teachers are college-educated, there is less variation in school quality compared to many other countries. Thus, Denmark presents itself as a best-case scenario in that if we can detect interaction effects between genetic influences, family SES, and school SES, they will likely exist in other systems too. In Denmark, children are as a general rule allocated to schools based on their school district. Districts usually span different neighborhoods with different socioeconomic compositions, and this creates variation in the parental SES composition of schools. Although variation in school quality is comparably low in Denmark, significant differences still exist across schools. Recent studies also show that teachers in Denmark sort into schools in ways that lead high-SES schools to have higher-quality teachers on average (Gensowski et al., 2020). Moreover, school SES composition reflects differences in peer environments and culture, also likely producing differences in learning environments.

4.3.2 Measurements

Our dependent variable is *compulsory school grade-point-average (GPA)*, which covers grades from a wide range of courses at the end of lower secondary school. This includes grades in the major reading, writing, grammar, and oral abilities, English (oral), mathematics, and physics/chemistry. GPA is based on grades awarded by the teacher during the school year and grades on final exams awarded by teachers and external examiners. Grades are measured on a 7-point scale ranging from -3 to 12 and include the values -3 (ECTS equivalent: F), 0 (F+), 2 (E), 4 (D), 7 (C), 10 (B), 12 (A) (Ministry of Children and Education, 2023). According to Table 4.2, girls have higher GPAs on average and slightly lower dispersion than boys. Moreover, female siblings have on average a slightly higher GPA than female twins in the full sample.

24 Although siblings mostly attend the same school, not all go to the same schools. In our sample, 75 percent attend the same type of school (as defined by family SES composition), and 67 percent attend the same school. We rely on attending the same school type instead of same school, but the choice is not expected to influence the results. For testing the school SES interaction, it is important that twin and sibling pairs attend the same type of school. Including children attending the same type of school but not the same school, leads on average to greater environmental similarity for the group of twins than the group of siblings, because twins are more likely to attend the same school. Our robustness checks show that assuming a lower environmental relatedness among siblings does not affect our conclusions.

Our two focal variables are *parental SES* and *school SES composition*. We measure SES by whether at least one parent has obtained a college degree (i.e., the equivalent of a bachelor's degree or higher). Parental education is the most stable and most important indicator of parental SES when predicting children's educational performance (Sirin, 2005).[25] We construct the school SES composition measure using the full population of children completing compulsory school between 2003 and 2014. We calculate by school and graduation cohort the share of children with college-educated parents and then divide this measure into two equally sized groups (cut at the median). As Table 4.2 shows, both the school SES composition variable and parental college variable are well balanced across the four groups.

Table 4.2. Descriptive statistics for the analytical sample.

	Male twins	Male siblings	Female twins	Female siblings
N_{pairs}	2,459	16,974	2,551	15,309
Average GPA	6.33	6.38	6.85	7.06
SD GPA	2.31	2.22	2.24	2.15
Proportion college-educated parents	0.40	0.43	0.38	0.41
Proportion high-SES schools	0.53	0.54	0.52	0.53

4.3.3 Methods

We apply the Classical Twin Design (CTD) to sibling and twin data to separate genetic and environmental sources of variance in educational performance. In the CTD, the similarity in an outcome among MZ twins is compared with the similarity among DZ twins. MZ twins are genetically identical while DZ twins share on average half of their segregating genes. Based on this information and the assumption that both types of twins share their shared environment to the same extent, the variance in an outcome can be decomposed into three components: genetic (*A*), shared environmental (*C*), and non-shared environmental variance (*E*) (Knopik et al., 2016). The genetic component captures additive genetic influences on educational performance, which include genetically influenced characteristics such as cognitive skills and non-cognitive skills (Demange et al., 2021; Krapohl et al., 2014). The genetic component may also capture more complex processes such as gene-environment correlations and interactions if these are present but unmodeled (Rijsdijk & Sham, 2002). Shared environmental variance includes all non-genetic sources of twin similarity in educational performance, including parental SES, shared school experiences, and neighborhood characteristics (Engelhardt et al., 2019). Non-shared environmental variance captures non-genetic sources of twin dissimilarity, including distinct events (e.g., differential treatment by parents or teachers, peer influence, illness) or events that twins experience differently. It also includes variance due to random measurement

25 Having a college degree or not is substantively relevant because it is an important dividing line in society. Moreover, in our case it is also practical to have binary variables given that it makes our results with multiple interactions easier to communicate (see Purcell, 2002).

error. Measured environmental characteristics, in our case family SES and school SES, are always shared between siblings and therefore explain shared environmental variance (Turkheimer et al., 2005). This means that the shared environmental variance that is left after taking into account the environmental measures refer to shared environmental variance not related to our family SES and school SES measures.

Our data do not contain information on zygosity. Therefore, we rely on data on SS twins and SS non-twin sibling pairs. Previous studies used twins' sex composition to approximate their zygosity (e.g., Erola, Lehti, Baier, & Karhula, 2021; Figlio et al., 2017; Pokropek & Sikora, 2015; Rodgers, Rowe, & May, 1994). Opposite-sex (OS) twin pairs are always DZ and share, just as non-twin sibling pairs, on average half of their genes.[26] SS twin pairs are a mixture of DZ and MZ twin pairs and thus genetically more similar, yet, the exact average genetic correlation is unknown. There are different approaches to deal with this. The approach that we use is based on the assumption that among SS twins the mixture of MZ and DZ twins is fifty-fifty (see, e.g., Rodgers et al., 1994). The genetic relatedness of SS twins then equals the average of MZ and DZ twins' genetic relatedness, that is, .75.[27] We performed several robustness analyses to check this assumption, and the assumption that SS twins and sibling pairs share their environment to the same extent. If the genetic relatedness of twins is lower than .75, this results in larger genetic variance and smaller shared environmental variance. And if siblings do not share their shared environment to the same extent as twins, this leads to a smaller genetic variance and larger shared environmental variance. Importantly, while the imposed genetic relatedness could affect the levels of the *ACE* components, it cannot affect the gender gaps in these components.

We use structural equation models using full information maximum likelihood estimation in Stata to decompose the total variance in educational performance into *A*, *C*, and *E*. The twin and sibling pairs differ in their (assumed) genetic correlation (.75 vs. .50, respectively). The shared environment is assumed to be perfectly correlated between twins and between siblings, while the non-shared environment is uncorrelated between twins and between siblings as it only contributes to within-pair differences. Based on this parameterization, the expected covariances in the educational performance of SS twins and siblings are:

$$Cov_{SS\ twins} = .75A + C \quad (4.3)$$

$$Cov_{SS\ siblings} = .50A + C \quad (4.4)$$

26 We do not use OS twins (nor OS siblings). Comparing SS and OS twins with unknown zygosity provides insufficient information to identify a model in which both *ACE* components and sex differences are separately identified. Additionally, using opposite-sex twins requires the assumption that they have only lower similarity in educational performance than SS twins because they are less genetically similar, not because they differ in sex (Figlio et al., 2017). Excluding OS pairs may lead to more robust results if gender differences exist (see Pokropek and Sikora, 2015).

27 Other approaches include estimating the genetic relatedness based on the number of SS and OS twin pairs in the data or use available twin registry information on the number of MZ twins among same-sex twins. In our case, this results in a somewhat lower estimated genetic relatedness for SS twins (.70).

Therefore, *A* can be estimated by:

$$A = (Cov_{SS\,twins} - Cov_{SS\,siblings})/(.75 - .50) \quad (4.5)$$

Shared environmental variance is the residual of the twin or sibling covariance not accounted for by *A*. Variance unexplained by *A* and *C* results from non-shared (or unique) environmental influences including measurement error (*E*).

We estimate the covariances and resulting variance components for different subgroups, also known as non-parametric gene-environment interaction analyses. The subgroups include family SES (low vs. high), school SES (low vs. high), gender (boys vs. girls), and their intersections. The means and variances of educational performance were constrained to be equal between twins and siblings and allowed to differ between the different subgroups (i.e., family SES, school SES, and gender). We rely on the standardized variance components, but also examine the unstandardized ones. Focusing on only the standardized variance components would conceal the underlying processes. By including the unstandardized components, we can investigate if a gene-environment interaction is driven by absolute changes in both *A* and *C*, or only by a change in one of the two. We performed several checks to test whether our results are robust to different modeling specifications, which we report at the end of the results (see section 4.4.4).

4.4 Results

4.4.1 ACE decompositions by gender, family SES, and school SES

First, we investigate to what extent the *ACE* decomposition in educational performance differs by gender, family SES, and school SES separately. In the full sample, the proportional genetic contribution (i.e., heritability) is 68.6% for boys and 76.3% for girls, but this gender difference is not statistically significant at a five percent level (p = .082). Both environmental sources of variance are larger for boys than for girls, but only the difference in non-shared environmental variance is statistically significant (difference = 5.8, p = .002).

To examine general (non-gender specific) patterns of a gene-environment interplay in the Danish student population, we break down the *ACE* decomposition by family SES and school SES pooled over both genders (see Table 4.3). For family SES, we find that the proportional genetic contribution (i.e., heritability) is 86% in low-SES families and 75% in high-SES families. Additionally, we find larger contributions of the shared (yet, not statistically significant) and non-shared environment in high-SES families than in low-SES families (see Table 4.3). These results are consistent with the compensation interaction pattern. The unstandardized components show that the finding of lower genetic variance in high-SES compared to low-SES families results from both smaller absolute genetic variance and larger environmental variance (see Table 4.3). The decrease in genetic variance is stronger than the increase in environmental variance, leading to the less total variance in educational performance in high-SES families.

For school SES, the minor (and non-significant) differences in genetic and environmental influences are in the same direction as family SES (see Table 4.3). Although we find some

differences by school SES, the genetic and environmental influences are largely similar between low-SES and high-SES schools. Thus, we find no support for a general gene-environment interaction for school SES.

Table 4.3 ACE decomposition for GPA separately by family SES and school SES.

	1. Family SES						2. School SES					
	Low-SES (*N*=21,734)		**High-SES (*N*=15,559)**		**Difference**		**Low-SES (*N*=17,373)**		**High-SES (*N*=19,920)**		**Difference**	
Mean	6.04 (0.01)	***	7.60 (0.01)	***	1.57 (0.02)	***	6.10 (0.02)	***	7.20 (0.01)	***	1.10 (0.02)	***
Variance	4.68 (0.04)	***	3.89 (0.04)	***	-0.79 (0.05)	***	4.92 (0.04)	***	4.41 (0.04)	***	-0.51 (0.06)	***
% *A*	85.74 (3.05)	***	74.94 (4.05)	***	-10.80 (5.07)	*	77.77 (3.37)	***	74.79 (3.12)	***	-2.99 (4.59)	
% *C*	11.47 (1.87)	***	16.23 (2.39)	***	4.77 (3.03)		17.68 (2.04)	***	20.16 (1.88)	***	2.48 (2.78)	
% *E*	2.80 (1.30)	*	8.83 (1.79)	***	6.03 (2.21)	**	4.55 (1.46)	***	5.05 (1.35)	***	0.50 (1.99)	
A	4.01 (0.14)	***	2.91 (0.16)	***	-1.10 (0.21)	***	3.83 (0.17)	***	3.30 (0.14)	***	-0.53 (0.22)	*
C	0.54 (0.09)	***	0.63 (0.10)	***	0.09 (0.13)		0.87 (0.10)	***	0.89 (0.09)	***	0.02 (0.13)	
E	0.13 (0.06)	*	0.34 (0.07)	***	0.21 (0.09)	*	0.22 (0.07)	**	0.22 (0.06)	***	0.00 (0.09)	

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs. Results are based on models with an average genetic relatedness of .75 for SS twins.

4.4.2 Gene x SES interactions by gender

To examine how gene-SES interactions depend on gender, we further break down the variance decompositions by gender. We do this for two separate models, one for family SES and one for school SES. Figures 4.1 and 4.2 show the results, and we refer to Table D1.1 in Appendix D1 for more details including standardized and unstandardized *ACE* estimates, total variances, and means.

We find that the gene-family SES interaction differs by gender (see Figure 4.1). The relative variance components show that for boys, but not for girls, *A* is statistically significantly lower in high-SES families (difference = 21.25, *se* = 7.73, p =.006), while both *C* (difference = 10.91, *se* = 4.53, p =.016) and *E* (difference = 10.33, *se* = 3.45, p = .003) are significantly higher (see Figure 4.1). We observe this interaction for boys because we find no gender differences in the *ACE* decomposition in low-SES families, while we do find these differences in high-SES families. In high-SES families, *A* is considerably lower for boys than for girls. For boys in high-SES families, 66% of the variance in educational performance can be attributed to genetic variance; for girls, 90%. Boys in these

families are thus more dependent on environmental influences. This result holds for both the shared environment and the non-shared environment (see Figure 4.1a). The unstandardized variance components (Figure 4.1b) show that the gene-SES interaction for boys is driven by both absolute lower genetic variance in high-SES families and higher non-shared environmental variance. The unstandardized variance components also show a statistically significant decrease in genetic variance for girls with increasing SES. Given that for girls the shared and non-shared environmental variance decrease with a similar degree, we do not find this when we look at the relative components.

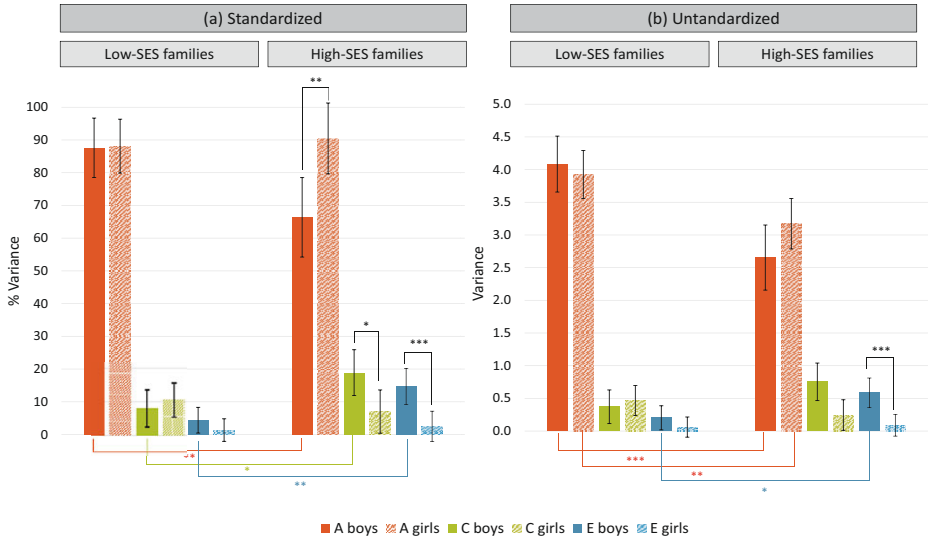


Figure 4.1 ACE decomposition for boys and girls by family SES.

Note: *p<.05, **p<.01, ***p<.001 (two-tailed). Genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variance including 95% CI. Results are based on models with an average genetic relatedness of .75 for SS twins.

In contrast to family SES, school SES barely shows any gender differences (see Figure 4.2). We do not find any gene-environment interaction with the school SES environment when looking at the standardized components (Figure 4.2a). The relative contributions of *A*, *C*, and *E* do not statistically significantly differ between low- and high-SES schools among both boys and girls. In low-SES schools, the non-shared environmental contribution to boys' education performance (E = 8.82%, se = 2.35, p<.001) is significantly larger (p = .007) than for girls (E = 0.70%, se = 1.89, p = .710). The unstandardized variance shows less genetic variance for girls in high-SES schools. Other than that, there are no statistically significant gender differences in the genetic and environmental sources of variance for children attending low- and high-SES schools.

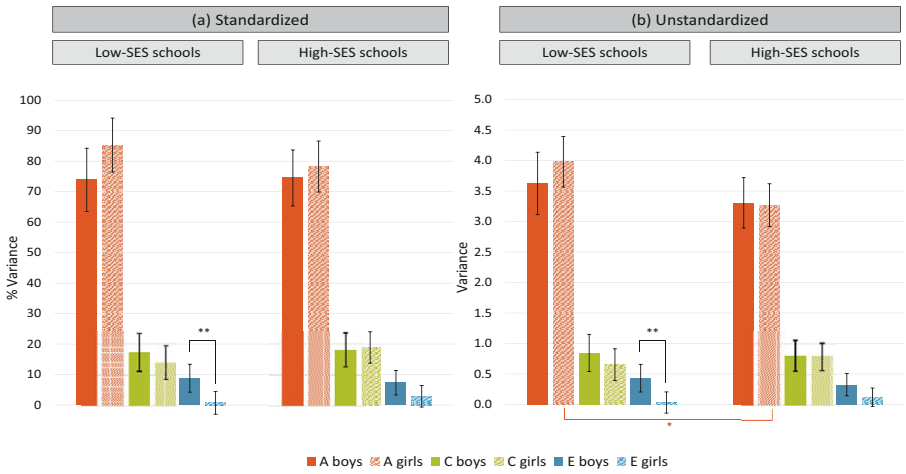


Figure 4.2 ACE decomposition for boys and girls by school SES.

Note: *p<.05, **p<.01, ***p<.001 (two-tailed). Genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variance including 95% CI. Results are based on models with an average genetic relatedness of .75 for same-sex twins.

4.4.3 Gene x family SES x school SES interactions by gender

Although high-SES children more often attend high-SES schools, this relationship is far from perfect. In our data, 25% of high-SES children attend low-SES schools, whereas 38% of low-SES children attend high-SES schools. In Figure 4.3 (and Table D1.2 in Appendix D1), we report the *ACE* decomposition for boys and girls for the interaction between school and family SES.

We find that the moderating effect of family SES that we found for boys is almost entirely driven by children attending low-SES schools. For boys attending low-SES schools, standardized *A* is significantly lower in high-SES families compared to low-SES families (difference = 36.29, *se* = 14.37, *p* =.011), while both *C* (difference = 21.56, *se* = 8.19, *p* =.009) and *E* (difference = 14.74, *se* = 6.57, *p* = .025) are significantly larger in high-SES families. For boys in high-SES schools, on the other hand, we do not find such interaction for family SES. Here we only find that boys' non-shared environmental variance is significantly larger in high-SES families than in low-SES families (difference = 10.61, *se* = 4.46, *p* = .018). This stronger gene-family SES interaction (for boys) in low-SES schools is in line with the idea of families and schools acting as substitutes. The unstandardized results (Figure 4.4b) likewise show that the gendered gene-SES interaction is concentrated in low-SES schools.

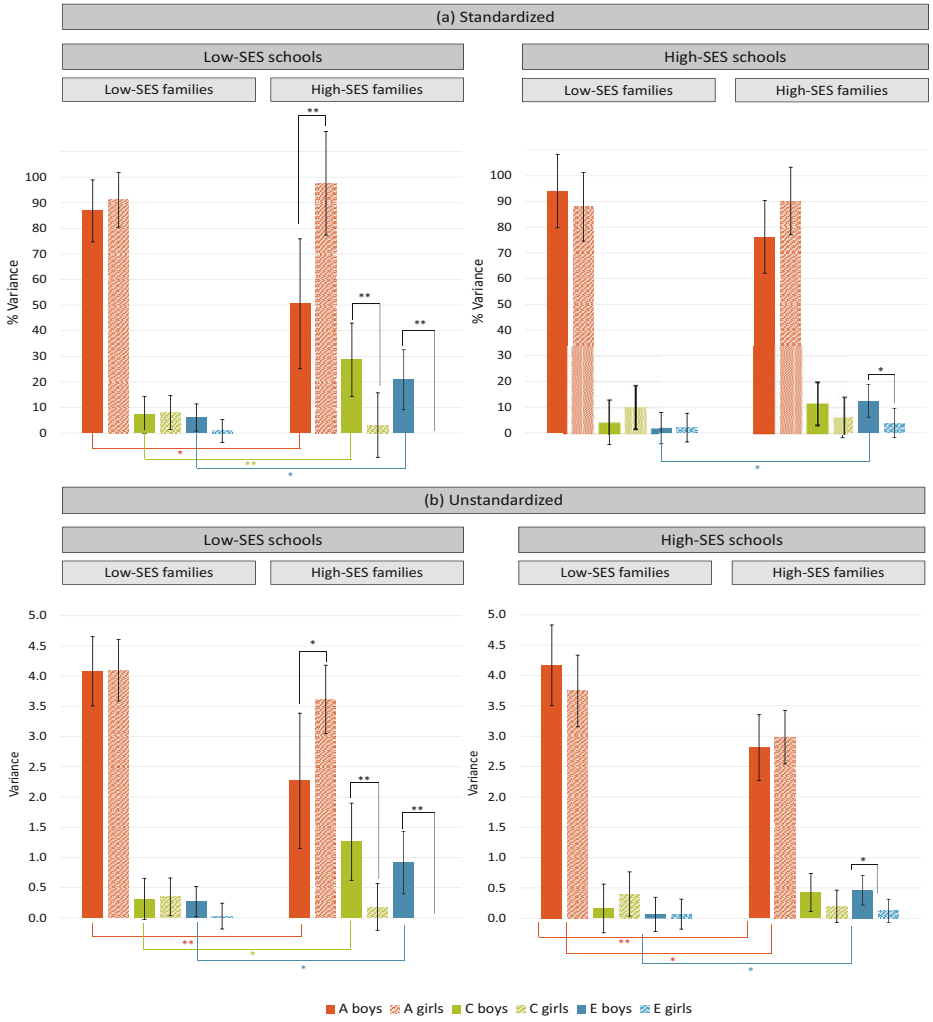


Figure 4.3. Standardized (a) and unstandardized (b) ACE decomposition for boys and girls by school SES x family SES.

Note: *p<.05, **p<.01, ***p<.001 (two-tailed). Genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variance including 95% CI. Results are based on models with an average genetic relatedness of .75 for SS twins.

4.4.4 Robustness checks

To test the robustness of our findings, we perform five auxiliary analyses (reported in Appendix D2). First, we estimate the interactions by gender, school, family SES, and their intersections based on the Intraclass Correlation Coefficients (ICCs) of twins and siblings. The interaction between genes and the SES environment involves changes in the twin and sibling (dis)similarities as a

function of SES. Parameterizing these (dis)similarities in ICCs instead of *ACE* components has the advantage that assumptions on the genetic and environmental relatedness of twin and sibling pairs are not needed, although it results in less precise information on the sources of (dis)similarities (Turkheimer & Horn, 2014). Results based on the ICCs are similar to those we report based on the *ACE* model. Second, we take into account that the genetic similarity of twins may not be exactly .75 and the environmental similarity of siblings may be lower than 1. To reflect both sources of uncertainties, we performed analyses with .70 and .80 as an estimate for the average genetic relatedness among SS twins. The gender differences and interactions by family and school SES are similar, pointing to that the substantive conclusions in the main analyses hold up. Third, we use alternative birth spacing among siblings (two instead of three years) to check if the potentially smaller environmental relatedness among them compared to twins influences our results. The gender differences and interactions by family and school SES are substantively similar, but we have a harder time detecting significant interaction terms which may stem from the loss of statistical power. Fourth, controlling for cohort fixed effects in our structural equation models does not change any of our results. Fifth, we test whether our results are robust against coding school SES into terciles. Our conclusions remain the same as those reported in the results section.

4.5 Conclusion and discussion

Children differ in their educational performance and one reason for this lies in genetic differences between them. Other common explanations include differences concerning family background, school environment, and gender. These influences are not independent. The importance of genetic influences on education depends on the social context, including the family and school context (Shanahan & Hofer, 2005). Also, non-genetic research suggests that boys' educational performance is more dependent on the school and family environment than that of girls (Autor et al., 2019; Legewie & DiPrete, 2012). We add to these strands of literature by considering gene-environment interactions at the intersection between family background, school SES context, and gender. Our empirical analyses on the educational performance of about 37,000 Danish twin and sibling pairs from population-wide administrative registers reveal three main findings.

First, we find a gene-environment interaction for family SES, but not for school SES. In both low-SES and high-SES families, genetic influences play a large role in explaining children's educational performance, but they do so to a greater extent in low-SES families. This compensation pattern is in line with the diathesis-stress model and the sociological idea of compensatory advantage (Paris, 1999; Pennington et al., 2009; Rende & Plomin, 1992; Shanahan & Hofer, 2005). In low-SES families, environmental risks factors are more often present such as stressful life events, which may increase the expression of genetic risks for lower performance. In high-SES families, on the other hand, such risks factors are less often present. Moreover, high-SES parents may be more likely to compensate for the expression of genetic risks, for example, by providing a more favorable literacy environment in case of higher genetic risks for poor reading (Pennington et al., 2009). This finding contradicts the bioecological model, in which case we would have found an enhancement pattern with more genetic variance in high-SES families as genetic potential would

be more fully realized in such environments (Bronfenbrenner, 1992; Bronfenbrenner & Ceci, 1994; Rowe et al., 1999; Scarr-Salapatek, 1971).

Second, we find that the gene-family SES interaction is further moderated by the child's gender. In high-SES families, the genetic influence is considerably lower for boys than for girls. Additionally, shared and non-shared environmental influences are larger for boys than girls in these families. We thus find a gene-family SES compensation interaction for boys but not for girls. This result is consistent with the idea that environmental factors have a larger impact on boys, due to differential investments and/or differential sensitivity (see, e.g., Autor et al., 2019). One interpretation of this finding is that high-SES parents provide more compensatory investments in boys such as tutoring, investments that eventually reduce their genetic risk toward lower educational performance.

Third, the moderating effect of family SES for boys is almost entirely driven by low-SES schools. In low-SES schools, but not in high-SES schools, genetic influences play a much smaller role and shared and non-shared environmental influences play a larger role for high-SES boys. In other words, for boys in low-SES schools, our findings are in line with the compensation interaction pattern, suggesting that genetic risks are more actualized in low-SES families and neutralized or compensated in high-SES families. One explanation for this finding is that especially the genetic risks towards lower school performance of boys in low-SES schools may be compensated by parental SES. Especially for boys, low-SES schools may provide a disadvantageous context, as boys are more responsive to an environment that is less learning-orientated and more disruptive (e.g., Legewie & DiPrete, 2012). It may be that especially with such risky school circumstances, a high-SES family environment has a compensating influence. This also corresponds with the idea that family and school influences substitute each other's influence, leading to a more pronounced gene-family SES interaction in lower-SES schools.

In the United States, the gene-SES interaction is more often found to follow the bioecological model, while the interaction was not found or even reversed in European countries (e.g., De Zeeuw & Boomsma, 2017; Tucker-Drob & Bates, 2016). Our finding of greater genetic variance in low-SES families contributes to the evidence pointing towards a reversed interactions outside the United States. It also contributes to the increasing empirical studies into gene-environment interactions in education within Europe. A recent cross-national study by Baier et al. (2022) find no significant gene-family SES interactions in GPA in Germany and Norway, but a significant gene-environment compensation interaction in Sweden. We likewise find a significant compensation interaction in Denmark. Country-specific welfare state arrangements and features of educational systems may play a role, but more comparative research is needed to understand how such contextual characteristics affect gene-environment interactions in education.

Furthermore, while we find that genetic influences play a larger role in low-SES families, this pattern does not extend to school SES. In contrast to the United States (Hart, Soden, et al., 2013) where larger genetic variance in low-SES schools is found, we do not find such an interaction in Denmark. This could be due to the fact that there is less variation in the school environment in Denmark than in the United States, partly as a result of the public funding of schools and relatively uniform teacher quality (all teachers are college-educated) in Denmark. Still, this does not mean that school SES does not play a role in Denmark. The gene-family SES interaction appears to be

heterogenous, as it is largely concentrated among boys attending low-SES schools. This finding highlights the importance of investigating the intersection of gender, family SES, and school SES. Thus, future gene-environment interaction studies could benefit from considering the multiplicity of contexts for complex outcomes such as educational performance (Seabrook & Avison, 2010).

Most prior twin studies likely did not investigate multiple interactions as it requires large and socioeconomically diverse twin samples. Twin samples with known zygosity often lack the power to detect (multiple) interaction effects. Moreover, twin samples are more likely to suffer from self-selection bias (Trejo et al., 2018). Relying on twin-sibling analyses based on population-wide register data, as we do, has the advantage that it comprises all socioeconomic strata and provides us with many cases to investigate the gene-environment interactions. A trade-off here is that these data do not contain zygosity information. While analyses with unknown zygosity also reduce statistical power (Eaves & Jinks, 1972), the sample size provided by population data can offset such issues.

Although the twin design applied to twin and sibling register data allows us to investigate the complex interplay between genes, gender, and family and school SES, it also comes with some disadvantages. First, we relied on comparing SS sibling types that differ in genetic relatedness: twins and non-twin siblings. This design relies on two assumptions: (i) twins' genetic relatedness is 0.75 and (ii) siblings' shared environmental relatedness is 1. Violation of these assumptions may influence the *ACE* decomposition. However, we do not find that these assumptions affect our conclusions. We provide several robustness checks that relax these assumptions, and the results are largely similar. Moreover, possible violations of the assumptions do not influence gender differences in *ACE*.

Second, our *ACE* moderation models do not allow for causal interpretations. The environmental moderators that we use, family SES and school SES, are endogenous (Schmitz & Conley, 2017). Therefore, the gene-environment interactions as reported in our study could mean that effects of genetic risks are weaker in high-SES families but could also reflect environmental constraints placed on the (genetic) variation in educational performance making the high-SES subgroup more homogenous. Although both may be related to inequality, we are mainly interested in the first interpretation. Yet, drawing such conclusion requires other research designs such as applying quasi-natural experimental approaches to gene-environment interaction models (Schmitz & Conley, 2017).

In conclusion, our study suggests that the importance of genetic and environmental influences on educational performance differs by (the intersections between) family SES, school SES, and gender. For boys, but not for girls, genetic influences play a considerably smaller role in high-SES families than in low-SES families. This interaction appears to mainly take place among high-SES boys attending low-SES schools. These findings show that it is important to consider the multiplicity of context. This may be especially important for educational outcomes, which are influenced by complex processes related to genes, gender socialization processes, and environments in the family and school. Focusing on only one aspect would lead to a misunderstanding of how genetic and environmental factors play a role.

CHAPTER 5

The development of inequality during primary education

Investigating the genetic and environmental sources underlying dispersion in learning

A slightly different version of this chapter has been submitted to an international journal as:
Stienstra, K. The development of inequality during primary education: Investigating the genetic and environmental sources underlying dispersion in learning.

Kim Stienstra is the sole author of this chapter.

5 Abstract

I investigate whether differences in educational performance (i.e., dispersion) and the underlying genetic and environmental differences are reproduced, exacerbated, or compensated during schooling (Grades 1-5, approximately age 6-11). I use longitudinal data comprising reading and mathematics test scores of around 5,500 same-sex and opposite-sex twin pairs, identified in the Netherlands Cohort Study on Education. Results of the biometric latent growth models lead to three main conclusions. First, the dispersion in initial educational performance that exists at the start of formal education is compensated during the schooling period, but the extent to which this is related to a decrease in genetic and environmental differences varies by educational domain. For reading, genetic differences decrease while environmental differences are largely reproduced over the school career. For mathematics, the decrease in dispersion results from decreasing environmental differences, while genetic differences are reproduced. Second, new influences are coming into play during schooling, mostly new genetic influences. Combining the development of initial sources of dispersion with these new influences results in an increase in the total dispersion in educational performance during primary education. Third, measured school characteristics did not account for (the development of) dispersion, suggesting that quality differences between schools likely play a limited role. However, schools may still be of importance in a more complex way. Via gene-environment correlations and interactions, schools may still play a role in the decrease of the dispersion in initial performance and the new genetic influences that are coming into play, which is an important direction for future studies.

This work was supported by the NWO research talent grant for the project 'Quality and inequality: The compensatory and multiplicative effects of school quality'(NWO: 406-18-557, awarded to Stienstra, Knigge, and Maas) and a grant for the usage of the Netherlands Cohort Study of Education (NCO) data by the Netherlands Initiative for Education Research (NRO: 40.5.22326.0022, awarded to Stienstra).

I am grateful to Ineke Maas, Antonie Knigge, and Rolf van der Velden for their valuable comments and discussion on this work. I thank Esmee Bosma for her help with the data cleaning of the school quality indicators. Earlier versions of this work were presented at the Migration and Social Stratification Seminar (Utrecht, 2022), the Interuniversity Workgroup Social Inequality and Life Course (ISOL) meeting (Amsterdam, 2022), the 41st Congress of the German Sociological Association (Bielefeld, 2022), and the Learning and Work Seminar (Maastricht, 2023). I thank the participants at these meetings for their helpful feedback.

5.1 Introduction

Inequalities in educational performance contribute to divergence in life outcomes, including educational attainment, income, health, and social outcomes. Important questions for public policy are, therefore, why some children perform better than others and how inequalities can be reduced. When children begin formal education, they differ already considerably in their skill levels. These differences reflect the different environments children originate from including their families (Downey et al., 2004; Reardon, 2019), as well as genetic differences between children (De Zeeuw et al., 2016). The question addressed in this paper is, how does schooling affect such initial differences? Pre-existing inequalities could be reproduced, exacerbated, and/or compensated by schools. However, what happens and why remains unclear (Downey & Condron, 2016).

A common way to assess the role of the school versus non-school environments in explaining inequality is by measuring the extent that children change in their educational performance over time (i.e., with learning gains/growth or value-added models) (Timmermans & van der Werf, 2017). While differences in performance when children enter school can only be due to non-school factors, differences in growth are due to a mixture of school and non-school factors (Von Hippel & Hamrock, 2019). To further disentangle how inequality is shaped by school versus non-school influences, studies applied a seasonal comparison design (Cooper et al., 1996). This is done by comparing learning gains during the school year (influenced by both school and non-school factors) with learning gains or losses during the summer break (influenced by non-school factors).

Such prior studies often investigated the average learning gains along the lines of ascriptive background characteristics, including gaps in learning between children from different socioeconomic status (SES) backgrounds, boys and girls, and different races or ethnicities (Alexander et al., 2007; Downey et al., 2004, 2022; Downey & Condron, 2016; Passaretta & Skopek, 2018). The influence of such ascriptive characteristics on educational performance and learning growth has been labeled 'inequality of educational opportunity' (Van de Werfhorst & Mijs, 2010). Although important and insightful, the emphasis on inequality of educational opportunity comes at the expense of understanding the broader question of how schooling affects the total variance in educational performance (Condron et al., 2021; Montt, 2011). The total variance in educational performance is another type of inequality in learning, labeled by Van de Werfhorst and Mijs (2010) as 'inequality as dispersion'. Focusing on dispersion seems particularly relevant given that ascribed characteristics such as SES only explain a small proportion of the variance in educational performance and growth (Condron et al., 2021; Downey et al., 2004; Montt, 2011; Von Hippel et al., 2018). Most of the differences thus remain unexplained by these characteristics (Condron et al., 2021).

To gain a more profound understanding of how schools shape inequality, I focus on inequality in learning as indicated by the dispersion in pupils' scores on national achievement tests in reading and mathematics taken throughout primary education. More specifically, I ask: 'Does dispersion of educational performance when children enter school increase or decrease during primary education, and why?'. To answer this, I take a behavioral genetics approach. This has the advantage that not only dispersion in performance throughout primary education can be examined, but also different sources underlying the dispersion. By using data on twins, insight can be provided

into how much of the total variance in educational performance and learning growth can be attributed to genetic variance, shared environmental variance (i.e., environmental aspects shared by twins that make them more similar, such as parental influence), and non-shared environmental variance (i.e., environmental aspects that affect twins individually, making them more dissimilar).

Prior behavioral genetics studies have shown that genetic differences explain a relatively large part of the variance in educational performance at a single measurement occasion, but differences in learning growth have only rarely been studied (Kievit et al., 2021). The relatively few twin studies that investigate learning growth mostly focus on the development of reading (and, to a lesser extent, numeracy) skills in early childhood instead of educational performance over the whole primary school period. These studies generally find that the proportion of variance that is related to genetic influences (i.e., heritability) is lower for growth than for initial performance (Christopher et al., 2013; Grasby & Coventry, 2016; Hart, Logan, et al., 2013; Logan et al., 2013; Petrill et al., 2010). This indicates that learning over time has a stronger environmental component than performance at a single measurement point. The family and/or school environment may play a role in this. However, prior behavioral genetics studies usually do not include measured environmental characteristics as specific underlying sources of dispersion. The study by Christopher et al. (2013) is a notable exception, as they included parental education and schools' average reading and writing scores. Although these measures influenced average initial performance and growth rates, they did not account for significant shared environmental variance. In this study, I include measured school and family characteristics to further examine the specific shared environments that play a role.

I use data from the Netherlands Cohort Study on Education (in Dutch: *Nationaal Cohortonderzoek Onderwijs* [NCO]), comprising longitudinal measurements of pupils' performance on national achievement tests in primary schools derived from school administrative systems (Haelermans et al., 2020). These data provide a unique opportunity to investigate the development of dispersion in reading and mathematics because of three great advantages. First, the test scores are scaled in such a way that children's achievement is measured on a single continuous scale throughout primary education (via Item Response Theory), making the test scores comparable between schools, grades, and measurement moments. This overcomes the problem of some previous studies where results on the change in achievement gaps during schooling were biased by test score scaling artifacts (Von Hippel & Hamrock, 2019). Second, given that children are frequently tested in Dutch primary education, I can rely on ten measurement points. Prior studies measuring learning growth based on two time points face the risk that pupils may have a very high or low score on the first test by chance and will perform worse or better on the second test, respectively (Caro et al., 2009). The resulting decrease in dispersion would then be driven by regression toward the mean. Relying on many time points minimizes this source of bias. Lastly, the administrative nature of the data including the many observations allows for the identification of sufficient twins to apply biometric latent growth models (McArdle et al., 1998). In this way, I can investigate how much of the dispersion at the beginning of schooling is related to genetic and environmental differences between children, and whether these underlying sources of dispersion increase or decrease over the primary school career. Moreover, these models provide information on genetic and environmental influences that come into play later during primary education (i.e.,

are unrelated to initial performance). Another advantage of the administrative nature is that data on children's family and school environment is available or can be linked, allowing me to further examine the shared environmental influences that play a role (in this case, school quality, school SES, and family SES).

Altogether, this will provide insight into the extent to which dispersion in educational performance in general – and the underlying genetic and (non-)shared environmental differences in specific – is exacerbated or compensated during the primary school period. This is important to scrutinize for both research and policy, as it provides insight into how inequality develops and whether more attention should be paid to identifying and counteracting inequality-generating mechanisms before or during schooling.

5.2 Theory

When children enter primary school, they differ in their educational performance levels. The dispersion in educational performance could increase, decrease, or remain stable throughout primary education. The development of the total dispersion in performance during schooling consists of two parts. First, initial sources of dispersion in performance (i.e., differences present at the start of primary education) could increase or decrease. Second, new sources of dispersion could come into play during the schooling period. I first outline environmental and genetic reasons for dispersion in initial performance. Then, I discuss how initial environmental and genetic dispersion in educational performance might develop over time. Lastly, I discuss the new sources of dispersion that could come into play over the school career.

5.2.1 Differences in initial educational performance

Children's initial performance can be seen as the outcome of the advantages that children bring to school (Hanselman & Fiel, 2017). These advantages relate to individual characteristics (e.g., cognitive ability) and early learning opportunities. Since children have not been exposed to the primary school environment, the learning opportunities are largely related to the family environment (Reardon, 2019). Depending on parental SES, children receive different learning opportunities at home. Studies show time and again that children from more advantaged backgrounds perform better in school (Sirin, 2005). Such SES differences in performance are already present when children start schooling (Passaretta & Skopek, 2021), emphasizing the role of early educational opportunities as provided by children's home environments (Bradley & Corwyn, 2002). These early educational opportunities include, for example, access to books and educational toys, and parental stimulation of learning to count and read, which are all more prevalent in high-SES families (Bradley & Corwyn, 2002). Besides the home environment, there are also other environmental conditions shaping children's early educational opportunities including neighborhood conditions (e.g., presence of public libraries), high-quality childcare, and preschool programs (Reardon, 2019).

Initial educational performance does not only reflect differences in children's environmental circumstances, but also genetic differences between children. Genes influence educational

performance via embodied characteristics that are valued in school, such as cognitive and non-cognitive skills (Freese & Jao, 2017; Krapohl et al., 2014). For Dutch school pupils at the early stage of primary education (around age 6-8), it has been found that the relative genetic contribution (i.e., heritability) is rather high for different educational domains (arithmetic, reading, reading comprehension, and spelling), accounting for 60 to 80% of the variance (De Zeeuw et al., 2016). Such genetic influences do not operate in a vacuum but are expressed in correlation and interaction with the environment (Sameroff, 2009; Tucker-Drob et al., 2013). Children with different genotypes evoke different responses from their environments and increasingly select (and are selected in) environments based on early genetically influenced behaviors and characteristics (Tucker-Drob et al., 2013). For example, children's reading ability may be more stimulated for children who have a genetic predisposition towards reading because they are more likely to pick up a book (i.e., active gene-environment correlation) and parents may be more likely to buy a book for them (i.e., evocative gene-environment correlation) than for children whose genetic predisposition is less inclined towards reading (Knigge et al., 2022).

In short, when children enter school, they can be expected to differ in their educational performance because of the different environmental conditions they are exposed to prior to schooling and because of genetic differences between them. These genetic differences are not fixed and immutable but reflect the realization of genetic endowments via environmentally mediated processes.

5.2.2 Development of differences in educational performance

Children differ in their initial performance at the start of primary education and these differences are related to both genetic and environmental differences between children. These differences could persist throughout primary school. In this case, the initial differences in educational performance would be reproduced (see Figure 5.1 – reproduction). It is, however, questionable whether differences in initial performance are simply reproduced over the school career (Dumont & Ready, 2020; Pfost et al., 2014). There are opposing ideas on how initial differences in performance develop over time. Initial environmental and genetic influences affecting educational performance at the beginning of school could become more important over the school career (Figure 5.1 – exacerbation), or less important (Figure 5.1 – compensation). Depending on whether initial differences are exacerbated or compensated during the school career, dispersion in educational performance would increase or decrease, respectively. This does not necessarily mean that the *total* dispersion in educational performance changes in the same way over time, because dispersion does not only result from the initial dispersion but also from new sources of dispersion that come into play over time. If there are such new influences, some children may progress faster in their educational performance than others, leading to an increase in the total dispersion in performance over the school career.

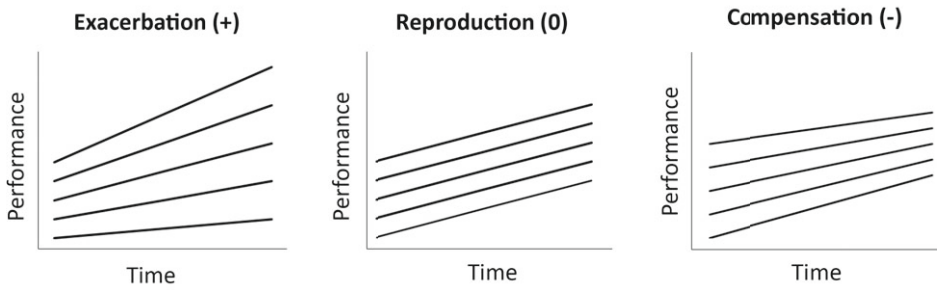


Figure 5.1 Possible associations between initial performance and growth.

Exacerbation of initial differences

Initial differences in performance can increase over time, which is expressed as a positive association between initial performance and learning growth. Pupils who initially perform well would then show the greatest learning gains, while those lacking the initial advantages fall further behind (Kwiatkowska-White et al., 2016). This has also been referred to as a 'Matthew effect' or 'cumulative advantage' (DiPrete & Eirich, 2006; Stanovich, 1986). Exacerbation of initial educational differences could occur because environmental differences increase over time and/or because genetic differences increase over time. This is expected to largely operate via interactions between the initial environmental and genetic influences on performance at the start of schooling and later (school) influences.

Environmental differences in performance, which are expected to mostly relate to family background, could increase during primary education due to cumulative advantage processes. In this case, increasing dispersion is not due to an increased influence of parental SES per se but rather the legacy of pre-existing SES inequalities (Hanselman, 2018). High-SES pupils may increasingly benefit from their SES background during schooling, which could be explained by the idea that 'skills beget skills'. Children's skills gained early in life (e.g., in effective home environments) increase children's capacity to benefit from later instruction in school (Heckman, 2000; Sørenson & Hallinan, 2016). For example, high-SES children enter the school with better academic preparation because high-SES parents emphasize verbal skills more (e.g., by engaging children more in conversations) and provide more experiences that foster learning (e.g., taking children to cultural events) (Bradley & Corwyn, 2002). They may therefore develop their skills faster in school, for instance, because they understand instructional materials better (Hanselman, 2018). Certain school practices can also contribute to such 'rich-get-richer' dynamics. For example, if teachers have higher expectations of high-SES pupils, pupils with different SES backgrounds will get different learning opportunities. High-SES pupils will then further improve their educational performance while low-SES pupils are more hindered in their learning, increasing the SES differences in performance over time. Initial family background differences could also increase over time because SES influences on educational performance increase. Parental SES and parental involvement positively relate to learning growth (Fan, 2001). This could be because

high-SES parents are better able to adapt their children's school experiences to their children's (perceived) needs (e.g., arrange special educational resources for their child such as a gifted program) (Horvat et al., 2003). Parental interventions in school could complement children's learning opportunities, making them especially effective for advantaged pupils (Hanselman, 2018).

Similarly, dispersion in initial performance could increase because initial genetic differences increase over time. This process has been labeled 'amplification' in the behavioral genetics literature (Plomin, 1986; Tucker-Drob et al., 2013). Existing genetic differences could increase via transactional processes. Children select and evoke environments and responses based on (small) genetically influenced differences in educational performance, which will lead to increased genetic differences over time (Briley & Tucker-Drob, 2013). For example, genetically influenced differences in early cognitive or non-cognitive abilities (e.g., motivation, intellectual interest) may lead some pupils to take more advanced educational material or spend more time in intellectually stimulating activities (Tucker-Drob et al., 2013). Dispersion in educational performance will then increase over time. Additionally, there may be genetic influences that are not very important for educational performance at the beginning of the schooling period but become increasingly important in later grades (Briley & Tucker-Drob, 2013). This could relate to general cognitive processes that contribute to learning progress (e.g., memory, attention control) (Welsh et al., 2010), but also to genetically influenced non-cognitive characteristics that are important for growth. It has been shown, for example, that motivation and learning strategies were strong predictors of achievement growth but were less strongly related to initial performance (Murayama et al., 2013).

The exacerbation of genetic differences in children's performance can be facilitated by school practices, similar to how this works for SES differences. Initially high performing pupils (whose high performance can be partly explained by genetic endowment) may get praised more often by teachers or receive more attention from them, leading to increased motivation and self-efficacy that positively reinforces performance (Ready, 2013). Also learning opportunities could be differentiated by performance. If students are grouped based on their prior performance within or between classrooms (i.e., ability grouping or tracking), this may widen the genetic (and environmental) differences underlying performance and thereby increase the gap between low- and high-performing students (Hallinan, 1988; Welsh et al., 2010).

Compensation of initial differences

Initial differences in educational performance could also decrease over time. In that case, initially lower-performing pupils have higher growth rates than the pupils who initially performed higher (Kwiatkowska-White et al., 2016). This would be expressed as a negative association between initial performance and learning growth and would contribute to a decrease in dispersion in performance over the school career.

Dispersion in initial performance could be compensated because the SES influence on performance may become less important over the school career. If this is the case, pupils who initially benefited from their SES background at the start of primary school benefit less over time. Conversely, those who were hindered by their SES background in their initial performance are hindered less throughout primary school. One reason for such a pattern is that schools tend

to provide most benefits for the children who need it most, including pupils from lower-SES backgrounds and those with lower initial performance (Downey et al., 2022). For example, the instructions in class could be directed to the level of average or below-average pupils, and teachers may also spend more time with them (Ready, 2013; Wright et al., 1997). Initially low-performing pupils may then improve faster (and high-performing pupils may not be challenged enough), leading to a compensatory growth pattern. Schools could also compensate for initial disadvantage in a less active way. The school environment varies less than non-school environments (Downey et al., 2004). Children learn in groups and share the same teacher, educational material, and learning goals, whereas learning at home is less standardized and less homogeneous. Going to a standardized school environment thus makes that high-SES pupils are less able to continue reaping the benefits of their family environment while low-SES pupils may be less hampered by their family environment. This would decrease dispersion in educational performance.

A more genetically related explanation can be provided by the developmental lag model. According to this model, children differ in the rate at which cognitive skills develop, so children who initially have a lower performance will catch up over time (Francis et al., 1996). Some children, especially those who grow up in poverty, may have a delay in mental processes that support effective learning and have a large genetic component, including working memory and attention control (Welsh et al., 2010). If they develop these processes later during the school years, this may result in faster learning growth rates for them. It has also been hypothesized that children with reading disabilities may have a lag in the maturation of the brain and poor readers could catch up to their peers as the brain matures (Francis et al., 1996).

New sources of dispersion

During primary education, new influences on educational performance can come into play. These are influences that are not related to initial educational performance but consistently and systematically affect later performance. Such new influences increase the dispersion in performance by definition. The new influences that come into play over time could be environmental and genetic. Environmental influences can be expected to mostly relate to the school environment, which becomes prominent in children's learning when they start formal education. Children in high-quality schools may develop their skills faster than those in low-quality schools (Borghans et al., 2015b; Reardon, 2019). School effectiveness research identified numerous factors that could play a role, including teaching practices, whether teachers have high expectations of students, classroom climate, and instructional effectiveness (Scheerens, 2000). For example, Connor et al. (2009) found that children who received more precise individualized instruction had stronger literacy skill growth.

The family environment could play a role too, but probably to a much lesser extent. To be a *new* source of dispersion, the family influences should reflect influences that come into play during schooling that are unrelated to performance at the start of schooling. This could be, for example, parental school involvement (e.g., in school activities and via teacher contact) (Fan, 2001). Most family influences are likely influences that already operate before children enter school and continue to have an influence during schooling.

Besides new environmental influences, there could be new genetic influences that come into play over time, a process that has been labeled 'innovation' in the behavioral genetics literature (Plomin, 1986; Tucker-Drob et al., 2013). New genetic influences may come into play as part of natural development: biological maturation across childhood may activate genes (Briley & Tucker-Drob, 2013; Welsh et al., 2010). Next to the emergence of novel genetic influences as a function of age, also (new) environments could activate genes (Briley & Tucker-Drob, 2013). For example, it has been found that new genetic influences on reading skills arise when children move from kindergarten to the first grade of primary school (Byrne et al., 2007). This could occur because the activation of genes is dependent on certain experiences, but also via genetically influenced characteristics that were irrelevant to performance in the home and pre-school context but become increasingly important for performance in primary school (Briley & Tucker-Drob, 2013). The latter could relate cognitive characteristics that are irrelevant for initial performance but increasingly important over time due to the different (more complex) content covered in later grades (e.g., abstract thinking). It could also be related to genetically influenced non-cognitive characteristics such as learning motivation and intellectual interest (Tucker-Drob & Harden, 2012b, 2012c).

5.2.3 Empirical evidence and hypotheses

'Traditional' (i.e., not genetically informed) studies into the development of educational performance point toward a compensatory pattern, where the gap between the lowest and highest performers decreases over time. A meta-analysis of reading development showed that 42% of the findings supported a compensation pattern over all aspects of reading, 25% a stable achievement gap, and 23% a Matthew effect (Pfost et al., 2014).

Genetically informed studies not only provide insight into whether dispersion in educational performance is exacerbated or compensated during the primary school period but also provide insight into the underlying sources of dispersion. Concerning genetic dispersion, studies generally find that genetic influences on learning are smaller than those on initial performance and that new genetic influences come into play over time (Kievit et al., 2021). Conversely, (shared) environmental influences on learning growth are more substantial than those on initial performance (Christopher et al., 2013; Grasby & Coventry, 2016; Hart, Logan, et al., 2013; Logan et al., 2013; Petrill et al., 2010). There are indications that the shared environmental influences on growth are largely unique, that is, unrelated to initial performance (Grasby & Coventry, 2016; Hart, Logan, et al., 2013; Petrill et al., 2010). Since environmental influences on initial performance are thought to largely capture family background effects, these unique environmental influences on learning growth potentially capture the effect of schools. However, most studies do not include measured characteristics of the environment and one study that did so was not able to explain shared environmental variance by school and family variables (Christopher et al., 2013).

All in all, there are two opposing ideas of how inequality in educational performance develops over the school career. Pre-existing differences in educational performance could follow a cumulative pattern, leading to an exacerbation of initial dispersion over time. Alternatively, there could be a compensatory growth pattern, where initial dispersion decreases over time. Empirical evidence seems to point towards compensatory growth, yet results are mixed, and it is not

clear under which conditions accumulation or compensation takes place. Hence, I formulate two hypotheses: *initial dispersion in educational performance increases over the school career (H1a)* and *initial dispersion in educational performance decreases over the school career (H1b)*. I will explore to what extent the increasing or decreasing dispersion over the school career is related to the influence of genes or the environment. Furthermore, I will investigate the new genetic and environmental sources of dispersion that come into play over the school career and what this means for the total dispersion in educational performance. Lastly, I will examine the role of the environment more specifically by investigating the influences of measured school quality, school SES, and parental SES.

5.3 Data and methods

5.3.1 Data

I use data on national tests on reading and mathematics of the pupil monitoring system (*leerlingvolgsysteem* in Dutch, abbreviated as LVS) from Grade 1 until Grade 5 (equivalent to the Dutch grades 3-7, ages 6-11).[28] These data are collected as part of the Netherlands Cohort Study on Education (NCO). Under certain conditions, the so-called NCO-LVS-data are accessible for scientific research via Statistics Netherlands (CBS).[29] A description of the NCO is provided by Haelermans et al. (2020). Since 2014, primary schools must use a pupil monitoring system, but schools are free to choose the test provider. The LVS developed by the Central Institute for Test Development (*Cito*) is the main provider chosen by most schools (Nusche et al., 2014), and these test results are included in the dataset. Children take the tests on fixed moments twice a year: a midterm test (M-test, administered between mid-January and mid-February) and an end-of-term test (E-test, administered in June). The test scores are stored by schools' administration systems. As part of the NCO, all primary school boards were contacted with the request to share the LVS data. Participating schools receive a report on the learning growth of their pupils twice a year. After the permission of schools, the administration system exported the data to Statistics Netherlands, which pseudonymized the student and school ID. If parents objected to the export of their child's data, the data were not exported for those children. The project is ongoing, and I use the most recent data that are available for researchers at the moment of this study. These include data originating from four data exports that took place between the end of November 2020 and the beginning of August 2021 (NCO, 2022). More than 1900 schools decided to participate, which comprises almost 30% of the total number of primary schools in the Netherlands.

The data that were requested by the NCO comprise school years 2013/2014 to 2020/2021 and gradually consist of more grades over the school years. This means that for school year 2013/2014 only data for Grade 1 is available, for 2014/2015 Grades 1 and 2, etcetera (see Figure 5.2). After

28 In the Netherlands, children participate in kindergarten for two years from approximately age 4 to 6. This is part of the publicly funded primary school system and counted as the first two grades (OECD, 2016).

29 Further information and the procedure to request the data can be found on the websites of CBS (https://www.cbs.nl/en-gb/onze-diensten/customised-services-microdata/microdata-conducting-your-own-research) and NCO (https://www.nationaalcohortonderzoek.nl/onderzoek).

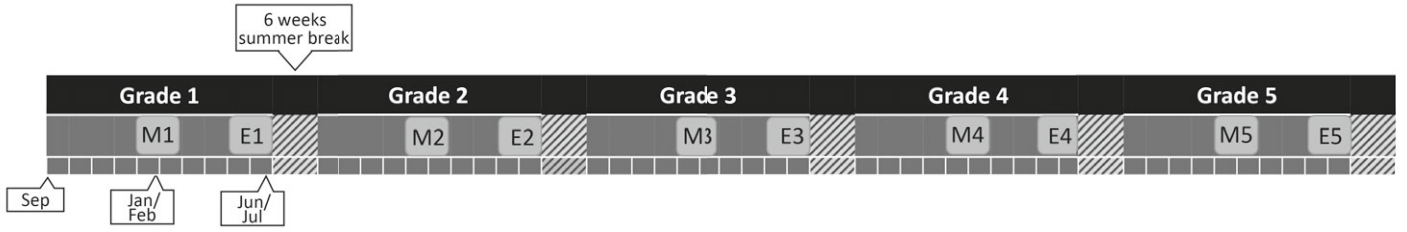
data cleaning by the NCO (e.g., removing duplicates, excluding tests for special education; see NCO, 2022), the NCO-LVS-data consist of 438,355 pupils. In less than 2% of the test records in the whole dataset, the test is not administered at the right time (NCO, 2022). Therefore, in a few cases, the variable name of the test score may indicate that the score belongs to, for example, the E-test, while this does not match with the variable indicating the corresponding test type (e.g., test type indicates an M-test). This could be because children can take a test that fits their level, for example, the M4-test for pupils who are ahead of their level when the E3-test is taken (Tomesen et al., 2018). This is not a problem because the underlying measurement technique (see Measurements section) makes sure that scores remain comparable. I focus on the test time points and not on the test type. This means that, for example, E1 refers to the test taken at the end of Grade 1, although a few pupils did not take the E1 test at that time but a test that belongs to a prior or later test moment.

In addition to the LVS data, I use information from other microdata of Statistics Netherlands (*CBS*). The files that I use are the municipal personal records database 2021 (CBS, 2021), the linked parent-child data 2021 (CBS, 1995), educational attainment file 2020 (CBS, 2019), personal income files 2013-2019 (CBS, 2011), primary school registrations 2013-2019 (CBS, 2014), and school quality data from the Inspectorate of Education.

5.3.2 Selections and selectivity

I use pupils' scores on the reading and mathematics tests of Grades 1-5, taken during the school years 2013/2014 to 2019/2020. As I am interested in general learning growth, I exclude the tests that were taken during and after the school closure due to COVID-19 (i.e., all the tests from the M-test of 2019 onward). For the respective cohorts, test results are available for 378,378 pupils. Twin pairs are identified in the linked parent-child data combined with the municipal personal records database, based on children who have the same legal parents and the same birth month and year (birthday is not available because of privacy reasons). Linking this to the NCO-LVS-data shows that of the 378,378 pupils, 10,803 individuals can be identified as twins.[30] The percentage of twins in the data is 2.9%, which is comparable to the percentage of twins in the Dutch population (Webbink et al., 2006). I exclude twins that do not attend the same primary schools ($N = 21$), reducing the sample to 10,782 twins. I also exclude 49 twins with outliers (see section 5.3.3). This leads to the final analytical sample of 10,733 twins ($N_{pairs} = 5{,}699$) of which I can analyze the learning growth in reading of 5,304 twin pairs and mathematics of 5,576 twin pairs. Note that the number of twins is lower than two times the number of twin pairs because for some pairs there is only information on one twin. Twin pairs that are incomplete can still be analyzed. Individuals who miss one or more observations on the test (e.g., due to absence on the test day) are still included in the analyses if there is data on the co-twin and/or other test occasions using Full-Information Maximum Likelihood (FIML) estimation (Arbuckle, 1996). Missingness on covariates is also dealt with using FIML.

30 This number also includes some individual twins of whom the co-twin is not in the NCO-LVS-data, leading to incomplete twin pairs. Excluding these incomplete twin pairs did not change the results.



2013/2014	m	r m								
2014/2015	m	r m	r m	r m						
2015/2016	m	r m	r m	r m	r m	r m				
2016/2017	m	r m	r m	r m	r m	r m	r m	r m		
2017/2018	m	r m	r m	r m	r m	r m	r m	r m	r m	r m
2018/2019	m	r m	r m	r m	r m	r m	r m	r m	r m	r m
2019/2020	m		r m		r m		r m		r m	

Figure 5.2 Overview of the tests throughout the school years used in this study.

Note: M = mid-term test, E = end-of-term test, r = reading, m = mathematics.

Potential selectivity in the sample could result from two main sources: selectivity in the NCO-LVS-data and selectivity due to relying on twins. First, the NCO-LVS-data could be somewhat selective at the school and pupil level. Some schools may be more likely to export their data, and there can be selectivity in the preference for using the tests provided by *Cito* rather than one of the other test providers. Also on the pupil level, selectivity could result from parents who did not permit sharing the test data of their children. There could also be missing test data. Exclusion from taking tests is rare but could occur if pupils are absent (e.g., due to illness) or have a very large learning loss. Yet, the number of pupils that are excluded from the tests is generally relatively small (Haelermans, Korthals, et al., 2022). The representativeness of the sample compared to the full population on several school and pupil background variables has been investigated by Haelermans et al. (2021).[31] They find that pupils with a second-generation migration background, pupils who attend larger schools, and schools with a larger percentage of lower-educated parents were overrepresented. However, these differences between the sample and the total population were small and controlling for the selectivity of the sample with inverse probability weighting resulted in only minor differences in the analyses (Haelermans et al., 2021).

Second, relying on twins may induce selectivity. Therefore, I check whether the twin sample differs from the full NCO-LVS sample concerning parental education, parental income, school composition, and school quality. Compared to the full sample, twins have parents with a higher educational level and income and attend schools with a higher SES (see Table E1.1 in Appendix E1). However, although statistically significant, these differences are small. Moreover, the small differences between the twin sample and full sample do not result in differences in (the development of) dispersion (see Appendix E1).

5.3.3 Measurements

Standardized test scores

The measures of pupils' reading and mathematics performance are derived from the LVS of *Cito*. I use the M- and E-tests for Grades 1 to Grade 5. Grade 5 is the penultimate grade of primary school in the Netherlands. I do not use the tests taken in Grade 6, because the timing of the tests is different, and the final test is the Central End-of-primary-school test (*Centrale Eindtoets*) of which the scaling of scores differs. There is no mid-term assessment for reading in Grade 1, therefore, I rely on all the subsequent tests for reading. Figure 5.2 shows an overview of the different test moments by domain used in this study. Almost 5% of twins repeated a grade between grades 1 and 5. In the case of grade repetition, I use the last test score. Twin pairs with an outlier on either the intercept and/or slope (defined as M ± 3 S.D.) are excluded from the analyses. As a robustness check, I also perform the analyses excluding those twin pairs of whom one or both twins repeated a grade between Grade 1 and Grade 5 (see Tables E2.1 and E2.2, Appendix E2),

31 Although the raw data are the same for the present study and the study by Haelermans et al. (2021), the sample is not. Amongst others, the cohort selection differs, with the most important difference being that the study by Haelermans et al. (2021) includes test occasions during the COVID-19 pandemic. Schools were more likely to abstain from administering the tests for a larger share of pupils during the pandemic. The sample of the current study does not include this source of selectivity.

and where outliers were not excluded (see Tables E2.3 and E2.4, Appendix E2). This did not substantially change the results.

The reading tests assess the understanding of written texts, including both factual and literary content. Pupils are presented with a series of texts and have to answer multiple-choice questions. The mathematics tests assess both abstract problems (involving addition, subtraction, multiplication, and division) and applied problems that describe a concrete task (e.g., involving time or money). All tests are shown to have high validity and reliability with reliability coefficients around .90 (Hop et al., 2016; Tomesen et al., 2018).[32] The test provider *Cito* used Item Response Theory (IRT) to relate the answers of pupils on the test to a score on a continuous scale that reflects the underlying latent performance level. IRT weighs which questions are answered (in) correctly to come to an estimation of pupils' performance level. An advantage of IRT is that comparable scales can be created across different tests, allowing comparison over time and thus a calculation of student growth trajectories in primary school (Scheerens et al., 2012). When the test changes and includes new questions, questions from earlier tests are used as an 'anchor' for the continuous scale.

Ideally, I would have measured initial performance before children enter school. Instead, the first observation of pupils' performance is in Grade 1 when children are around 6 years old. Since this is the first grade where formal teaching starts, it can still be expected that differences in performance capture mostly influences of pupils' (family) background instead of school influences which likely become more important over time.

Parental SES

I measure parental SES by a factor score based on parental education and income. I use the father's and mother's highest attained level of education coded according to the International Standard Classification of Education (ISCED) 2011. For this, I used administrative data from the highest education database. For income, I rely on the father's and mother's percentile scores of their personal yearly income for which I look at the year of twins' starting cohort at school (Grade 1). I constructed a factor score for SES using FIML estimation to handle missing data (see Table E3.1, Appendix E3).[33]

32 To assess the construct validity, *Cito* assessed amongst others the association with other (older generation) LVS-tests, to what extent the test items represent a unidimensional construct, and whether the items are of good quality. Criterium validity is not applicable for the LVS tests, because the tests have no predictive measurement pretension.

33 Educational attainment is sometimes missing. The educational attainment register data are based on diverse registrations of completed education at governmentally funded institutions, combined with data from the Labor Force Survey collected on a sampling basis. The coverage is high, but the data do not include the whole population. Diplomas in higher education are registered from the mid-1980s onwards and there is no (reliable) register data available for education obtained abroad, long-term corporate training, or privately funded education (which is relatively rare in the Netherlands). This may lead to some selectivity, but since I rely on a factor score based on multiple indicators (i.e., father's and mother's income and education) this is less problematic compared to relying on education as the only SES indicator.

School SES

To measure school SES composition, I rely on data of all pupils that attended the school in the period that I study (2013-2019). For each year, I took the average of the parental SES measure. Next, I averaged this over the years.

School quality

I constructed a factor score to measure school quality based on the indicators that the Dutch Inspectorate of Education uses to assess whether schools meet a certain quality standard. These indicators were available for the years 2000-2011, and sometimes until 2019. Generally, schools are visited by the inspectorate every four years and the set of indicators that were used differed over the years. This data structure means that it is impossible to measure school quality per year or even a couple of years. Therefore, for each indicator, the average of all available years is taken. Items were mostly measured on a three-point scale (insufficient, sufficient, good) or a four-point scale (bad, insufficient, sufficient, good). Sometimes, also a two-point scale was used (insufficient, sufficient; no, yes). I conducted Exploratory Factor Analyses (EFA) with Full Information Maximum Likelihood (FIML) estimation in Mplus: one for all the items related to school resources leading to two dimensions, and one for school climate leading to seven dimensions. Altogether, this led to nine dimensions of school quality: (1) range of educational activities, (2) (implementation of) school curriculum, (3) guidance of educational needs, (4) parental involvement, (5) monitoring and evaluating (special needs) students, (6) learning climate, (7) social climate, (8) safety, and (9) quality assurance. Based on these dimensions, I constructed one overall school quality factor in a third (confirmatory) factor analysis.[34] Given the numerous latent variables and items, it was not possible to integrate the full measurement model with the analytical model. Therefore, I saved the factor score and included it in the analytical model as a single variable while imposing a measurement error correction. More information on this correction – as well as further details on the procedure, items, and factor analyses – are provided in Appendix E3.

Control variable

In all models, I control for sex differences (girls = 0, boys = 1) by including it as a covariate with an influence on the intercept and slope. All descriptive statistics are presented in Table 5.1.

34 Three dimensions (parental involvement, social climate, and safety) had a low loading on the overall factor. Therefore, as an alternative measure, I constructed a factor score excluding these dimensions. Using this alternative measure did not lead to substantially different results. Also analyzing the resource and climate dimensions separately, or all the nine quality dimensions separately, did not lead to other conclusions (see Appendix E2).

Table 5.1 Descriptive statistics for the reading (*N*=10,214; N_{pairs}=5,304) and mathematics (*N*=10,616; N_{pairs}=5,576) samples.

	Reading			Mathematics		
Variable	***N***	**Mean**	**SD**	***N***	**Mean**	**SD**
M1	-	-	-	9,505	116.68	32.88
E1	5,084	116.18	28.15	8,169	142.41	29.14
M2	6,755	136.14	27.62	8,142	163.40	29.74
E2	6,327	144.02	26.35	6,769	185.06	29.44
M3	6,704	156.35	26.46	6,709	202.42	28.61
E3	4,029	159.19	25.74	5,406	215.51	28.02
M4	5,402	171.50	24.86	5,409	227.90	27.35
E4	2,945	180.70	26.59	4,058	241.07	26.49
M5	4,089	191.95	26.54	4,113	254.25	27.68
E5	1,682	197.42	25.79	2,559	265.85	27.74
Male	10,214	0.49	-	10,616	0.50	-
School quality	10,179	0.10	0.48	10,579	0.10	0.47
School SES	10,214	0.49	0.27	10,616	0.49	0.27
Parental SES	10,180	0.09	0.86	10,578	0.09	0.85

Note: N = number of individuals, M = mid-term test, E = end-of-year test.

5.3.4 Latent growth model

To examine the dispersion in educational performance over time, latent growth modeling (LGM) is used. A linear latent growth model, presented in Figure 5.3, consists of two latent factors. The first factor, the *intercept*, estimates the average initial educational performance and individual differences in initial performance ($\sigma^2_{intercept}$). The factor loadings on the underlying test scores are fixed to one, given that the intercept is constant over time. All test scores have residual variance, capturing time-specific and measurement error sources of variation across time. The second factor, the *slope*, describes the average learning growth rate and individual differences in the growth rate (σ^2_{slope}). The factor loadings for the slope are fixed at the time scores. The spacing between the different tests is not equal. There are about five months between the M_t-test and E_t-test, and seven months between the E_t-test and M_{t+1}-test. To reflect this, the loadings are fixed with intervals of .4 year and .6 year. The first measurement occasion (i.e., E1 for reading, M1 for mathematics) is fixed at zero. The resulting mean of the slope factor can be interpreted as the change between time score 0 and time score 1 and thus reflects the average growth over one year. The covariation between the intercept and slope is freely estimated and shows whether higher initial performance is associated with stronger learning growth (i.e., positive covariance, increasing dispersion) or weaker learning growth (i.e., negative covariance, decreasing dispersion).

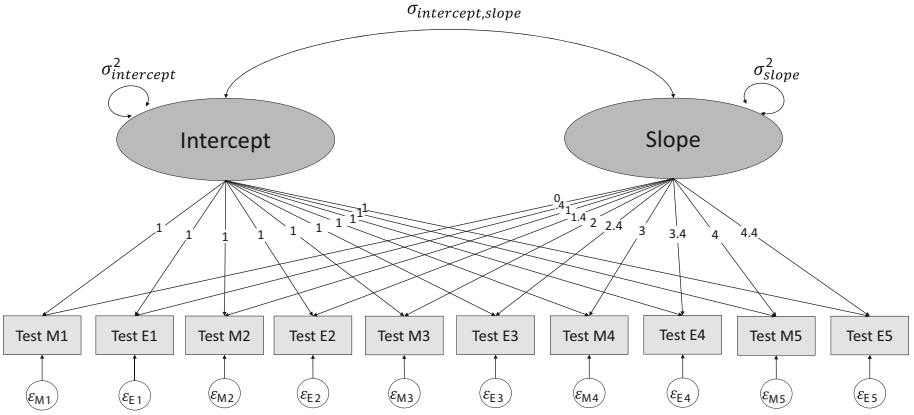


Figure 5.3 Latent growth model.

Note: M = mid-term, E = end-of-year. For reading, there is no M1-test. In this case, the first observation refers to the E1 test with a loading of the slope fixed at 0, and the other loadings fixed with a .6 interval between the E_t and M_{t+1} tests and .4 between the M_t and E_t tests (i.e., 0.6 for M2, 1 for E2, 1.6 for M3, etc.).

5.3.5 Twin model

To investigate the extent to which genetic and environmental differences play a role in initial performance and learning growth, I apply a twin design (see Figure 5.4). The classical twin design is based on the comparison of the resemblance in an outcome between identical (i.e., monozygotic; MZ) and fraternal (i.e., dizygotic; DZ) twins. These twin types differ in their genetic similarity (MZ twins share all of their genes at conception, and DZ twins on average half of their segregating genes) and are assumed not to differ in their shared environmental similarity.[35] Therefore, if MZ twins have more similar educational performance than DZ twins, this indicates that education is genetically influenced. Based on this idea, one can use structural equation modeling (SEM) to decompose the variance in educational performance into three latent components capturing additive genetic variance (*A*), shared environmental variance (*C*), and non-shared environmental variance (*E*) (Knopik et al., 2016).

Since I rely on administrative data, I do not have information on whether twins are MZ or DZ. Therefore, I follow the strategy of previous research (Erola et al., 2021; Figlio et al., 2017) and compare same-sex (SS) and opposite-sex (OS) twins. OS twins are always DZ, hence they share half of their segregating genes. SS twins are a mixture of MZ and DZ twins and therefore their genetic similarity is somewhere between .5 and 1. The exact genetic correlation of SS twins (rSS_G) is unknown, but the mixture of MZ and DZ twins among SS twins is often assumed to be fifty-fifty

35 This is called the Equal Environment Assumption (EEA). Although MZ twins can be expected to grow up in more similar environments and to be treated more similarly (Joseph, 2014), there is not much evidence that this relates to differences in educational outcomes. Studies showed that the EEA is unproblematic for many outcomes (for an overview, see (Felson, 2014)), including educational outcomes (Fletcher & Conley, 2013; Mönkediek, 2021).

(see, e.g.,Marks et al., 2023; Rodgers et al., 1994). Therefore, the average of MZ and DZ twins' genetic relatedness can be used, that is, an rSS_G of .75. As robustness checks (see Appendix E2), I also conduct the analyses with .70 and .80 as alternative values.[36]

In the SEM model, I constrain the genetic correlation (*A1-A2*) to be .75 for SS twins and .5 for OS twins, and the shared environmental correlation (*C1-C2*) to be 1 for both SS and OS twins. The resulting covariance in educational performance between twin 1 and twin 2 is $Cov_{SS} = .75a^2 + c^2$ for SS twins and $Cov_{OS} = 0.5a^2 + c^2$ for OS twins. The latent variance components *A*, *C*, and *E* are set to a variance of 1. Path coefficients *a*, *c*, and *e* represent the effects of the latent factors on educational performance. The variance is equal to the square of the path coefficient; hence the *ACE* model can be written mathematically as:

$$V_{educ} = a^2 + c^2 + e^2 = V_A + V_C + V_E \quad (5.6)$$

For standardized components, each variance component is made proportional to the total variance. For example, relative genetic contribution (i.e., heritability) is obtained by:

$$SV_A = \frac{V_A}{V_{educ}} \quad (5.7)$$

It is important to be aware of what the *ACE* components do and do not capture. The latent component *A* provides an omnibus measure of genetic variance in educational performance, which captures differences in intelligence but also numerous non-cognitive traits that have a genetic component such as self-efficacy and behavioral problems (Krapohl et al., 2014). Similarly, the *C* and *E* components provide omnibus measures of the environment, capturing influences that make twins more similar (e.g., related to the family, school, and neighborhood) and dissimilar (e.g., differential experiences and treatment, different friends, luck, measurement error), respectively. Educational performance has been shown to have a moderate to large heritability and a relatively smaller amount of environmental variance (De Zeeuw et al., 2015). That is not to say that environmental influences are not important. Children with a genetic predisposition towards being good at learning would not learn to read, write, and perform calculations without

36 There are reasons to expect a higher or lower rSS_G. The, rSS_G could be lower, first, because the MZ/DZ ratio among SS twins may not be fifty-fifty. The DZ rate could be higher due to a higher maternal age and usage of assisted reproductive technologies, amongst others (Glasner et al., 2013). Therefore, alternatively, Weinberg's differential rule can be used to estimate the rSS_G based on the number of SS and OS twins in the data (Weinberg, 1901). This leads to an estimated relatedness of $rSS_G = 1*(N_{SS} - N_{OS} / N_{SS}) + 0.5*(N_{OS} / N_{SS}) = 1*(3544 - 2155 / 3544) + 0.5*(2155 / 3544) = .70$. Second, the presence of assortative mating increases the genetic relatedness of DZ twins relative to MZ twins (Loehlin et al., 2009). This implies that the difference in genetic relatedness between SS and OS twins becomes smaller. This smaller difference is captured by using rSS_G of .70 instead of .75, while holding rOS_G constant at .50. Contrarily, I also use a higher value of rSS_G to reflect that SS twins may not only have more similar performance because they are genetically more similar, but also because they are of the same sex. Theoretically, one would adjust shared environmental correlation of OS twins downwards to consider that their environments are less similar because they are of different sexes. However, genetic and shared environmental relatedness account for the same pattern in the data (cf. Spinath et al., 2004). Therefore, I use a higher value of rSS_G, which is practically the same as using a lower value for the shared environmental correlation of OS twins.

being taught so by teachers and parents. If children's reading, writing, and mathematics skills increase over time, as well as the heritability of these skills, this can be seen as an achievement of parents and teachers and should not be seen as a sign of genetic determinism (Asbury & Plomin, 2014). Moreover, the *ACE* components capture *differences* in educational performance. If all children are raised in equally supportive families and attend equally good schools, there would be limited environmental variance in educational performance and most of the differences would be accounted for by genetic differences between children (Asbury & Plomin, 2014). Supportive families and good schools would then increase *average* educational performance but not explain much of the *differences* between children, because these environments are rather universal.

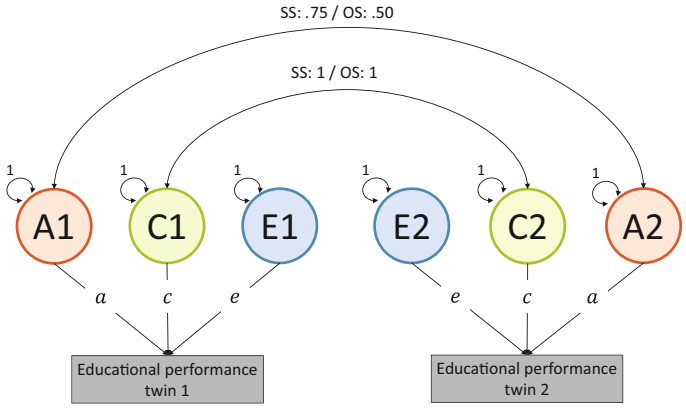


Figure 5.4 The classical twin design applied to same-sex (SS) and opposite-sex (OS) twins decomposing the variance in educational performance into genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) components.

5.3.6 Analytical strategy

I apply the twin design to a latent growth model. This combination has also been referred to as a biometric latent growth model (McArdle et al., 1998). The variance in the intercept and slope, as well as their covariance, is decomposed into the *ACE* components (see Figure 5.5). It could be that the genetic and environmental influences on initial performance (i.e., the intercept) and learning growth (i.e., the slope) are the same, but there may also be genetic and environmental influences that only influence growth. Therefore, so-called Cholesky decomposition is used to decompose two sets of *ACE* components capturing variance related to the intercept (V_{A1}, V_{C1}, V_{E1}) and the slope (V_{A2}, V_{C2}, V_{E2}). The paths *a11*, *c11*, and *e11* reflect the genetic and environmental effects on the intercept, and *a21*, *c21*, and *e21* show to what extent these same influences also predict the slope. The paths *a22*, *c22*, and *e22* show the effects that are independent of the intercept and thus uniquely influence the slope. The total variance in the intercept and slope is given by:

$$V_{intercept} = a11^2 + c11^2 + e11^2 = V_{A1} + V_{C1} + V_{E1} \quad (5.8)$$

$$V_{slope} = \left(a21^2 + a22^2\right) + \left(c21^2 + c22^2\right) + \left(e21^2 + e22^2\right) = V_{A2} + V_{C2} + V_{E2} \quad (5.9)$$

The covariance between the intercept and slope can likewise be decomposed in *ACE* components. This covariance decomposition tells us how much of the covariation/correlation between the intercept and slope is accounted for by shared genetic and environmental effects (de Vries et al., 2021). This is done by calculating the respective covariance over the total covariance, where the total covariance of the intercept *(i)* and slope (*s*) is given by:

$$\begin{aligned} COV_{intercept,\ slope} &= (a11 \times a21) + (c11 \times c21) + (e11 \times e21) \\ &= COV_{Ai,As} + COV_{Ci,Cs} + COV_{Ei,Es} \end{aligned} \quad (5.10)$$

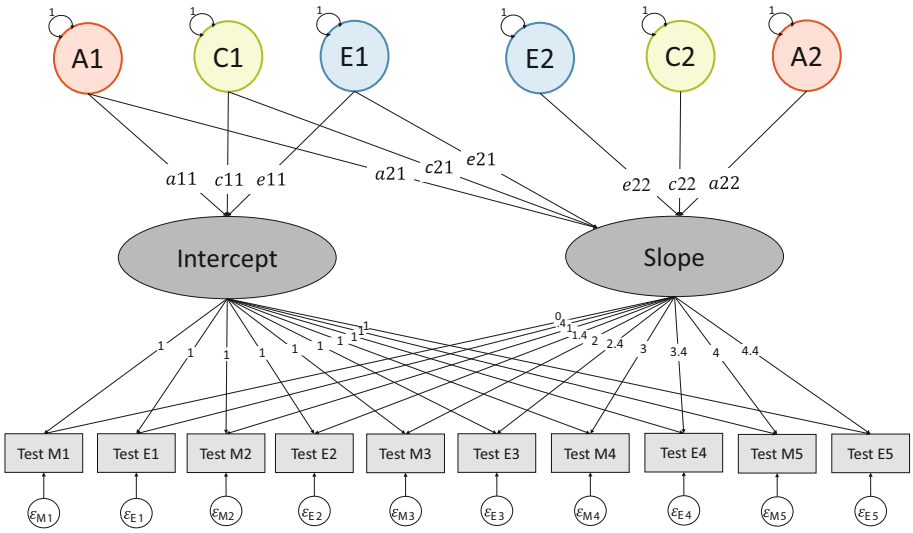


Figure 5.5 Biometric latent growth model decomposing the variance in the intercept, slope, and their covariance, into genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variance.

Note: M = mid-term, E = end-of-year. For reading, there is no M1-test. In this case, the first observation refers to the E1 test with a loading of the slope fixed at 0, and the other loadings fixed with a .6 interval between the E_t and M_{t+1} tests and .4 between the M_t and E_t tests (i.e., 0.6 for M2, 1 for E2, 1.6 for M3, etc.).

Based on the variance in the intercept, variance in the slope, and their covariance, the expected total variance (i.e., dispersion) at timepoint *t* can be calculated using the following formula (McArdle et al., 1998; Muthén & Khoo, 1998; Reynolds et al., 2002):

$$V_{performance|t} = a11^2 + \left(2 \times t \times (a11 \times a21)\right) + (a21 \times t)^2 + (a22 \times t)^2 + c11^2 + \left(2 \times t \times (c11 \times c21)\right) + (c21 \times t)^2 + (c22 \times t)^2 + e11^2 + \left(2 \times t \times (e11 \times e21)\right) + (e21 \times t)^2 + (e22 \times t)^2 \quad (5.11)$$

I split this into two parts to examine the development of initial dispersion in performance (i.e., all paths associated with the intercept, namely paths *11* and *21*) from the development of all the new influences that uniquely contribute to the variance in learning growth (i.e.., all paths that are only associated with the slope, namely paths *a22*, *c22*, and *e22*).

To identify the model, the variances of the intercept and slope are fixed to zero because their variances are now fully explained by the *ACE* components. Furthermore, the means and variances of the intercepts and slopes are constrained to be the same for twin 1 as for twin 2, and the same for SS and OS twins.[37] Also, the residual variances for the underlying test scores were constrained to be equal for both twins and SS and OS twins.[38] The model includes correlations between the residual variances of each test score of twin 1 and twin 2, separately for SS and OS twins. This is because these residuals could include meaningful twin similarity that is unrelated to the latent factors due to time-specific factors that are shared by twins (e.g., lower test results due to family disruptions) (Christopher et al., 2013). Moreover, model comparisons show that including twin correlations between the residuals improved the model fit (see Table E4.1, Appendix E4).

I fit a series of LGMs, which all control for sex differences in the intercept and slope. The standard errors are clustered by the school identifier to take the hierarchical data structure into account. To determine whether learning growth follows a linear pattern, I fit biometric growth models with a linear slope and a quadratic slope. Model comparison using chi-square difference test shows that for both domains the model with a quadratic slope fits the data better (see Table E4.2, Appendix E4). However, the model including the quadratic slope shows that the quadratic slope has a small effect and little variance. Given this, and the fact that nonlinearity complicates the estimation and interpretation of model parameters (Dominicus, 2006), I continue with the linear model in the biometric analyses.

The biometric LGM with a linear slope examines genetic and environmental differences in initial performance, learning growth, and their covariation. The covariation between initial performance and learning growth can be positive or negative, and would indicate if inequality is

37 I used Wald-tests for parameter constraints to test whether the means and variances of the intercept and slope statistically significantly differ between twin 1 and twin 2 and the SS and OS groups. This was not the case, hence, the means and variances across twin 1 and 2 and SS and OS twins could be equalized.

38 Some studies additionally constrain the residual variances equal across time points (see Christopher et al., 2013; Reynolds et al., 2005), but imposing these constraints did not improve model fit (see Appendix E4).

increasing or decreasing, respectively.[39] A negative association between initial performance and growth could appear because of genetic or (shared) environmental effects but could also be a statistical phenomenon caused by regression to the mean (Ready, 2013). This is less likely to form a problem in the current design, as relying on the many time points counteracts this source of bias (Caro et al., 2009). Moreover, measurement error is part of the residual variances or the non-shared environmental variance component in the model, so we can separate this from genetic and shared environmental influences.

Lastly, I will more specifically investigate the contribution of schools by including school quality and school SES as covariates in the model, with an influence on the intercept and the slope. I also include parental SES to take socioeconomic selection in schools into account (Borghans et al., 2015a; Robert, 2010). Both twins are raised in the same family and attend the same school. That means that the school and family characteristics are part of the shared environment. To the degree that these measures account for shared environmental variance, their inclusion reduces the relative contribution of the residual shared environment and increases the relative contribution of genes (Christopher et al., 2013).

I perform several supplementary analyses, which are presented in Appendix E2. The results show that the conclusions are robust to alternative model specifications.

5.4 Results

5.4.1 Differences in initial educational performance

Figure 5.6 shows the total dispersion in reading and mathematics performance over the primary school period, including the underlying genetic and environmental sources of dispersion. When formal education starts in Grade 1, there is already considerable variation in educational performance. The total variance in the intercept is 512.82 for reading and 694.21 for mathematics (see Table 5.2). As can be seen in Figure 5.6, and more detailed in Table 5.2, genetic variance is the major source contributing to differences in initial educational performance. For reading, genetic differences account for 84% of the differences in initial performance in Grade 1. For initial mathematics performance, this is somewhat lower but still high, namely 70%. When it comes to the role of the environment, the shared environment is almost as important as the non-shared environment. The shared and non-shared environment account for, respectively, 8% and 7% of the variance in initial reading performance (but note that the shared environmental variance in reading is not statistically significant, see Table 5.2). For mathematics, both the shared and non-shared environment account for 15% of the variance in initial performance.

39 This shows the development of different sources of dispersion over the school career – which is the focus of this paper – but does not provide insight in the more complex underlying mechanisms. For example, the parental resources that gave high-SES children an advantage at the start of education could continue to have an influence on later performance (i.e., *ACE* affect both initial performance and learning growth). Alternatively, it could also be that high-SES children start with an advantage and perform higher in Grade 1, and that these higher initial performance levels affect their learning growth (i.e., the influence of *ACE* on learning growth is mediated by initial performance).

5.4.2 Development of differences in educational performance

Children improve their educational performance over time; there is on average statistically significant learning growth for reading (b = 18.36, p<.001) and mathematics (b = 31.78, p<.001). These learning growth rates refer to the average improvement in reading and mathematics test scores per year. Not all children improve at the same pace, which is captured by the statistically significant variance of the slopes for reading and mathematics (respectively 12.32 and 21.58, see Table 5.2). The presence of variance in the learning slope implies that the total dispersion in educational performance changes during the primary school career. This change results from two developments: (i) the change in initial differences, which depending on whether these differences increase or decrease would lead to more or less dispersion, and (ii) new influences that come into play over time leading to more dispersion.

The differences in initial educational performance decrease over time (see Figure 5.6). This is reflected by the negative covariance between initial performance and growth, shown in Table 5.2. The intercept and slope are negatively and statistically significantly associated for reading ($COV_{i,s}$ = -14.08, p<.001) and mathematics ($COV_{i,s}$ = -38.81, p<.001). Transforming these covariances into correlations show that the association between the intercept and slope is relatively weak for reading (r = -.18, p<.001), but more substantial for mathematics (r = -.32, p<.001). These negative associations indicate a compensatory growth pattern where those children who initially have a lower test score grow faster than those who initially performed higher (and vice versa). This is in line with hypothesis 1b: initial dispersion in educational performance decreases over the primary school career.

While the total amount of differences in initial performance decreases during schooling, Figure 5.6 shows that it is not the case that all the underlying sources of dispersion decrease. The initial differences in *reading* performance decrease over time due to a decrease in the initial genetic differences. This is reflected in the statistically significant genetic covariance as shown in Table 5.2 ($COV_{Ai,As}$ = -13.42, p=.040). There is neither a statistically significant change in initial shared environmental differences ($COV_{Ci,Cs}$= 2.53, p = .496) nor in non-shared environmental differences ($COV_{Ei,Es}$= -3.19, p=.327) in reading performance. The initial differences in *mathematics* performance decrease over time due to a decrease in initial shared environmental differences ($COV_{Ci,Cs}$= -26.97, p=.012) and non-shared environmental differences ($COV_{Ei,Es}$= -17.57, p=.001). Here, the change in initial genetic differences is not statistically significant ($COV_{Ai,As}$= 5.73, p=.713). The decreasing shared environmental differences likely reflect the reducing impact of family background, given that the initial sources of performance differences at the start of schooling are assumed to be largely related to the home environment (given the limited exposure to the school environment at that moment). These family background differences decrease and approach zero near the end of primary education (see Figure 5.6).

In addition to the changes in pre-existing sources of dispersion in educational performance, there are also new sources of dispersion coming into play. For both reading and mathematics performance, these new sources mainly include new genetic influences, although there are also new non-shared environment influences (see Figure 5.6 and Table 5.2). For mathematics performance, the new genetic variance (9.15) and non-shared environmental variance (2.12) do not reach statistical significance at the five percent level (see Table 5.2). This is probably a power

issue, since the underlying path estimates are statistically significant (see Table E5.1, Appendix E5). The new influences increase dispersion in performance by definition. Combining the (decreasing) dispersion in initial performance with the dispersion that is coming into play during the school period (and thus is unrelated to the initial situation), leads to an increase in the overall dispersion in educational performance (see Figure 5.6). This means that at the beginning of formal education (Grade 1) children differ in their performance levels but near the end of formal education (Grade 5) they differ even more. The new genetic and non-shared environmental influences that are coming into play outweigh the decrease in initial genetic differences (for reading) and initial shared and non-shared environmental differences (for mathematics).

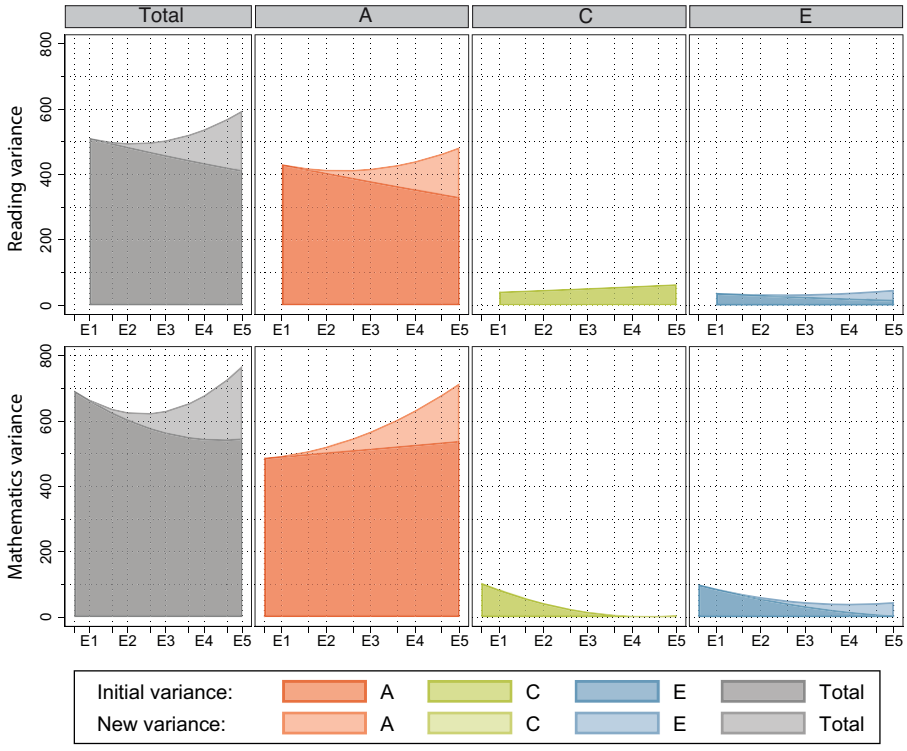


Figure 5.6 The development of the total, genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variance in reading and mathematics performance.

Note: Controlled for sex. Based on the variance in the intercept and slope and their covariance, reported in Table 5.2.

Table 5.2 Estimates of the *A, C, E* variance and covariance components of the intercept and slope for reading (N_{pairs}=5,304) and mathematics (N_{pairs}=5,576).

	Reading				Mathematics			
	Total	**A**	**C**	**E**	**Total**	**A**	**C**	**E**
Intercept mean	125.57 ***				125.93 ***			
Slope mean	18.36 ***				31.78 ***			
Unstandardized								
Intercept variance	512.82 ***	432.27 ***	42.41	38.14 **	694.21 ***	488.25 ***	105.24 *	100.73 ***
Slope variance	12.32 ***	9.97 ***	0.15	2.20 *	21.58 ***	9.21	7.19 *	5.18 **
common with intercept	0.84	0.42	0.15	0.27	10.05	0.07	6.91	3.06 *
unique to slope	11.49 ***	9.55 ***	0.00	1.93 *	11.54	9.15	0.27	2.12
$COV_{I,S}$	-14.08 ***	-13.42 *	2.53	-3.19	-38.81 ***	5.73	-26.97 *	-17.57 ***
Standardized								
Intercept variance		0.84 ***	0.08	0.07 **		0.70 ***	0.15 *	0.15 ***
Slope variance		0.81 ***	0.01	0.18 *		0.43	0.33 *	0.24 **
common with intercept		0.03	0.01	0.02		0.00	0.32	0.14 *
unique to slope		0.78 ***	0.00	0.16 *		0.42	0.01	0.10
$COV_{I,S}$		0.70 *	0.13	0.17		0.11	0.54 *	0.35 ***

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Controlled for sex. *A* = genetic, *C* = shared environment, *E* = non-shared environment. For all the underlying path estimates and standard errors, see Table E1 in Appendix E.

Altogether, I find that initial dispersion in reading and mathematics performance at the beginning of primary education decreases during the primary school career. This is due to a decrease in genetic differences for reading, and a decrease in shared and non-shared environmental differences for mathematics. However, the total dispersion in reading and mathematics performance is increasing over time. This is because new influences are coming into play during the school career, mainly new genetic influences but also some non-shared environmental influences. The increase due to the new influences is stronger than the decrease of the initial differences, leading to an overall increase of dispersion in performance over the primary school career. Although the focus is on dispersion (i.e., unstandardized variance components), I also investigated how the relative contributions of genetic, shared environmental, and non-shared environmental factors develop over the school career (see Figure E5.2, Appendix E5). These largely reflect the patterns of the absolute variances. Additionally, it shows that the relative contributions of genes, the shared environment, and non-shared environment to the total amount of variance in reading performance do not change in a statistically significant way. For mathematics performance, the relative contribution of genes increases while the relative contribution of both sources of environmental variance decreases.

5.4.3 Role of the school environment

The finding that initial dispersion decreases during the schooling period, and that there is shared environmental variance in learning growth, suggests that the school environment could play a role. I include measured characteristics of the school environment – school quality and school SES – to further investigate this. Given socioeconomic selection in schools (Borghans et al., 2015a; Robert, 2010), I also include parental SES. The results are presented in Table 5.3.

As initial performance mainly reflects influences before the start of schooling, I do not expect that school characteristics affect the intercept. However, the results show that school quality is positively associated with the mathematics intercept (β = .05, p = .008) and that school SES is positively associated with the intercepts of both reading performance (β = .04, p = .016) and mathematics performance (β = .04, p = .015). Children who attend higher-SES and higher-quality schools thus have on average higher initial educational performance. Although statistically significant, the effects are small. One S.D. increase in school SES is associated with only 0.04 S.D. increase in initial reading and mathematics performance. For the association between school quality and initial mathematics performance, this is only 0.05. Parental SES is more substantially related to initial performance (β = .31, p<.001 for reading; β = .27, p<.001 for mathematics). Comparing the total variance in the intercept between Model 2 and Model 1 shows that the inclusion of all the measured characteristics explains 10% and 8% of the total variance in the reading and mathematics intercepts, respectively.

In contrast to the expectation, I do not find evidence for a positive association between school quality and the slope for either reading or mathematics. For school SES, there is a statistically significant positive association with the reading slope. Children in higher-SES schools have higher learning growth rates in reading (β = .11, p < .001), but this effect is rather weak. Parental SES is a statistically significant, yet also rather weak, predictor of both reading (β = .11, p < .001) and mathematics growth (β = .06, p = .023). Although some of the measured environmental

characteristics are thus significant predictors of average learning growth, their effects are small in magnitude and their impact on explaining differences in learning growth is limited. Of the total variance in the reading and mathematics slope, only 4% and 1% of the variance can be explained. For the development of reading performance, this means that the small amount of shared environmental variance that was present can entirely be explained by the measures (see Table 5.3, and Figure E5.1 in Appendix E5). For mathematics performance, the amount of unexplained shared environmental variance is hardly affected by the inclusion of the school and family measures.

Table 5.3 Estimates of the biometric growth models including covariates for reading (N_{pairs}=5,304) and mathematics (N_{pairs}=5,576).

	Reading		Mathematics	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Mean				
Intercept mean	125.57 ***	123.26 ***	125.93 ***	123.09 ***
$\beta_{school\ quality}$		0.02		0.05 **
$\beta_{school\ SES}$		0.04 *		0.04 *
$\beta_{parental\ SES}$		0.31 ***		0.27 ***
Slope mean	18.36 ***	17.61 ***	31.78 ***	31.46 ***
$\beta_{school\ quality}$		0.04		-0.04
$\beta_{school\ SES}$		0.11 ***		0.04
$\beta_{parental\ SES}$		0.11 ***		0.06 *
Variance				
Intercept variance	512.82 ***	459.08 ***	694.21 ***	636.98 ***
SV_A	0.84 ***	0.91 ***	0.70 ***	0.71 ***
SV_C	0.08	0.00	0.15 *	0.12
SV_E	0.07 *	0.09 ***	0.15 ***	0.18 ***
Slope variance	12.32 ***	11.85 ***	21.58 ***	21.33 ***
SV_A	0.81 ***	0.80 ***	0.43	0.40
SV_C	0.01	0.00	0.33 *	0.35 **
SV_E	0.18 *	0.20 *	0.24 *	0.25 **

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Controlled for sex. SV_A, SV_C, SV_E refer to standardized genetic, shared environmental, and non-shared environmental variance, respectively. Estimates of school quality, school SES, and parental SES are standardized regression coefficients (beta's) on the intercept and slope.

5.5 Conclusion and discussion

In this study, I aimed to provide more insight into the development of dispersion in educational performance throughout primary education by taking a behavioral genetics approach. By investigating the learning growth of twins identified in the Netherlands Cohort Study of Education, I studied if the pre-existing genetic and environmental differences between children are reproduced, exacerbated, or compensated during the schooling period. I also showed to what extent new sources of dispersion are coming into play, and what this means for the total dispersion in performance. Moreover, by investigating the underlying sources of differences more insight could be provided into the role of the school environment. This led to three main findings.

First, the dispersion in initial educational performance at the start of formal education becomes smaller over time, consistent with a compensatory growth pattern. Dispersion in performance at the beginning of primary education can be largely attributed to genetic differences between children, but environmental differences are also present. For reading, these genetic differences decrease during the school career, leading to the decrease in the dispersion of initial reading performance. For mathematics, the decrease in the dispersion of initial performance is driven by a decrease in shared and non-shared environmental differences. Genetic differences could become smaller because children differ in the rate at which cognitive skills develop, so those who initially have a lower performance may catch up over time (Francis et al., 1996). Another potential mechanism is that those who perform low at the beginning of primary school (e.g., because of lower innate ability or because their genetic potential is not fully realized in the period before Grade 1), improve faster because they benefit more from their school environment or receive more attention. This results in gene-environment correlations that are captured in the genetic component. The decreasing environmental differences in mathematics may be partly driven by low-SES children who have lower initial performance but improve faster during school than high-SES children. Low-SES children may benefit more from school inputs (e.g., resources, academic climate, higher levels of motivation and aspiration) because they are less likely to find these at home (Coleman et al., 1966; Rumberger & Palardy, 2005).

Second, new influences are coming into play over time. If these new sources of dispersion are combined with the development of the initial dispersion in performance, I find that the total dispersion in educational performance increases. The new sources of dispersion are mainly new genetic influences, but to a small extent also new non-shared environmental influences. The increase in genetic differences during schooling due to new genetic influences could be understood from the processes of innovation (Briley & Tucker-Drob, 2013). For example, exposure to new learning activities and peer groups may lead to the activation of genes. However, although innovation occurs during the schooling period, it does not necessarily reflect the influence of exposure to the school environment. It could be related to the family environment as well, for example if changes in the family environment (e.g., parental divorce, move to a new place) lead to the expression of new genetic influences. Also, genetically influenced characteristics that become relevant for later performance (e.g., intellectual interest) could be expressed via processes in the family (Tucker-Drob & Harden, 2012b). Lastly, it is possible that the new genetic influences are

related to neither the school nor family environment, but come into play as a function of age (Briley & Tucker-Drob, 2013)

Third, the results suggest that the role of schools in the development of dispersion is likely limited. If all children would perform better and learn faster in certain schools, this would be captured in the contribution of the shared environment. The role of the shared environment is more important in explaining differences in learning growth than in explaining differences in initial performance for mathematics. However, the shared environmental factors affecting learning growth in mathematics are almost entirely common with those affecting initial mathematics performance. These shared environmental influences are probably more related to the non-school environment than the school environment, given the limited exposure to school before the first measurement occasion in Grade 1. Including measures of the school environment also suggest that the role of schools in (compensatory) learning growth is likely limited. School quality and school SES are only weakly associated with average learning growth and barely account for the (shared environmental) variance in learning growth.

Interestingly, there are differences between educational domains. Genetic differences in initial performance play a larger role for reading than mathematics. Language skills are for an important part acquired before children enter school, while mathematic skills are typically learned in school (Passaretta & Skopek, 2021). Hence, it can be expected that initial performance differences in the language domain are more strongly affected by children's family environment than mathematics. It may be counterintuitive that the component reflecting family influences – the shared environment component – is relatively small for reading and that genetic differences play a large role. This finding can be understood from transactional models of development (Sameroff, 2009). Genetic differences between children are translated to genetic differences in educational performance via environmentally mediated processes (Freese, 2008). Based on children's genetic predisposition towards reading, children may seek out literacy activities or evoke specific responses from others (e.g., provided with books by their parents) (Knigge et al., 2022; Pennington et al., 2009). The family environment may facilitate or constrain such processes (Ruks, 2022). Given the greater importance of the (early) family environment for the development of reading skills than mathematics skills, this may explain the greater realization of genetic endowment for reading. The idea that reading is more strongly the domain of the family and mathematics more of the school, is also supported by the findings on the associations between initial performance and learning growth. For mathematics the negative association is stronger than for reading, indicating that initial dispersion in mathematics is more strongly reduced during the primary school career than initial dispersion in reading.

While the results suggest that schools' influence in explaining compensatory learning growth is probably limited, it could also be related to the limitations of this study concerning the measurement of the school environment. It could be that schools play a role in a way that I cannot capture. First, the school measures reflect between-school effects, while the within-school (classroom) environment may be particularly important for educational performance and growth. For example, pupils who are taught by more effective teachers learn faster (Burgess et al., 2022), but school-level indicators do not provide enough information on this. Second, the school and classroom environment need to be stable and systematically related to learning growth to

explain (shared environmental) variance in growth. It could be that these environments are not so systematically related to growth, but they could be still of importance for performance at specific time points. For example, if children get assigned different teachers each year and these teachers vary in quality, teacher effectiveness would not be observed as strong predictor of learning growth because its influence is not systematic over the school career. But in the grade(s) where children are taught by an effective teacher, their performance may still be higher than if they were assigned a less effective teacher. Third, initial educational performance is not measured at the start of primary education (age 4) but two years later. Hence, it could be that initial performance is already affected by the school environment (i.e., the kindergarten period). In the Netherlands, children participate in kindergarten from approximately age 4 to 6, which is part of the publicly funded primary school system (OECD, 2016). Although formal learning of reading and mathematics starts from age 6 onwards, there is gradual preparation for this during kindergarten via play, structure, and independence, amongst others. The school environment influences this process, and these school factors may continue to exert an influence on performance in the later grades. Thus, the overlap in shared environmental factors contributing to both initial performance and growth does not necessarily arise from the non-school environment.

To gain more insight into the role of schools, future studies could use different measures and methods. For example, by using more specific measures on the classroom level, such as classroom composition, instructional activities, and teacher effectiveness. Since children typically get a different teacher each year, ideally longitudinal data on teachers' effectiveness and activities is needed, which is difficult and costly to obtain. Alternatively, classroom influences could be identified as a latent factor capturing all (unmeasured) classroom influences affecting performance. This can be estimated by comparing twins who are in the same classroom with twins who are in different classrooms (Byrne et al., 2010; Grasby et al., 2020; Stienstra et al., 2023). Unfortunately, a classroom identifier is not available in the NCO-LVS-data. Lastly, future research could broaden the scope and investigate what dispersion looks like if there is no schooling. It could be that differences in the school environment are indeed of limited importance in explaining dispersion in learning, especially in a country as the Netherlands where nearly all primary schools are governmentally funded, and schools attended by pupils from more disadvantaged backgrounds receive additional funding (Ritzen et al., 1997). But even if school differences are small, the fact that children are exposed to a rather universal school environment may still be of great importance for the development of (genetic and environmental) dispersion in educational performance. Combining twin data with a seasonal comparison design or with a design incorporating the school closure due to the COVID-19 pandemic would provide a novel way to gain more insight into the extent to which inequality as dispersion – and the underlying genetic and environmental sources – is dependent on the school environment.

Despite the issues related to the measurement of school quality, the observed decreases of dispersion in performance are unlikely to be explained by differences between schools. Individual differences in learning growth are much smaller than individual differences in initial performance. The finding that differences in learning growth are for a substantial part driven by shared environmental differences, as I find for mathematics, does not have large practical implications if the magnitude of the effect is small (Christopher et al., 2013). Hence, there is more

to gain by altering the initial situation (i.e., prior Grade 1), than the situation during schooling. Given that it is more difficult to intervene in the family environment, one way to do this is via longer duration and/or higher quality early formal childhood education and care (i.e., preschool) for low-SES children than high-SES children. Preschool attendance could foster children's skills, especially for low-SES children (Ghirardi et al., 2023). Hence, preschools could reduce shared environmental differences in initial educational performance (Tucker-Drob, 2012). Given my finding that shared environmental influences on initial performance are common with shared environmental influences on growth, expanding (high-quality) preschool access could have lasting consequences as it would improve the initial situation but possible also later performance via improved learning growth. Moreover, preschools may affect dispersion in initial educational performance via gene-environment transactions. For example, talented children are recognized as such and consequently receive an enriched environment to maximize their talents (Diewald et al., 2015). It can be expected that this occurs less effectively in lower-SES families. However, the preschool environment may substitute for this, especially if low-SES children receive more (higher quality) early childhood education and care.

Dispersion in educational performance is not necessarily problematic because society requires differently skilled individuals (Strello et al., 2021). However, dispersion is more difficult to defend if it is related to the family environment, or if it is produced by schools leaving children with lower performance or ability behind. This research showed that family influences as captured by shared environmental differences are reproduced for reading and compensated for mathematics. Yet, shared environmental differences contribute relatively little to the dispersion in performance. The lion's share of dispersion in performance during primary education in the Netherlands is related to genetic differences and already present at the beginning of formal education in Grade 1. This likely captures more complex relations between children's genetic endowment and their (family) environment. These genetic differences are also compensated for reading and reproduced for mathematics, leading to a decrease of dispersion in initial performance over the primary school career. New genetic differences are expressed during schooling. When these new influences are combined with the initial dispersion in performance, the total dispersion increases during schooling. Whether the decrease of dispersion in initial performance and the new genetic influences that are coming into play during schooling are produced by the school or non-school environment remains a question for future research.

APPENDICES

A. Supplements to Chapter 1

A.1. Classical Twin Design

The Classical Twin Design (CTD) provides insight into how much of the variance in an outcome is accounted for by genetic and environmental variance. Two sources of environmental variance can be distinguished: shared environmental variance (*C*) and non-shared environmental variance (*E*). *C* captures all non-genetic influences that are shared between twins and contribute to twin similarity and is thus a between-family component. *E* captures all non-genetic influences that contribute to twin dissimilarity and is a within-family component. For the genetic variance, *A*, it is a bit more complex. As monozygotic (MZ) twins are genetically identical, there is no difference in genetic variance within twin pairs but only between MZ twin pairs. Dizygotic (DZ) twins share on average half of their segregating genes, hence, among DZ twins A reflects both within and between twin pair variances.

Comparing the similarity in educational performance between MZ twin pairs and DZ twin pairs allows for the estimation of the *ACE* components. The similarity (i.e., covariance or correlation) in the educational performance of MZ and DZ twins reared together can be estimated using the following formulas (Neale & Maes, 2004):

$$Cov_{MZ} = a^2 + c^2 = V_A + V_C \quad (A.1)$$

$$Cov_{DZ} = .5a^2 + c^2 = .5V_A + V_C \quad (A.2)$$

Dividing the covariances by the total variance V_{total}, which is the sum of the variance components ($V_{total} = V_A + V_C + V_E$), gives the twin correlations. The separate variance components can be calculated by:

$$V_A = 2 \times (Cov_{MZ} - Cov_{DZ}) = (Cov_{MZ} - Cov_{DZ})/.5 \quad (A.3)$$

$$V_C = (2 \times Cov_{DZ}) - Cov_{MZ} = Cov_{DZ} - .5V_A \quad (A.4)$$

$$V_E = V_{total} - Cov_{MZ} \quad (A.5)$$

The expected variances and covariances of educational performance within MZ and DZ twins can also be written as the following two model-implied covariance matrices, $\Sigma(\theta)$:

$$\Sigma_{MZ}(\theta) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} V_A + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} V_C + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} V_E \quad (A.6)$$

$$\Sigma_{DZ}(\theta) = \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix} V_A + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} V_C + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} V_E \quad (A.7)$$

Each matrix shows the relation between twin pairs for the stated variance components. The off-diagonal of the genetic relatedness matrix reflects the genetic relatedness of MZ twins (1) and DZ twins (.5) (Hunter et al., 2021).

The variance components can be standardized (i.e., standardized variance; *SV*). For example, the proportion of the genetic variance component to the total variance in educational performance, which is called heritability, is given by:

$$SV_A = \frac{a^2}{a^2 + c^2 + e^2} = \frac{V_A}{V_{total}} \quad \text{(A.8)}$$

Multigroup structural equation modeling (SEM) is used to decompose the variance into the *ACE* components (Rijsdijk & Sham, 2002). The most basic model includes two groups (i.e., MZ and DZ twins), one observed outcome variable for twin 1 and twin 2, and the latent components *A*, *C*, and *E* for both variables (see also Figure 1.2). For both zygosity groups and both twins within a pair, means and variances are constrained to be equal (note that this is an assumption that has to and can be tested). The latent components *A*, *C*, and *E* are set to a variance of 1. The correlation between *C1* and *C2* is set to 1 for both zygosity groups, reflecting the equal environment assumption. The correlation between *A1* and *A*2 differs between the groups, reflecting the difference in genetic relatedness. It is set at 1 for MZ twins, and .5 for DZ twins.

The identification and interpretation of the *ACE* components relies on several assumptions, which are simplifications that are needed for modeling. These assumptions are often testable, and prior research as well as my own robustness checks show that in general the underlying assumptions are met (see also Table 1.1).

A.2. CTD when zygosity is unknown

In most chapters of this dissertation, as well as in other studies (Baier et al., 2022; Erola et al., 2021; Figlio et al., 2017; Pokropek & Sikora, 2015; Scarr-Salapatek, 1971), the CTD is fitted by comparing the similarity in an outcome among same-sex (SS) twins with that of opposite-sex (OS) twins. While the genetic relatedness of OS twins is known (i.e., .50 because they are DZ by definition), the genetic relatedness of SS twins (rSS_G) is unknown. Because SS twins are a combination of MZ and DZ twins, the average rSS_G will fall within the range .5 to 1. The estimated rSS_G (and the rOS_G of .5) can be incorporated in the CTD similar to the genetic relatedness of MZ and DZ twins. That is, constraining the correlation between *A1* and *A2* to be rSS_G for SS twins, and .5 for OS twins. There are commonly two ways to calculate the rSS_G.

1. Assume that the average genetic relatedness of SS twins is .75

This is based on the assumption that among SS twins, half will be MZ and half will be DZ. Given this fifty-fifty mixture, the average of the genetic relatedness of MZ and DZ twins can be used, leading to an rSS_G of .75 (Marks et al., 2023; Rodgers et al., 1994).

2. Estimate the average genetic relatedness of SS twins

Using Weinberg's differential rule, the proportion of MZ and DZ twins within the group of SS twins can be calculated (Weinberg, 1901). According to this rule, the probability of male births equals the probability of female births. Therefore, among DZ twins the number of SS equals the number of OS twins. The total number of DZ twins is thus twice the number of OS twins. The proportion of MZ and DZ twins within SS pairs in the data can be estimated by:

$$p_{MZSS} = (N_{ss} - N_{OS})/N_{SS} \quad \text{(A.9)}$$

$$p_{DZSS} = 1 - p_{MZSS}. \quad \text{(A.10)}$$

Accordingly, the average genetic relatedness among SS twins is:

$$rSS_G = \left(1 \times p_{MZSS}\right) + \left(.5 \times p_{DZSS}\right). \quad \text{(A.11)}$$

Assumptions and potential limitations

The application of the CTD to SS and OS twins rely on the same assumptions as the CTD applied to MZ and DZ twins (see Table 1.1 for an overview). In addition to these standard assumptions, the CTD applied to SS and OS twin data relies on the assumption that SS twins are only more similar than OS twins because SS twins are on average genetically more similar, not for other non-genetic reasons (Figlio et al., 2017). If sex differences play a role, this assumption is violated. SS DZ twins will then be more similar to one another than OS DZ twins, because the first are from the same sex and the latter are not. This would result in an upward bias in estimates of genetic variance and a downward bias in shared environmental variance. The presence of sex influence can be tested (see Chapter 3) and one can correct for a possible bias by using a larger value for

rSS_G.[40] All analyses based on SS and OS data (Chapter 3, 4 and 5) are therefore conducted using three values of rSS_G (i.e., .70, .75, and .80). The conclusions are robust to the different values.

Second, when investigating the interaction between the *ACE* components and an environmental moderator, it is assumed that the rSS_G is the same across all values of the moderator. However, the MZ/DZ birth ratio may depend on parental SES. While MZ twin births are thought to be largely the result of a random event, DZ twin births are more common when assisted reproductive technologies are used, when the mother is older, taller, has a higher BMI, and smokes, amongst others (Glasner et al., 2013). These factors may differ by SES, hence, the estimated rSS_G may also differ by SES. If this is the case, a gene-SES interaction may then result from the bias resulting from assuming the same rSS_G for all SES groups. This did not appear to be the case (see Chapter 3).

Third, one should be aware that misspecification of rSS_G leads to biased estimates. When zygosity is known, the A is calculated as $2 \times (Cov_{MZ} - Cov_{DZ})$, or, to put differently, $(Cov_{MZ} - Cov_{DZ})/.50$, where .50 is the difference in genetic relatedness (Δr_G) of MZ and DZ twins. The difference in genetic relatedness between twin types becomes smaller when SS and OS twins are compared (Δr_G of .30, .25, or .20, depending on whether rSS_G of .70, .75, or .80 is used, respectively). The smaller Δr_G, the larger the estimated *A* and the smaller the estimated *C*. This is no problem if the Δr_G reflect real differences, for example, if a smaller Δr_G results from the larger share of DZ twins among SS twins and/or the presence of assortative mating. However, the *A* and *C* will be over or underestimated if the assumed Δr_G is larger or smaller than the true difference. We investigate the robustness of the results by modelling different plausible values of rSS_G in the range of .70 to .80 (see also Figlio et al., 2017).

Lastly, applying the CTD to data with unknown zygosity is less powerful compared to when zygosity is known. Generally, the more distinct the genetic and environmental relatedness of the relatives that are sampled, the greater the power to distinguish the *A, C*, and *E* (Medland & Hatemi, 2009). If Δr_G is relatively large, as is the case when comparing MZ and DZ twins, there is less information required to distinguish the covariance structure of MZ twins with that of DZ twins. This is because the only difference in the implied covariance structure is the genetic correlation. When Δr_G becomes smaller (i.e., when relying on SS/OS data), the covariance structures become more similar for the two groups. In order to fit the implied covariance matrixes more precisely to the covariance matrices generated from the empirical data, larger sample sizes (and thereby smaller estimated standard errors) are needed (Lyu & Garrison, 2023).[41] Thus, when using SS/OS twin data, greater sample sizes are needed to obtain the same level of power as a CTD applied to MZ/DZ data. Given the usage of population data for the models with unknown zygosity in this dissertation, the sample sizes are much larger than those of twin samples with known zygosity. This counteracts the reduced power resulting from unknown zygosity.

40 Theoretically, one would adjust the shared environmental correlation of OS twins downwards to consider that their environments are less similar because they are of different sexes. However, genetic and shared environmental relatedness account for the same pattern in the data (cf. Spinath et al. 2004), so it makes sense to only adjust one at a time. Practically, increasing rSS_G is similar to decreasing the shared environmental correlation of OS twins.

41 See Lyu and Garrison (2023) for more information – including a mathematical derivation – on the relation between power, sample size, and genetic relatedness.

B. Supplements to Chapter 2

B.1. Operationalization parental educational attainment

Table B1.1. Coding of level of parental education (twin age 10, 7, 3 or 1) into ISLED scores.

Twin age 10 and 7	Twin age 3 and 1	ISLED	Label
(1) Primary school	(1) Lower education	22.98	Primary school or first stage of basic education
(2) Few years 'mavo' or 'mulo' (lower general secondary education)		22.98	Primary school or first stage of basic education
(4) Few years 'LTS' (lower technical education)		22.98	Primary school or first stage of basic education
(6) Few years 'havo/VWO' (higher general secondary education/pre-university education)		22.98	Primary school or first stage of basic education
(5) Finished 'LTS' (lower technical education)		29.34	Lower secondary school, technical training [lts]
	(2) Lower secondary school, lower general secondary education	37.31[a]	Lower secondary school, technical training [lts] or lower secondary school, theoretical training [mulo, mavo]
(3) Finished 'mavo' or 'mulo' (lower general secondary education)		45.27	Lower secondary school, theoretical training [mulo, mavo]
(8) Few years 'MBO' (lower vocational education)		45.27	Lower secondary school, theoretical training [mulo, mavo]
(9) Finished 'MBO' (lower vocational education)		52.70	Upper secondary professional education [MBO]
	(3) vocational education 'MBO', general secondary education 'havo, VWO'	62.31[b]	Upper secondary professional education [MBO], higher secondary school [havo], or pre-scientific secondary school [VWO]
(7) Finished 'havo/VWO' (higher general secondary education/pre-university education)		67.11[c]	Higher secondary school [havo] or pre-scientific secondary school [VWO]
(10) Few years 'HBO' (university of applied sciences) or university		67.11[c]	Higher secondary school [havo] or pre-scientific secondary school [VWO]
(11) Finished 'HBO' (university of applied sciences)	(4) higher vocational education 'HBO'	77.93	Tertiary professional education [HBO]
(12) Finished university	(5) scientific education	87.13	Tertiary scientific education [WO]
(13) Post-doctoral education		92.63[d]	Post-doctoral education and second stage of tertiary education, PhD

[a] Average of 29.34 (lower secondary school) and 45.27 (lower general secondary education).
[b] Average of upper secondary professional education (52.70), higher secondary school (62.30), and pre-scientific secondary school (71.92).
[c] Average of 62.30 (higher secondary school) and 71.92 (pre-scientific secondary school).
[d] Average of 90.63 (post-doctoral education) and 94.62 (second stage of tertiary education, PhD).

B.2. Robustness checks different operationalization classroom allocation

As a robustness check, we used a stricter definition of classroom allocation where we also excluded twins who went to different schools and those who were in the same school but in different grades. This leads to a sample of 3,443 twin pairs (880 MZ twin pairs in the same classroom, 596 MZ twin pairs in different classrooms, 1,444 DZ twin pairs in the same classroom, and 693 DZ twin pairs in different classrooms). Similar to the main analyses, we first checked the assumptions of equal means and variances using likelihood-ratio tests. Distinguishing a classroom component requires the variance in educational performance to be equal between the different groups (i.e., MZ and DZ twins in the same class, and MZ and DZ twins in different classes). However, in contrast to the main analyses, the assumption of equal variances between twins in the same and different classrooms cannot be met when we use the stricter operationalization of classroom allocation. Since there are different ways to deal with this, we provide three robustness analyses. First, we fitted an *ACE* model for MZ and DZ twins in the same and different classrooms where we allow the variances to be different (see Analysis 1). This is probably the most appropriate test, but less informative as it does not allow for decomposing the fourth classroom variance component. Therefore, as alternative tests, we ignored that the variances are unequal and nevertheless performed analyses including a classroom component. In Analysis 2, we ignored the violation of equal variances by fitting the model with a classroom component while we constrained the variances across groups to be equal (although the LR-test showed that they cannot be equalized). Lastly, in Analysis 3, we estimated the model with a classroom component while we equalized the variances by standardizing the *Cito* scores within each group. This standardization artificially creates equal variances between the groups.

All three tests lead to the same conclusions as those reported in the main text: there is classroom variance that significantly decrease with increasing parental education. The evidence for a compensation effect is thus robust against a different operationalization of classroom allocation and different ways of modeling.

Analysis 1: ACE model

We decomposed the variance in *Cito* scores (unstandardized) for MZ and DZ twins in the same and different classes into *ACE* components, where we allowed the variances to differ between twins in the same and different classrooms. Although we do not directly estimate a classroom component, the *C* and *E* components for twins in the same and different classes allow us to infer whether there is a classroom effect. If twins are in the same class, the classroom is part of their shared environment. If there are classroom effects, one would thus expect twins in the same class to be more alike (larger *C*) than twins in different classes. Similarly, for twins in different classes, the classroom is part of their unique, non-shared environment. Hence, these twins can be expected to have a larger *E* than twins in the same class if classroom effects are present.

Table B2.1 shows larger shared environmental variance for twins in the same classroom (V_C = 9.17) than for twins in different classrooms (V_C = 6.37), indicating a classroom effect. Standardizing the variance components also shows that the shared environment explains a larger proportion of the variance in educational performance if twins are in the same classroom (12.4%) than if twins are in different classrooms (10.4%). While the absolute non-shared environmental

variance is similar in both groups, the standardized results show that the non-shared environment explains more variance for twins in different classrooms (24.0%) than for twins in the same classroom (19.2%). This is also consistent with classroom influences being present.

Table B2.2 shows the sources of variance for twins in the same and different classrooms moderated by parental education. The difference in shared environmental variance between twins in the same and different classrooms becomes smaller with increasing parental education (see Figure B2.1). The same applies to non-shared environmental variance. With lower levels of parental education, there is a larger difference in these sources of environmental variance, indicating classroom influences. These become smaller and are eventually absent with increasing parental education.

Analysis 2: ACE model including classroom component with unstandardized Cito scores

Table B2.3 shows the results of the model including a classroom component where we equalized the variance in educational performance for twins in the same and different classes, even though the LR-test showed that the variances differ. Model 1 shows that there is very small and non-significant classroom variance (V_{CL} = 0.25). Only 0.3% of the variance in educational performance can be explained by classroom effects.

Classroom variance decreases significantly with increasing parental education (b_{cl} = -0.02, p = .042), as shown in Model 2. Figure B2.2 plots the moderation of the unstandardized and standardized variance components. In both cases, the results show a pattern of compensation. For children with the lowest educated parents (primary education - ISLED 22.98), 5.2% of the variance in educational performance can be attributed to the classroom context. For the average parental education (higher and pre-scientific secondary school – ISLED 67.11), 2.4% can be attributed to classroom effects, and for children with the highest educated parents (postdoctoral education – ISLED 92.63) this is 0.9%.

Analysis 3: ACE model including classroom component with standardized Cito scores

Another way to deal with unequal variances is by standardizing the *Cito* scores within each group. The results are presented in Table B2.4. There is significant classroom variance ($\boldsymbol{V_{CL}}$ = 0.04, p = .037), making up 4.5% of the total variance in educational performance. Classroom variance decreases significantly with increasing parental education (b_{cl} = -0.003, p = .020), as shown in Model 2. Figure B2.3 plots the moderation of the unstandardized and standardized variance components. In both cases, the results show a pattern of compensation. For children with the lowest educated parents (primary education - ISLED 22.98), 10.6% of the variance in educational performance can be attributed to the classroom context. For the average parental education (higher and pre-scientific secondary school – ISLED 67.11), 6.3% can be attributed to classroom effects. For children with the highest educated parents (postdoctoral education – ISLED 92.63), this is 3.5%.

Table B2.1 Results of twin model for unstandardized educational performance for MZ twins in the same classroom (N_{pairs} = 880), DZ twins in the same classroom (N_{pairs} = 1,444), MZ twins in different classrooms (N_{pairs} = 596), and DZ twins in different classrooms (N_{pairs} = 693).

	Same classroom		Different classrooms	
	Estimate	**se**	**Estimate**	**se**
Intercept	551.06 ***	(3.83)	551.06 ***	(3.83)
a	7.13 ***	(0.27)	6.35 ***	(0.38)
c	3.03 ***	(0.55)	2.52 **	(0.83)
e	3.78 ***	(0.13)	3.83 ***	(0.17)
V_A	50.76 ***	(3.77)	40.31 ***	(4.81)
V_C	9.17 **	(3.36)	6.37	(4.20)
V_E	14.27 ***	(0.96)	14.70 ***	(1.30)
V_{Total}	74.20 ***	(1.94)	61.38 ***	(2.15)
Freely estimated parameters	33			
LL	-29756.05			
AIC	59578.11			

Note: Estimates are unstandardized, controlled for year of birth and sex in all models. *Se* = standard error, V_A = genetic variance, V_C = shared environmental variance, V_E = non-shared environmental variance, *Vtotal* = total variance, *LL* = loglikelihood, *AIC* = Akaike Information Criterion.

*p<.05, **p< .01, ***p< .001 (two-tailed tests).

Table B2.2. Results of twin moderation model for unstandardized educational performance for MZ twins in the same classroom (N_{pairs} = 880), DZ twins in the same classroom (N_{pairs} = 1,444), MZ twins in different classrooms (N_{pairs} = 596), and DZ twins in different classrooms (N_{pairs} = 693).

	Same classroom		Different classrooms	
	Estimate	**se**	**Estimate**	**se**
Intercept	535.04 ***	(3.64)	535.04 ***	(3.64)
Parental education	0.16 ***	(0.01)	0.16 ***	(0.01)
a	9.67 ***	(0.58)	8.31 ***	(1.58)
c	1.35	(1.89)	5.28	(3.74)
e	4.56 ***	(0.45)	6.36 ***	(0.73)
b_a	-0.04 ***	(0.01)	-0.03	(0.02)
b_c	0.00	(0.01)	-0.06	(0.05)
b_e	-0.01	(0.01)	-0.04 ***	(0.01)
Freely estimated parameters	48			
LL	-44098.16			
AIC	88292.32			

Note: Estimates are unstandardized, controlled for year of birth and sex in all models. *Se* = standard error, *a* = genetic path, *c* = shared environmental path, *e* = non-shared environmental path, *b* = moderation by parental education, LL = loglikelihood, *AIC* = Akaike Information Criterion.

***p<.001 (two-tailed tests).

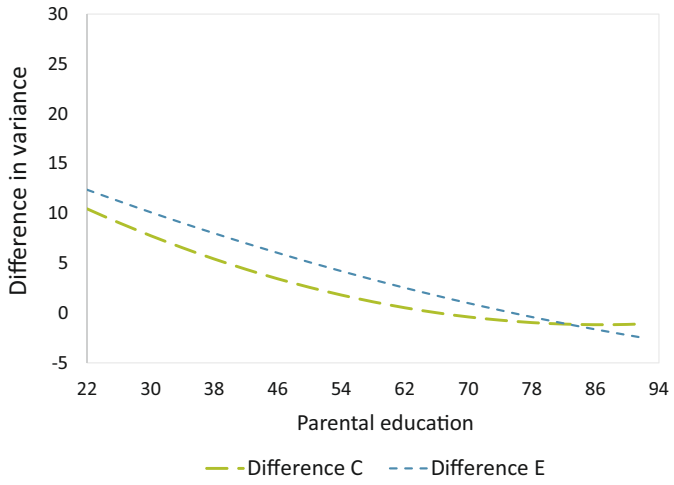


Figure B2.1 Difference in shared environmental (C) and non-shared environmental (E) variance in unstandardized educational performance between twins in the same and different classes, moderated by parental education.

Note: Results based on Table B2.

Table B2.3. Results of twin models for unstandardized educational performance for MZ twins in the same classroom (N_{pairs} = 880), DZ twins in the same classroom (N_{pairs} = 1,444), MZ twins in different classrooms (N_{pairs} = 596), and DZ twins in different classrooms (N_{pairs} = 693).

	Model 1		Model 2	
	Estimate	**se**	**Estimate**	**se**
Intercept	550.31 ***	(3.82)	534.85 ***	(3.63)
Parental education			0.16 ***	(0.01)
a	6.89 ***	(0.22)	9.53 ***	(0.48)
c	2.84 ***	(0.22)	0.00	(0.00)
cl	0.50	(1.43)	2.81 *	(1.10)
e	3.79 ***	(0.12)	4.80 ***	(0.40)
b_a			-0.04 ***	(0.01)
b_c			0.00	(0.00)
b_{cl}			-0.02 *	(0.01)
b_e			-0.02 **	(0.01)
V_A	47.45 ***	(3.02)	[a]	
V_C	8.07 **	(2.68)	[a]	
V_{CL}	0.25	(1.41)	[a]	
V_E	14.34 ***	(0.93)	[a]	
V_{Total}	70.10 ***	(1.50)	[a]	
Freely estimated parameters	31		44	
LL	-29765.90		-44102.47	
AIC	59593.79		88292.94	

Note: Estimates are unstandardized, controlled for age and sex in all models. *Se* = standard error, V_A = genetic variance, V_C = shared environmental variance, V_{CL} = classroom variance, V_E = non-shared environmental variance, V_{total} = total variance, *LL* = loglikelihood, *AIC* = Akaike Information Criterion.

[a] Not applicable, because in the moderation model the size of a variance component depends on the level of parental education.

*p<.05, **p<.01, ***p<.001 (two-tailed tests)

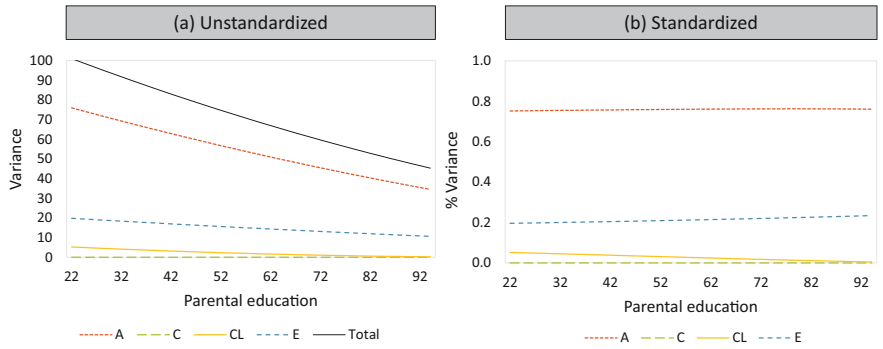


Figure B2.2 Decomposition of the unstandardized (a) and standardized (b) variance in unstandardized educational performance moderated by parental education.

Note: Sources of variance include genetic (*A*), shared environmental (*C*), classroom (*CL*), and non-shared environmental (*E*) variance. Results based on Table B3 Model 2.

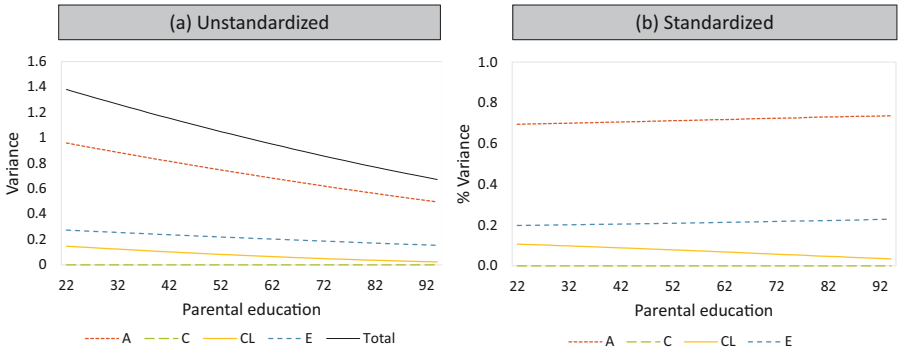


Figure B2.3. Decomposition of the unstandardized (a) and standardized (b) variance in standardized educational performance moderated by parental education.

Note: Sources of variance include genetic (*A*), shared environmental (*C*), classroom (*CL*), and non-shared environmental (*E*) variance. Based on Table B4 Model 2.

Table B2.4. Results of twin models for standardized educational performance for MZ twins in the same classroom (N_{pairs} = 880), DZ twins in the same classroom (N_{pairs} = 1,444), MZ twins in different classrooms (N_{pairs} = 596), and DZ twins in different classrooms (N_{pairs} = 693).

	Model 1		Model 2	
	Estimate	**se**	**Estimate**	**se**
Intercept	1.51 ***	(0.46)	-0.24	(0.44)
Parental education			0.02 ***	(0.00)
a	0.81 ***	(0.03)	1.06 ***	(0.06)
c	0.31 ***	(0.07)	0.00	(0.01)
cl	0.21 ***	(0.05)	0.46 ***	(0.11)
e	0.21 ***	(0.05)	0.56 ***	(0.05)
b_a			0.00 ***	(0.00)
b_c			0.00	(0.00)
b_{cl}			0.00 *	(0.00)
b_e			0.56 ***	(0.05)
V_A	0.65 ***	(0.04)	[a]	
V_C	0.10 *	(0.04)	[a]	
V_{CL}	0.04 *	(0.02)	[a]	
V_E	0.20 ***	(0.01)	[a]	
V_{Total}	0.99 ***	(0.02)	[a]	
Freely estimated parameters	31		44	
LL	-15532.94		-29905.32	
AIC	31127.88		59898.64	

Note: Estimates are unstandardized, controlled for age and sex in all models. *Se* = standard error, *VA* = genetic variance, *VC* = shared environmental variance, *VCL* = classroom variance, *VE* = non-shared environmental variance, *Vtotal* = total variance, *LL* = loglikelihood, *AIC* = Akaike Information Criterion.

[a] Not applicable, because in the moderation model the size of a variance component depends on the level of parental education.

*p<.05, **p<.01, ***p<.001 (two-tailed tests).

C. Supplements to Chapter 3

C.1. Measurement model school quality

Several steps were taken to construct this variable. First, we correlated all the items with the average *Cito* score of the school.[42] There were only a few items that correlated well with average educational performance, so we included also low positive correlations (larger than ~.030) that were statistically significant. Second, we divided the items into items that are related to school resources and to school climate. Third, we ran two Exploratory Factor Analyses, one for school resources and one for school climate. Several criteria were used for extracting factors: Kaiser's criterion (eigenvalue >1), inspection of the scree plot, and interpretability criteria. The interpretability criteria were: there should be at least two items with substantial loading (>.3), items that load on a factor should share conceptual meaning, and the rotated factor pattern should demonstrate a simple structure (no cross-loadings; items that load high on one factor have low loadings on the other factors). We extracted the resulting number of factors and excluded items that did not load (loading <.3) on any of the factors. All the resulting models have a good model fit (RMSEA, CFI, SRMR).

This procedure led to two dimensions of school resources (Table C1.1), and seven dimensions of school climate (Table C1.2). We saved these nine dimensions as nine factor scores and constructed one overall factor out of them by using the nine factor scores in a new Confirmatory Factor Analysis (CFA). As an alternative operationalization, we excluded the dimensions that had a low loading on the overall school quality factor (Table C1.3) and/or had a low correlation with schools' average *Cito* score (Table C1.4). This alternative operationalization based on the remaining six dimensions did not lead to substantially different results and therefore we kept all the nine dimensions in.

To correct for measurement error, we calculated the error variance of the nine school quality indicators (that are saved as factor scores) and specified this in the measurement model when constructing the overall school quality factor. Also, we specified the error variance for the overall school quality factor when it was included as a single indicator variable in the analytical model. The error variance (see Brown, 2015) is calculated as:

$$Var(error) = Var(X) \times (1 - \rho_X) \qquad (C.1)$$

where *Var(error)* is the error variance of the single indicator variable *X*, *Var(X)* is the total variance of the single indicator variable, and ρ_X is the reliability coefficient of the single indicator variable. The reliability coefficient is calculated by:

$$\rho_X = \frac{VA \times Lsum^2}{VA \times Lsum^2 + Vsum + 2 \times Csum} \qquad (C.2)$$

42 We did this to optimize our measure for school quality. We also extracted the factor scores without dropping the items that did not correlate with school-level educational performance. The resulting school quality variable had a slightly lower main effect on students' educational performance, but it did not substantially change our results.

where *VA* is the variance of the factor, *Lsum* is the sum of all item loadings, *Vsum* is the sum of all residual variances of the items, and *Csum* is the sum of all residual covariances (which is zero if no covariance parameters are included in the measurement model, as in our case). We have two sets of single indicator variables. First, the nine school quality dimensions, and second, the two second-order overall school quality factors. The resulting reliability coefficients and error variances can be found in Table C1.3.

Table C1.1 Factor analysis for the resource items ($N_{schools}$ = 9,709).

Label	Range of educational activities		(Implementation of) school curriculum	
The school provides a wide range of activities aimed at acquiring knowledge, insight, skills, and attitudes.	.86	(.01)		
The school ensures that there is coherence between the subjects concerning cross-curricular skills.	.50	(.02)		
General, cross-curricular skills and attitudes are part of the supply in the relevant subjects.	.87	(.02)		
The school ensures that the actual educational program corresponds to the planned program.			.47	(.05)
For groups of pupils, the school ensures a school curriculum that fits the educational needs of the pupils.			.45	(.05)
For individual pupils, the school ensures a school curriculum that fits the educational needs of the pupils.			.52	(.04)
The school curriculum is adjusted to the educational needs of the pupils.			.48	(.04)
The school curriculum prepares pupils for further education.			.39	(.06)
The school curriculum is partly aimed at the desired language development of the pupils.			.37	(.03)
School varies the amount of time for education, depending on the educational needs of the pupils.			.36	(.04)

Note: All factor loadings are statistically significant (p <.001, two-tailed test). Standard error in parentheses. Standardized effects.

Table C1.2 Factor analysis for the climate items ($N_{schools}$ = 9,709).

Item	Guidance of educational needs		Parental involvement		Monitoring and evaluating (special needs) students		Learning climate	Social climate	Safety	Quality assurance
The school ensures guidance for pupils with specific educational needs.	.95	(.01)								
The asks for help from pupils who drop out are analyzed by the school.	.40	(.02)								
The school takes care for additional guidance for pupils who require care.	1.02	(.01)								
The school has organized care and guidance in a structural way.	.33	(.02)								
The school has favorable conditions for the guidance of pupils with specific educational needs.	.33	(.02)								
The school involves parents/guardians in the guidance of the school career of the pupils.			.41	(.04)						
Parents are widely informed about the state of affairs at school.			.40	(.04)						
School contacts.			.92	(.02)						
Contact with parents.			1.01	(.02)						
The teachers systematically follow and analyze the progress in development of the pupils.					.49	(.02)				
The school identifies which pupils need care in an early stage.					.32	(.02)				
The school regularly evaluates the effects of the care.					.52	(.02)				
The school implements the care systematically.					.36	(.02)				
The school evaluates the results of the pupils yearly.					.46	(.02)				

Table C1.2 Factor analysis for the climate items ($N_{schools}$ = 9,709). (continued)

Item	Guidance of educational needs	Parental involvement	Monitoring and evaluating (special needs) students		Learning climate		Social climate		Safety		Quality assurance	
The school ensures the quality of the learning process.			.46	(.02)								
The school analyses the nature of the care that special need pupils require, based on the collected data.			.50	(.02)								
The teachers explain clearly.					.31	(.02)						
The teachers create a task-oriented work environment.					.67	(.03)						
The pupils are active and engaged.					.60	(.03)						
Staff and pupils interact positively.							.33	(.08)				
The school takes care of a pleasant, motivating environment for pupils.							.91	(.09)				
The atmosphere at school is stimulating.							.83	(.09)				
The school has a safety policy.									.53	(.06)		
The school ensures safety.									1.00	(.09)		
The involvement of staff, parents, students, and external parties in the self-evaluation is guaranteed.											.66	(.02)
The school has procedures, plans, and instruments for a self-evaluation.											.73	(.02)
School is accountable to stakeholders about the achieved educational quality.											.38	(.02)

Note: All factor loadings are statistically significant (p <.001, two-tailed test). Standard error in parentheses. Standardized effects.

Table C1.3 Factor loadings, estimated reliability, and error variance for the dimensions measuring school quality ($N_{schools}$ = 9,709).

	ρ	Error variance	SQ1		SQ2	
1. Range of educational activities	.82	.14	.61	(.02)	.53	(.02)
2. (Implementation of) school curriculum	.65	.02	.38	(.02)	.30	(.02)
3. Guidance of educational needs	.79	.19	.63	(.02)	.56	(.02)
4. Parental involvement	.91	.02	.34	(.02)	[a]	
5. Monitoring and evaluating (special needs) students	.61	.08	.58	(.02)	.70	(.02)
6. Learning climate	.55	.05	.55	(.02)	.66	(.02)
7. Social climate	.79	.04	.25	(.03)	[a]	
8. Safety	.77	.07	.26	(.03)	[a]	
9. Quality assurance	.65	.18	.66	(.01)	.62	(.02)
ρ			.72		.74	
Error variance			.09		.06	

Note: All factor loadings are statistically significant (p<.001, two-tailed test). Standard error in parentheses. Standardized effects. SQ = school quality.

[a] Excluded because of low factor loading and/or low correlation with average cito score.

Table C1.4 Correlations between the dimensions and factor scores for school quality ($N_{schools}$ = 9,709) and average cito score of the school ($N_{schools}$ = 7,622).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Range of educational activities											
2. (Implementation of) school curriculum	.27***										
3. Guidance of educational needs	.40***	.33***									
4. Parental involvement	.18***	.18***	.21***								
5. Monitoring and evaluating (special needs) students	.19***	.10***	.26***	.10***							
6. Learning climate	.27***	.11***	.27***	.09***	.58***						
7. Social climate	.16***	.16***	.11***	.19***	.04***	.02					
8. Safety	.14***	.19***	.07***	.20***	.07***	.06***	.48***				
9. Quality assurance	.37***	.18***	.30***	.25***	.41***	.27***	.12***	.12***			
10. Factor score school quality 1 (9 dimensions)	.66***	.44***	.66***	.39***	.65***	.62***	.28***	.29***	.70***		
11. Factor score school quality 2 (6 dimensions)	.57***	.35***	.59***	.23***	.78***	.75***	.12***	.14***	.66***	.96***	
12. Cito (school)	.07***	.09***	.10***	.05***	.14***	.14***	.04***	.05***	.100***	.16***	.17***

Note. ***p<.001 (two-tailed test).

Table C1.5 Factor loadings for the resources and climate dimensions of school quality ($N_{schools}$ = 9,709).

	Resources		Climate	
1. Range of educational activities	.71	(.02)		
2. (Implementation of) school curriculum	.42	(.02)		
3. Guidance of educational needs			.61	(.02)
4. Parental involvement			.33	(.02)
5. Monitoring and evaluating (special needs) students			.61	(.03)
6. Learning climate			.57	(.03)
7. Social climate			.23	(.03)
8. Safety			.25	(.03)
9. Quality assurance			.67	(.01)
ρ	.48		.66	
Error variance	.22		.10	

Note: All factor loadings are statistically significant (p<.001, two-tailed test). Standard error in parentheses. Standardized effects.

C.2. Measurement model socioeconomic status (SES)

Table C2.1. Descriptive statistics and factor loadings for SES.

	N	Mean	S.D.	SES
ISCED father	1,023,068	4.32	2.11	.68 (<.01)
ISCED mother	1,148,977	3.95	2.08	.80 (<.01)
Income (percentile) father	1,749,550	75.93	22.54	.43 (<.01)
Income (percentile) mother	1,620,932	44.29	25.86	.51 (<.01)

Note: All factor loadings are statistically significant (p<.001, two-tailed test). Standard error in parentheses. All items were standardized prior to the factor analysis.

C.3. Intraclass Correlation Coefficients

We compare the ICC of SS and OS twins with SS and OS siblings. Sibling pairs were selected by choosing a random sibling in a family with a co-sibling who is closest in age. In order to increase comparability of sibling pairs with twin pairs, we only included closely spaced siblings with a maximum age difference of three years. In addition, we made the same sample selections as for twins: birth cohorts 1994-2007, available cito scores, attending the same primary school, and exclude missings on school quality.

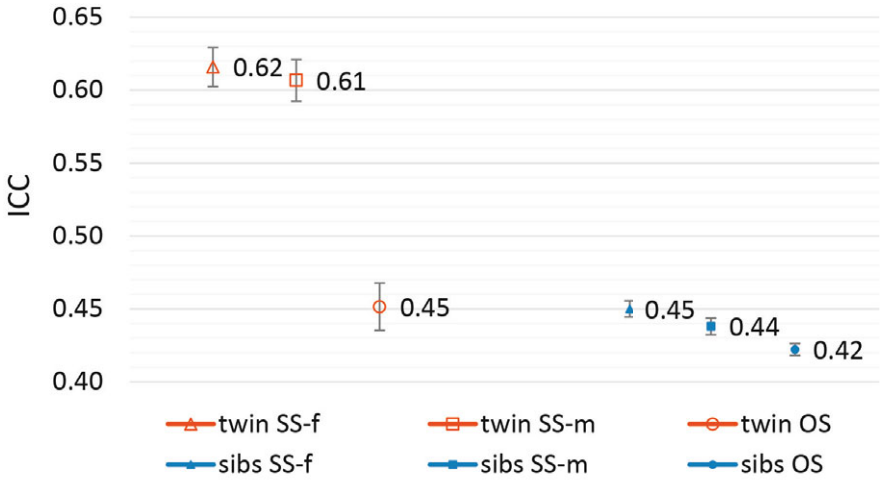


Figure C3.1. Intraclass correlation coefficient (ICC) for *Cito*-scores of same-sex female (SS-f), same-sex male (SS-m) and opposite-sex (OS) twin and sibling pairs, including 95% CI.

C.4. Additional results for different values of rSS_G

Table C4.1. *ACE* model for cito with main effects of school quality and school SES for different values of rSS_G ($N_{SSpairs}$ = 18,384, $N_{OSpairs}$ =11,050).

	rSS_G = .70						rSS_G = .75					
	Model 1		Model 2		Model 3		Model 1		Model 2		Model 3	
Parameter	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**
Intercept	534.63 ***	(0.08)	534.64 ***	(0.07)	534.65 ***	(0.07)	534.63 ***	(0.08)	534.64 ***	(0.07)	534.65 ***	(0.06)
a	9.28 ***	(0.06)	8.91 ***	(0.06)	8.33 ***	(0.06)	8.34 ***	(0.24)	8.34 ***	(0.23)	8.08 ***	(0.06)
c	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	2.85 ***	(0.48)	1.89 **	(0.71)	0.00	(0.00)
e	2.98 ***	(0.12)	3.35 ***	(0.10)	3.82 ***	(0.09)	4.16 ***	(0.16)	4.16 ***	(0.16)	4.31 ***	(0.07)
V_A	86.16 ***	(1.14)	79.29 ***	(1.08)	69.40 ***	(1.03)	69.61 ***	(3.95)	69.50 ***	(3.91)	65.32 ***	(0.96)
V_C	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	8.13 **	(2.73)	3.57	(2.67)	0.00	(0.00)
V_E	8.86 ***	(0.71)	11.20 ***	(0.70)	14.59 ***	(0.69)	17.27 ***	(1.32)	17.32 ***	(1.31)	18.56 ***	(0.64)
School quality	0.61 ***	(0.07)	0.25 ***	(0.06)	0.24 ***	(0.06)	0.61 ***	(0.07)	0.25 ***	(0.06)	0.24 ***	(0.06)
School SES			2.19 ***	(0.06)	0.92 ***	(0.06)			2.19 ***	(0.06)	0.92 ***	(0.06)
Parental SES					2.89 ***	(0.05)					2.89 ***	(0.05)
Freely estimated parameters	10		11		12		10		11		12	
Loglikelihood	-232015.70		-231114.62		-229734.35		-232015.68		-231112.87		-229725.54	
Scaling correction	2.57		2.38		2.28		2.70		2.50		2.29	
AIC	464051.41		462251.23		459492.70		464051.36		462247.74		459475.08	

Table C4.1. *ACE* model for cito with main effects of school quality and school SES for different values of rSS_G ($N_{SSpairs}$ = 18,384, $N_{OSpairs}$ =11,050). (continued)

	rSS_G = .80								
	Model 1			Model 2			Model 3		
Parameter	**Estimate**		**s.e.**	**Estimate**		**s.e.**	**Estimate**		**s.e.**
Intercept	534.63	***	(0.08)	534.64	***	(0.07)	534.66	***	(0.07)
a	7.62	***	(0.22)	7.61	***	(0.22)	7.61	***	(0.21)
c	3.73	***	(0.32)	3.06	***	(0.39)	1.66	*	(0.70)
e	4.80	***	(0.11)	4.81	***	(0.11)	4.81	***	(0.10)
V_A	58.01	***	(3.30)	57.93	***	(3.27)	57.96	***	(3.22)
V_C	13.92	***	(2.42)	9.36	***	(2.37)	2.76		(2.33)
V_E	23.07	***	(1.02)	23.11	***	(1.01)	23.11	***	(1.00)
School quality	0.61	***	(0.07)	0.25	***	(0.06)	0.24	***	(0.06)
School SES				2.19	***	(0.06)	0.92	***	(0.06)
Parental SES							2.89	***	(0.05)
Freely estimated parameters	10			11			12		
Loglikelihood	-232015.68			-231112.87			-229724.67		
Scaling correction	2.70			2.50			2.40		
AIC	464051.36			462247.74			459473.35		

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Controlled for sex and year of birth. All continuous independent variables are z-standardized prior to the analyses. Robust standard errors accounting for clustering at the school level are shown in parentheses. Parameters *a*, *c*, and *e* refer to unmoderated path coefficients capturing genetic, shared-environmental, and non-shared environmental influences, respectively.

C.5. Non-parametric gene-environment interaction

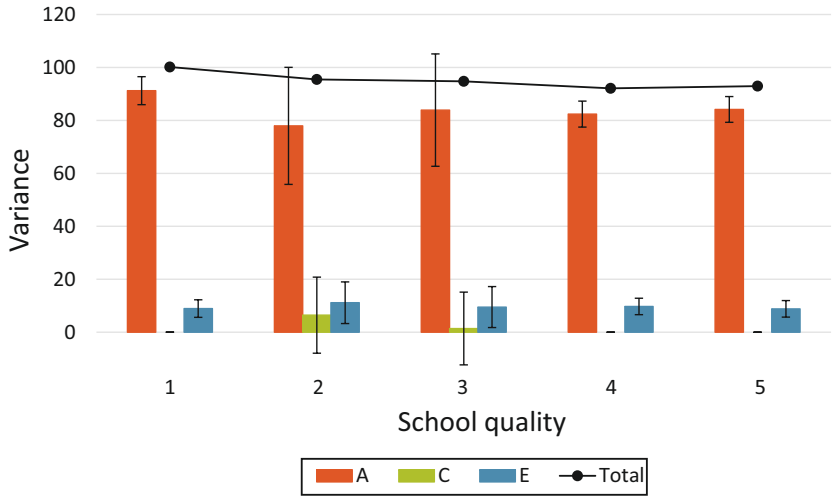


Figure C5.1. Unstandardized genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variances of educational performance moderated by quintiles of school quality in a non-parametric gene-environment interaction analysis, including 95% CI.

Note: Based on a model using a genetic correlation of rSS_G=.70, controlled for sex and year of birth.

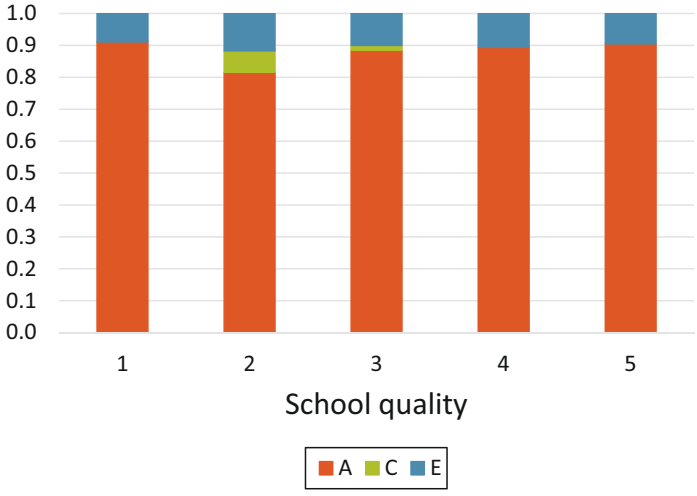


Figure C5.2. Standardized genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variances of educational performance moderated by quintiles of school quality in a non-parametric gene-environment interaction analysis.

Note: Based on a model using a genetic correlation of rSS_G=.70, controlled for sex and year of birth.

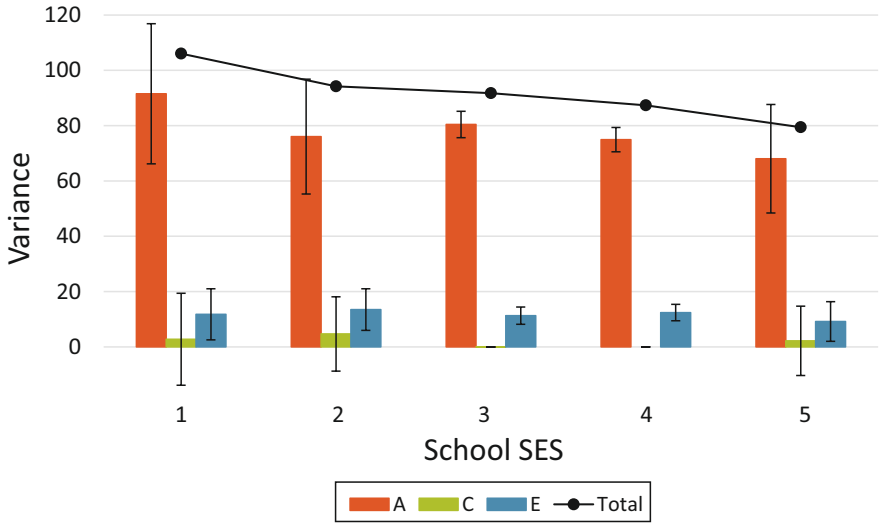


Figure C5.3. Unstandardized genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variances of educational performance moderated by quintiles of school SES in a non-parametric gene-environment interaction analysis, including 95% CI.

Note: Based on a model using a genetic correlation of rSS_G=.70, controlled for sex and year of birth.

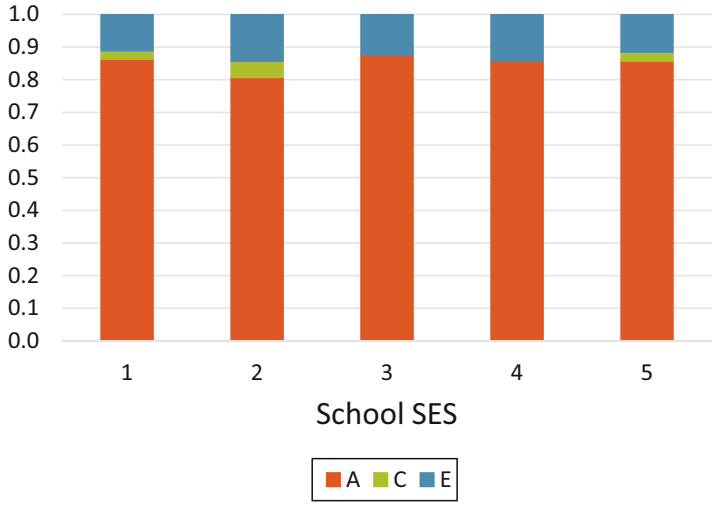


Figure C5.4. Standardized genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variances of educational performance moderated by quintiles of school SES in a non-parametric gene-environment interaction analysis.

Note: Based on a model using a genetic correlation of rSS_G=.70, controlled for sex and year of birth

C.6. Results school quality dimensions

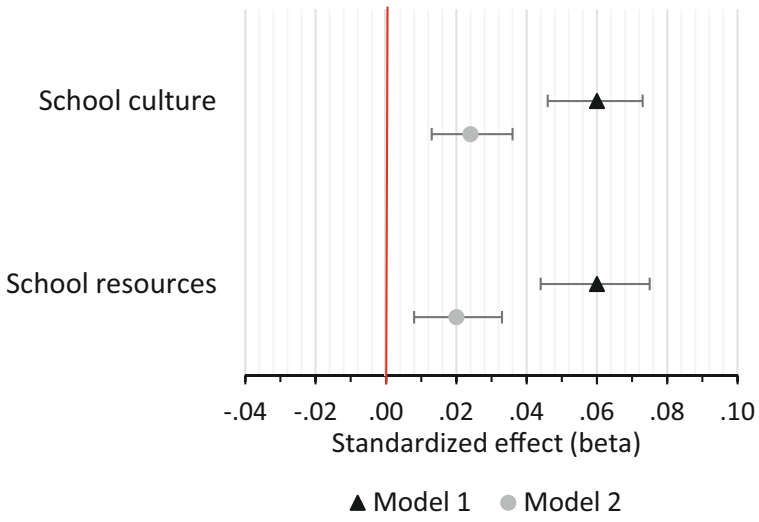


Figure C6.1. Main effects of the school climate and school resources on educational performance, including 95% CI.

Note: Based on separate analyses for each school quality factor (Model 1), controlled for school SES and parental SES (Model 2). Both models control for sex and year of birth. $N_{SSpairs}$ = 18,384; $N_{OSpairs}$ = 11,050.

Table C6.1 ACE model for cito with main effects and moderation effects of school quality (school resources or climate), school SES, and parental SES ($N_{SSpairs}$ = 18,384, $N_{OSpairs}$ =11,050).

	SQ = school resources				SQ = school climate			
	Model 1		Model 2		Model 1		Model 2	
Parameter	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**
Intercept	534.63 ***	(0.08)	534.67 ***	(0.07)	534.63 ***	(0.08)	534.67 ***	(0.06)
a	9.28 ***	(0.06)	8.21 ***	(0.07)	9.28 ***	(0.06)	8.21 ***	(0.07)
b_a*SQ*	-0.20 **	(0.06)	-0.01	(0.07)	-0.18 **	(0.06)	0.00	(0.06)
b_a*SchoolSES*			-0.31 ***	(0.06)			-0.31 ***	(0.06)
b_a*ParentalSES*			-0.61 ***	(0.06)			-0.62	(0.06)
c	[a]		0.25	(0.19)	0.00	(0.03)	0.25	(0.19)
b_c*SQ*	[a]		-0.33	(0.27)	0.00	(0.01)	-0.29	(0.25)
b_c*SchoolSES*			-0.85 ***	(0.23)			-0.86 ***	(0.22)
b_c*ParentalSES*			0.72 ***	(0.19)			0.73 ***	(0.19)
e	2.98 ***	(0.12)	3.92 ***	(0.09)	2.98 ***	(0.12)	3.92 ***	(0.09)
b_e*SQ*	0.09 ***	(0.12)	-0.01	(0.08)	0.05	(0.12)	-0.05	(0.09)
b_e*SchoolSES*			-0.04	(0.09)			-0.04	(0.09)
b_e*ParentalSES*			-0.23	(0.08)			-0.22 **	(0.08)
School quality	0.66 ***	(0.09)	0.21 **	(0.07)	0.62 ***	(0.07)	0.23 ***	(0.06)
School SES			0.91 **	(0.06)			0.91 ***	(0.06)
Parental SES			2.94	(0.05)			2.94 ***	(0.05)
Freely estimated parameters	11		21		13		21	
Loglikelihood	-217438.01		-214782.43		-217429.35		-214778.46	
Scaling correction factor	2.83		1.91		2.09		1.71	
AIC	434898.01		429606.87		434884.69		429598.92	

Note: *p<.05, **p<.01, ***p< .001 (two-tailed test). [a] Fixed to zero for model convergence. A genetic correlation of rSS_c = .70 is used. Controlled for sex and year of birth. All continuous independent variables are z-standardized prior to the analyses. Robust standard errors accounting for clustering at the school level are shown in parentheses. Parameters *a*, *c*, and *e* refer to the unmoderated path coefficients capturing genetic, shared-environmental, and non-shared environmental influences, respectively. The *b* coefficients refer to the moderation effects of *a, c*, and *e*, by school quality (i.e., resources or climate), school SES, and parental SES.

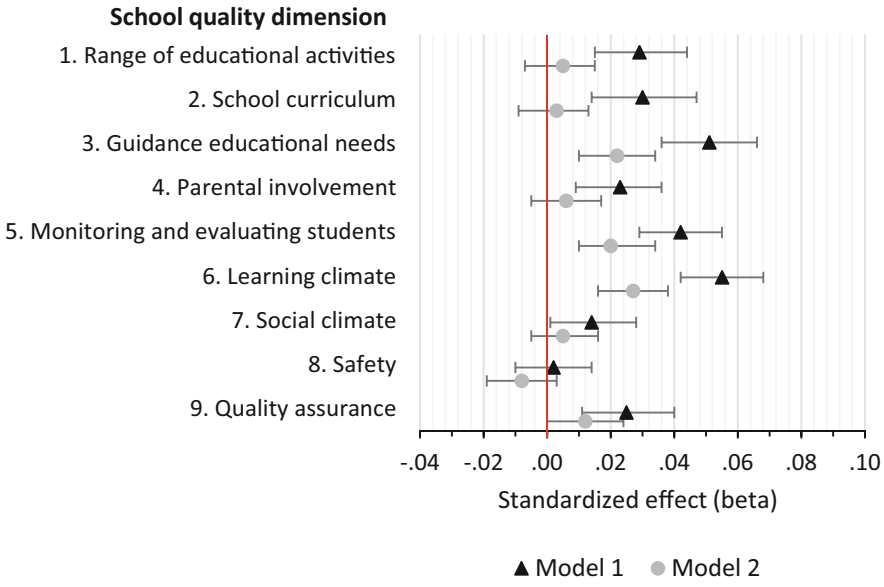


Figure C6.2 Main effects of the school quality dimensions on educational performance, including 95% CI.

Note: Based on separate analyses for each school quality dimensions (Model 1), controlled for school SES and parental SES (Model 2). Both models control for sex and year of birth. $N_{SSpairs}$ = 18,384, $N_{OSpairs}$ =11,050.

Table C6.2 Moderation effects of selected school quality dimensions (Model 1), controlled for the moderation by school SES and parental SES (Model 2) ($N_{SSpairs}$ = 18,384, $N_{OSpairs}$ =11,050).

	3. Guidance of educational needs				5. Monitoring and evaluating				6. Learning climate			
	Model 1		Model 2		Model 1		Model 2		Model 1		Model 2	
Parameter	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**
Intercept	534.63 ***	(0.08)	534.67 ***	(0.06)	534.63 ***	(0.08)	534.67 ***	(0.07)	534.63 ***	(0.08)	534.67 ***	(0.06)
a	9.24 ***	(0.08)	8.22 ***	(0.07)	9.29 ***	(0.06)	8.22 ***	(0.07)	9.23 ***	(0.14)	8.19 ***	(0.07)
b_a*SQ*	-0.09	(0.14)	0.00	(0.07)	-0.15 **	(0.05)	-0.05	(0.06)	-0.26	(0.14)	-0.10	(0.08)
b_a*SchoolSES*			-0.32 ***	(0.06)			-0.31 ***	(0.06)			-0.32 ***	(0.07)
b_a*ParentalSES*			-0.62 ***	(0.06)			-0.62 ***	(0.06)			-0.61 ***	(0.06)
c	-0.05	(0.96)	0.26	(0.18)	0.00	(0.04)	0.27	(0.19)	0.66	(1.01)	-0.22	(0.23)
b_c*SQ*	-0.78	(0.45)	-0.19	(0.21)	0.00	(0.00)	-0.09	(0.26)	0.46	(0.36)	0.65 *	(0.27)
b_c*SchoolSES*			-0.86 ***	(0.22)			-0.87 ***	(0.22)			0.77 **	(0.25)
b_c*ParentalSES*			0.72 ***	(0.19)			0.71 ***	(0.19)			-0.68 **	(0.20)
e	3.02 ***	(0.13)	3.91 ***	(0.09)	2.96 ***	(0.12)	3.91 ***	(0.09)	3.02 ***	(0.17)	3.93 ***	(0.09)
b_e*SQ*	0.01	(0.18)	-0.01	(0.09)	0.07	(0.10)	0.00	(0.08)	0.15	(0.17)	0.01	(0.10)
b_e*SchoolSES*			-0.05	(0.09)			-0.05	(0.09)			-0.04	(0.09)
b_e*ParentalSES*			-0.23 **	(0.08)			-0.23 **	(0.08)			-0.23 **	(0.08)
School quality	0.55 ***	(0.08)	0.22 **	(0.07)	0.43 ***	(0.07)	0.18 **	(0.06)	0.55 ***	(0.06)	0.24 ***	(0.05)
School SES			0.92 ***	(0.06)			0.92 ***	(0.06)			0.90 ***	(0.06)
Parental SES			2.94 ***	(0.05)			2.94 ***	(0.05)			2.94 ***	(0.05)

Table C6.2 Moderation effects of selected school quality dimensions (Model 1), controlled for the moderation by school SES and parental SES (Model 2) ($N_{SSpairs}$ = 18,384, $N_{OSpairs}$ =11,050). (continued)

	3. Guidance of educational needs				5. Monitoring and evaluating				6. Learning climate			
	Model 1		Model 2		Model 1		Model 2		Model 1		Model 2	
Parameter	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**	**Estimate**	**s.e.**
Freely estimated parameters	13		21		13		21		13		21	
Loglikelihood	-217454.58		-214781.47		-217463.41		-214781.28		-217435.49		-214771.26	
Scaling correction factor	2.26		1.72		2.22		1.80		2.76		2.08	
AIC	434935.16		429604.94		434952.82		429604.55		434896.98		429584.52	

Note: *p<.05, **p< .01, ***p< .001 (two-tailed test). Controlled for sex and year of birth. A genetic correlation of rSS_G=.70 is used. All continuous independent variables are z-standardized prior to the analyses. Robust standard errors accounting for clustering at the school level are shown in parentheses. Parameters *a*, *c*, and *e* refer to unmoderated path coefficients capturing genetic, shared-environmental, and non-shared environmental influences, respectively. The *b* coefficients refer to the moderation effects of *a*, *c*, and *e*, by the school quality dimension, school SES, and parental SES.

D. Supplements to Chapter 4

D.1. Supplementary tables

Table D1.1 ACE decomposition for boys and girls by family SES and school SES.

	1. Family SES			
	Low-SES family (*N* = 21,734)		High-SES family (*N* = 15,559)	
	Boys (*N* = 11,104)	Girls (*N* = 10,630)	Boys (*N* = 8,329)	Girls (*N* = 7,230)
Mean	5.70 (0.02) ***	6.39 (0.02) ***	7.28 (0.02) ***	7.98 (0.02) ***
Variance	4.66 (0.05) ***	4.45 (0.05) ***	3.99 (0.05) ***	3.51 (0.05) ***
% *A*	87.61 (4.63) ***	88.12 (4.20) ***	66.37 (6.19) ***	90.48 (5.52) ***
% *C*	8.01 (2.79) **	10.51 (2.61) ***	18.92 (3.57) ***	7.02 (3.37) *
% *E*	4.38 (2.01) *	1.37 (1.76)	14.71 (2.80) ***	2.51 (2.36)
A	4.08 (0.22) ***	3.92 (0.19) ***	2.65 (0.25) ***	3.17 (0.20) ***
C	0.37 (0.13) **	0.47 (0.12) ***	0.76 (0.15) ***	0.25 (0.12) *
E	0.20 (0.09) *	0.06 (0.08)	0.59 (0.12) ***	0.09 (0.14)

	2. School SES			
	Low-SES school (*N* = 17,373)		High-SES school (*N* = 19,920)	
	Boys (*N* = 8,966)	Girls (*N* = 8,407)	Boys (*N* = 10,467)	Girls (*N* = 9,453)
Mean	5.76 (0.02) ***	6.48 (0.02) ***	6.91 (0.02) ***	7.52 (0.02) ***
Variance	4.90 (0.06) ***	4.67 (0.06) ***	4.44 (0.05) ***	4.18 (0.05) ***
% *A*	73.89 (5.29) ***	85.29 (4.52) ***	74.49 (4.69) ***	78.25 (4.27) ***
% *C*	17.30 (3.12) ***	14.00 (2.81) ***	18.18 (2.79) ***	18.88 (2.63) ***
% *E*	8.82 (2.35) ***	0.70 (1.89)	7.33 (2.06) ***	2.88 (1.81)
A	3.62 (0.26) ***	3.98 (0.21) ***	3.30 (0.21) ***	3.27 (0.18) ***
C	0.85 (0.16) ***	0.65 (0.13) ***	0.81 (0.13) ***	0.79 (0.11) ***
E	0.43 (0.12) ***	0.03 (0.09)	0.33 (0.09) **	0.12 (0.08)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs.

Table D1.2. ACE decomposition for boys and girls by the intersection between family SES and school SES.

	Low-SES school							
	Low-SES family (*N* = 13,374)				High-SES family (*N* = 3,999)			
	Boys (*N* = 6,845)		Girls (*N* = 6,529)		Boys (*N* = 2,121)		Girls (*N* = 1,878)	
Mean	5.45 (0.02)	***	6.17 (0.06)	***	6.73 (0.04)	***	7.55 (0.04)	***
Variance	4.66 (0.06)	***	4.48 (0.02)	***	4.44 (0.11)	***	3.83 (0.10)	***
% *A*	86.82 (6.20)	***	91.16 (5.46)	***	50.53 (12.96)	***	97.59 (10.32)	***
% *C*	7.10 (3.70)		8.00 (3.39)	*	28.66 (7.31)	***	3.12 (6.44)	
% *E*	6.08 (2.71)	*	0.84 (2.29)		20.82 (5.99)	**	-0.71a (4.29)	
A	4.07 (0.29)	***	4.10 (0.26)	***	2.27 (0.58)	***	3.63 (0.29)	***
C	0.32 (0.17)		0.35 (0.16)	*	1.26 (0.33)	***	0.17 (0.20)	
E	0.27 (0.13)	*	0.03 (0.10)		0.91 (0.27)	**	0.02 (0.12)	

	High-SES school							
	Low-SES family (*N* = 8,360)				High-SES family (*N* = 11,560)			
	Boys (*N* = 4,259)		Girls (*N* = 4,101)		Boys (*N* = 6,208)		Girls (*N* = 5,352)	
Mean	6.10 (0.03)	***	6.74 (0.03)	***	7.47 (0.02)	***	8.13 (0.02)	***
Variance	4.40 (0.08)	***	4.21 (0.08)	***	3.70 (0.05)	***	3.31 (0.05)	***
% *A*	94.03 (7.28)	***	87.93 (6.82)	***	76.23 (7.22)	***	90.20 (6.69)	***
% *C*	4.17 (4.45)		10.05 (4.24)	*	11.36 (4.21)	**	6.02 (4.05)	
% *E*	1.80 (3.10)		2.02 (2.85)		12.41 (3.23)	***	3.77 (2.88)	
A	4.18 (0.34)	***	3.75 (0.30)	***	2.82 (0.28)	***	2.98 (0.23)	***
C	0.16 (0.21)		0.40 (0.19)	*	0.42 (0.16)	**	0.20 (0.14)	
E	0.06 (0.15)		0.07 (0.13)		0.46 (0.12)	***	0.12 (0.10)	

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs. [a] A negative estimate is actually impossible and could results from low power, small effect size, or model misspecification (see Lyu and Garrison, 2023 and the discussion in Appendix D2). We can force this estimate to have a lower bound of zero and the ACE decomposition would then become (94.6%, 4.7%, 0.6%), but this is generally discouraged (Lyu & Garrison, 2023; Verhulst et al., 2019) .

D.2. Robustness checks

D.2.1. Intraclass Correlation Coefficients

The interactions by gender, family SES, school SES, and their intersections can also be investigated based on the Intraclass Correlation Coefficients (ICCs) of twins and siblings instead of the *ACE* parameters that are derived from them. An advantage of this approach is that the assumptions that are needed to fit the *ACE* model, which are sometimes seen as controversial, can be relaxed (Turkheimer & Horn, 2014). In our case, relying on ICCs instead of *ACE* components also circumvents the potential problem that we do not know the exact genetic relatedness of SS twins.

Since heritability is computed by twice the difference between the ICCs for MZ and DZ twins, a larger difference between the ICCs results in a higher heritability. Therefore, if the ICCs for MZ and DZ twins would diverge as a function of SES, this is consistent with a larger heritability in higher-SES environments (i.e., enhancement, bioecological model) (Turkheimer & Horn, 2014). Applied to our study, we expect the ICCs to diverge between SS twins and siblings with increasing SES if it follows the enhancement pattern, and to converge if it follows the compensation pattern.

The results based on the ICCs are largely similar to what we found based on the *ACE* models. For school SES, our main analyses showed (i) a larger *E* for boys than girls in low-SES schools and (ii) no gene-environment interactions for boys and girls. For family SES, we found smaller *A* and larger *C* and *E* for boys in high-SES families, that is, a gene-environment interaction for boys in line with compensation. The findings based on the ICCs are largely in line with these (see Figure D2.1). The only difference is that although the converging lines for boys as a function of family SES follows a compensation pattern, the decrease in ICC of twin boys from .749 to .718 is statistically non-significant (p = .109).

When we look at the intersection between school SES and family SES, the findings based on the ICCs also support our conclusions based on the *ACE* models. As shown in Figure B2, the ICCs for boys in low-SES schools converge (i.e., compensation) while this is not the case for girls and in high-SES schools. This convergence for boys results both from a statistically significant decrease in the twin correlation (from .729 to .690, p = .027) and an increase in sibling correlation (from .504 to .537, p = .003) with increasing SES.

D.2.2. Alternative values for the average genetic relatedness of SS twins

We assumed that the average genetic relatedness of SS twins is .75, the average genetic relatedness of MZ and DZ twins. Additionally, we assumed that the shared-environmental relatedness among SS siblings is 1, while this may be slightly lower. A higher and lower genetic relatedness takes both uncertainties into account since genetic and shared environmental relatedness account for the same pattern in the data (see Spinath, Price, Dale, & Plomin, 2004). Therefore, we use .70 and .80 as alternative values for the genetic relatedness of twins. The results are presented in Tables D2.1-D2.4. As expected, the *ACE* estimates differ. Overall, the *A* is larger and *C* is smaller when a lower genetic relatedness is used (i.e., .70), while the reverse is true for the imposed genetic relatedness of .80. Importantly, the gender differences and interactions by family and school SES are similar, thus the conclusions based on the main analyses still hold.

It should be noted that the results sometimes show negative estimates, which are actually impossible. Alternative modeling strategies could be used to force the *ACE* estimates to be between 0 and 100%, but putting constraints on the higher and lower bound of the estimates is generally discouraged (Lyu & Garrison, 2023; Verhulst et al., 2019). Negative estimates are not unusual if the difference in genetic relatedness between the different pairs is less than .50, which is the case when comparing twins (with a genetic relatedness of .70, .75, or .80) with siblings (.50) (Lyu & Garrison, 2023). In general, it could occur from low power, small effect size, or model misspecification (Lyu & Garrison, 2023; Verhulst et al., 2019). For example, if educational performance is primarily affected by *A* and *C* and a relatively small number of pairs is used, confidence intervals of *E* will be large. The model will then lead to a negative *E* (Lyu & Garrison, 2023). Given that we perform *ACE* decompositions on smaller subgroups (by sex, family SES, and school SES), such a situation is not unlikely.

Also model misspecification may lead to negative estimates, if a too simplified model is used or if an *ADE* model would be better suited, for example (Lyu & Garrison, 2023). The twin/sibling correlations sometimes indicate non-additive genetic effects (*D*) when genetic relatedness among twins is set at .70. When zygosity is known, *D* is implicated if $r_{MZ} > 2*r_{DZ.}$ If zygosity is unknown and genetic relatedness is set at .70 for twins, *D* is implicated if $r_{SStwins} > 1.4*r_{SSsiblings.}$ This would become 1.5 and 1.6 times the sibling correlation if a genetic relatedness of .75 and .80 would be used, respectively. When we use a genetic relatedness of .70 for SS twins, we sometimes find indications of non-additive genetic effects.[43] Data on twins (and siblings) reared together do not provide enough information to identify both the *D* and the *C* components (Rijsdijk & Sham, 2002). Therefore, either an *ACE* or *ADE* model can be fitted. Since we want to compare our main results (based on .75) to those of alternative values, it is not desirable to compare *ACE*-models against *ADE*-models. Therefore, we always used *ACE* models, even if the twin/sibling correlations suggested otherwise. Although the estimates are sometimes impossible (e.g., *A* > 100%, *E<0%*), at least comparisons are possible.

D.2.3. Alternative birth spacing among SS siblings

Another way to check to what extent the potential smaller environmental relatedness among siblings (compared to twins) influences our results, is by using alternative birth spacing among siblings. As a robustness check, we select siblings that differ at most two instead of three years. This decreases the observed number of sibling pairs from 32,283 to 19,452. If birth spacing becomes closer, siblings' shared (family) environment can be expected to become more similar to one another and therefore also more similar to those of twins. As a result, smaller estimates of *A* and larger estimates of *C* can be expected.

As can be seen in Table D2.5 and D2.6 we indeed find smaller *A*'s and larger *C*'s now. Importantly, the gender differences in the *ACE* decomposition are substantively similar. However, in some cases, the differences are not statistically significant anymore.

43 Note that a similar pattern of correlations, that is, much greater MZ (SS twins) correlations than DZ (SS siblings) correlations is also consistent with the presence sibling contrast/competition effects (Neale & Maes, 2004; Rietveld et al., 2003), see also Table 1.1 in Chapter 1.

- In high-SES families, the difference between boys' and girls' *C* is no longer statistically significant.
- In the main analyses, we found that for boys *A* was significantly smaller in high-SES families and *C* and *E* were significantly larger. This interaction is now only statistically significant for the *E* component.
- When we look at the intersection between family SES and school SES (see Table D2.6), we still find that the interaction is driven by high-SES families in low-SES schools, but again, only statistically significant for the *E* component.

D.2.4. Cohort fixed effects

We included school cohort fixed effects when testing for gender differences in de gene-environment interplay for school SES and family SES (see Table D2.7) and their intersection (see Table D2.8). The results are substantively similar to those reported in the main text.

D.2.5. School SES terciles

We reran the analyses dividing school SES in terciles (i.e., low-, middle-, and high- SES schools) instead of quantiles based on the median value (i.e., low- and high-SES schools). The results of the gene-school SES interaction by gender are presented in Table D2.9 and show – similar to the main results – that there is no gene-school SES interaction for neither boys nor girls. None of the gender differences in *ACE* are statistically significant, except the gender difference in *E* in low-SES schools. Table D2.10 presents the results of the gene-family SES-school SES interaction by gender. These results also indicate that our conclusions are insensitive to the alternative specification of the SES groups. There is no gene-family SES interaction for girls in any of the school types. For boys, we do find a statistically significant gene-family SES interaction concentrated in low-SES schools but not in the other school types.

The gender differences in low-SES schools are more pronounced and more imprecisely estimated in low-SES schools when we divide school SES into terciles. This finding is a result of the loss of statistical power when breaking down SES into more groups. It should also be noted that the total sample size of these robustness analyses (N = 31,362) is somewhat lower than the main analyses (N = 37,293). This is due to the sample selection criterium that twins and siblings should attend the same school type, which in some cases no longer is the case when we distinguish three types of schools. Moreover, some of the estimates in Table B9 and B10 are imprecise and sometimes even impossible (e.g., >100 or <0). Ideally, we would use alternative modeling specifications, such as dropping a component if it is lower than zero or fitting an *ADE* instead of *ACE* model (see section 2 of Appendix D2). However, if we would do this, we could no longer compare the results between the several groups, or between the robustness analyses and the main analyses. Therefore, we used the same model specification as the main results. For the estimates the robustness checks with terciles, this implies that the focus should be on the patterns instead of the exact magnitude of the estimates.

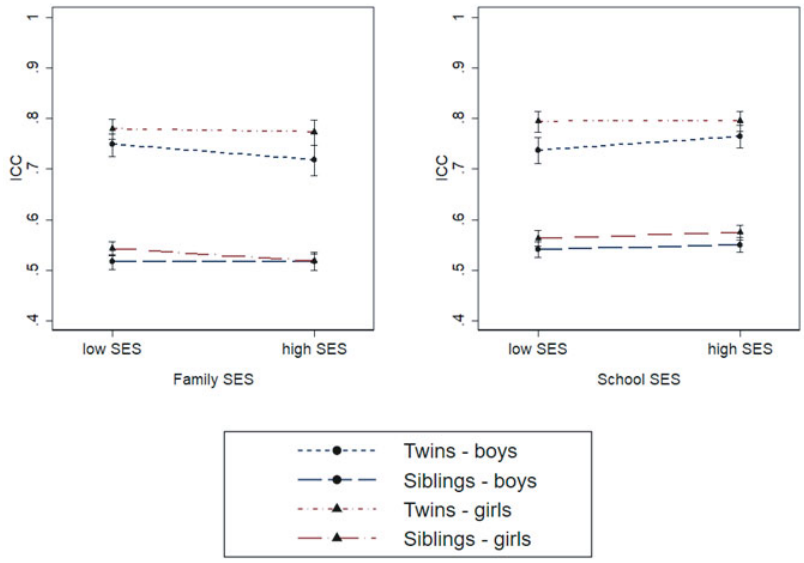


Figure D2.1. Intraclass Correlation Coefficient (ICC) for boys and girls by family SES and school SES, including 95% CI.

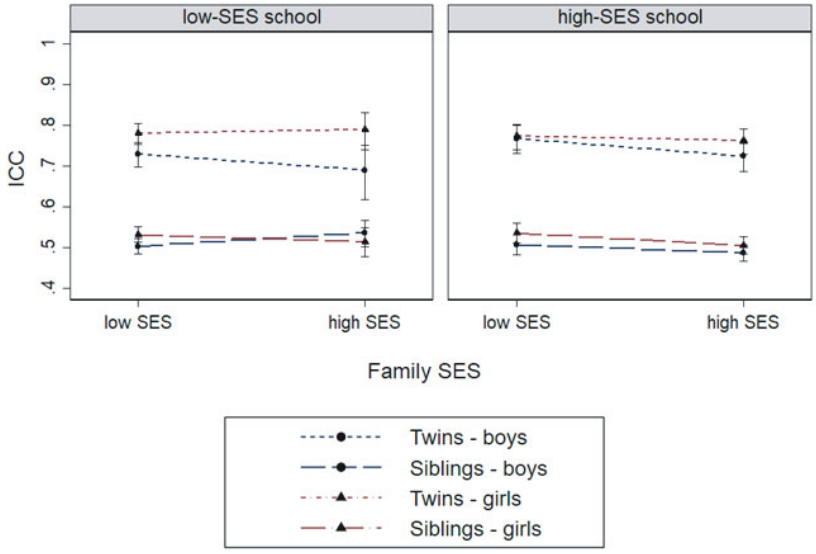


Figure D2.2. Intraclass Correlation Coefficient (ICC) for boys and girls by school SES x family SES, including 95% C

Table D2.1 ACE decomposition for boys and girls by family SES and school SES, genetic relatedness twins = .70.

	1. Family SES			
	Low-SES family (*N* = 21,734)		High-SES family (*N* = 15,559)	
	Boys (*N* = 11,104)	Girls (*N* = 10,630)	Boys (*N* = 8,329)	Girls (*N* = 7,230)
Mean	5.70 (0.02) ***	6.39 (0.02) ***	7.28 (0.02) ***	7.98 (0.02) ***
Variance	4.66 (0.05) ***	4.45 (0.05) ***	3.99 (0.05) ***	3.51 (0.05) ***
% *A*	109.52 (5.79) ***	110.15 (5.25) ***	82.96 (7.74) ***	113.10 (6.90) ***
% *C*	-2.94 (3.36)	-0.50 (3.12)	10.63 (4.33) *	-4.29 (4.05)
% *E*	-6.57 (2.56) **	-9.65 (2.26) ***	6.42 (3.56)	-8.80 (3.02) **

	2. School SES			
	Low-SES school (*N* = 17,373)		High-SES school (*N* = 19,920)	
	Boys (*N* = 8,966)	Girls (*N* = 8,407)	Boys (*N* = 10,467)	Girls (*N* = 9,453)
Mean	5.76 (0.02) ***	6.48 (0.02) ***	6.91 (0.02) ***	7.52 (0.02) ***
Variance	4.90 (0.06) ***	4.67 (0.06) ***	4.44 (0.05) ***	4.18 (0.05) ***
% *A*	92.36 (6.61) ***	106.62 (5.65) ***	93.12 (5.86) ***	97.81 (5.34) ***
% *C*	8.06 (3.77) *	3.34 (3.36)	8.86 (3.36) **	9.10 (3.15) **
% *E*	-0.42 (2.98)	-9.96 (2.43) ***	-1.98 (2.63)	-6.90 (2.32) **

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs.

Table D2.2 ACE decomposition for boys and girls by the intersection between school SES and family SES, genetic relatedness twins = .70.

	Low-SES school							
	Low-SES family (*N* = 13,374)				**High-SES family (*N* = 3,999)**			
	Boys (*N* = 6,845)		**Girls (*N* = 6,529)**		**Boys (*N* = 2,121)**		**Girls (*N* = 1,878)**	
Mean	5.45 (0.02)	***	6.17 (0.02)	***	6.73 (0.04)	***	7.55 (0.04)	***
Variance	4.66 (0.06)	***	4.48 (0.06)	***	4.44 (0.11)	***	3.83 (0.10)	***
% *A*	108.53 (7.75)	***	113.95 (6.82)	***	63.16 (16.20)	***	121.99 (12.91)	***
% *C*	-3.75 (4.46)		-3.39 (4.05)		22.34 (8.90)	*	-9.08 (7.70)	
% *E*	-4.78 (3.46)		-10.55 (2.94)	***	14.50 (7.57)		-12.91 (5.52)	*

	High-SES school							
	Low-SES family (*N* = 8,360)				**High-SES family (*N* = 11,560)**			
	Boys (*N* = 4,259)		**Girls (*N* = 4,101)**		**Boys (*N* = 6,208)**		**Girls (*N* = 5,352)**	
Mean	6.10 (0.03)	***	6.74 (0.03)	***	7.47 (0.02)	***	8.13 (0.02)	***
Variance	4.40 (0.08)	***	4.21 (0.08)	***	3.70 (0.05)	***	3.31 (0.05)	***
% *A*	117.54 (9.09)	***	109.91 (8.52)	***	95.29 (9.02)	***	112.75 (8.37)	***
% *C*	-7.59 (5.34)		-0.94 (5.07)		1.83 (5.09)		-5.25 (4.87)	
% *E*	-9.95 (3.97)	*	-8.97 (3.67)	*	2.88 (4.11)		-7.50 (3.69)	*

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs

Table D2.3 ACE decomposition for boys and girls by family SES and school SES, genetic relatedness twins = .80.

	1. Family SES			
	Low-SES family (*N* = 21,734)		High-SES family (*N* = 15,559)	
	Boys (*N* = 11,104)	Girls (*N* = 10,630)	Boys (*N* = 8,329)	Girls (*N* = 7,230)
Mean	5.70 (0.02) ***	6.39 (0.02) ***	7.28 (0.02) ***	7.98 (0.02) ***
Variance	4.66 (0.05) ***	4.45 (0.05) ***	3.99 (0.05) ***	3.51 (0.05) ***
% *A*	73.01 (3.86) ***	73.44 (3.50) ***	55.31 (5.16) ***	75.40 (4.60) ***
% *C*	15.31 (2.42) ***	17.85 (2.27) ***	24.45 (3.07) ***	14.56 (2.93) ***
% *E*	11.68 (1.64) ***	8.71 (1.43) ***	20.24 (2.31) ***	10.05 (1.93) ***

	2. School SES			
	Low-SES school (*N* = 17,373)		High-SES school (*N* = 19,920)	
	Boys (*N* = 8,966)	Girls (*N* = 8,407)	Boys (*N* = 10,467)	Girls (*N* = 9,453)
Mean	5.76 (0.02) ***	6.48 (0.02) ***	6.91 (0.02) ***	7.52 (0.02) ***
Variance	4.90 (0.06) ***	4.67 (0.06) ***	4.44 (0.05) ***	4.18 (0.05) ***
% *A*	61.57 (4.41) ***	71.08 (3.76) ***	62.08 (3.91) ***	65.21 (3.56) ***
% *C*	23.45 (2.69) ***	21.11 (2.44) ***	24.38 (2.41) ***	25.40 (2.29) ***
% *E*	14.97 (1.93) ***	7.81 (0.54) ***	13.54 (1.69) ***	9.40 (1.48) ***

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs.

Table D2.4 ACE decomposition for boys and girls by the intersection between school SES and family SES, genetic relatedness twins = .80.

	Low-SES school							
	Low-SES family (*N* = 13,374)				**High-SES family (*N* = 3,999)**			
	Boys (*N* = 6,845)		**Girls (*N* = 6,529)**		**Boys (*N* = 2,121)**		**Girls (*N* = 1,878)**	
Mean	5.45 (0.02)	***	6.17 (0.02)	***	6.73 (0.04)	***	7.55 (0.04)	***
Variance	4.66 (0.06)	***	4.48 (0.06)	***	4.44 (0.11)	***	3.83 (0.10)	***
% *A*	72.35 (5.16)	***	75.96 (4.55)	***	42.10 (10.80)	***	81.33 (8.60)	***
% *C*	14.34 (3.20)	***	15.60 (2.94)	***	32.87 (6.26)	***	11.25 (5.60)	*
% *E*	13.31 (2.22)	***	8.44 (1.86)	***	25.03 (4.95)	***	7.42 (3.49)	*

	High-SES school							
	Low-SES family (*N* = 8,360)				**High-SES family (*N* = 11,560)**			
	Boys (*N* = 4,259)		**Girls (N = 4,101)**		**Boys (*N* = 6,208)**		**Girls (*N* = 5,352)**	
Mean	6.10 (0.03)	***	6.74 (0.03)	***	7.47 (0.02)	***	8.13 (0.02)	***
Variance	4.40 (0.08)	***	4.21 (0.08)	***	3.70 (0.05)	***	3.31 (0.05)	***
% *A*	78.36 (6.06)	***	73.28 (5.68)	***	63.52 (6.02)	***	75.17 (5.58)	***
% *C*	12.00 (3.86)	***	17.38 (3.69)	***	17.71 (3.62)	***	13.54 (3.51)	***
% *E*	9.64 (2.53)	***	9.35 (2.32)	***	18.77 (2.66)	***	11.29 (2.36)	***

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs

Table D2.5 ACE decomposition for boys and girls by family SES and school SES (maximum sibling spacing of two years).

	1. Family SES			
	Low-SES school (*N* = 11,785)		High-SES school (*N* = 12,677)	
	Boys (*N* = 5,982)	Girls (*N* = 5,803)	Boys (*N* = 6,554)	Girls (*N* = 6,123)
Mean	5.68 (0.03) ***	6.40 (0.03) ***	6.85 (0.02) ***	7.46 (0.03) ***
Variance	5.05 (0.08) ***	4.82 (0.07) ***	4.60 (0.07) ***	4.33 (0.07) ***
% *A*	68.25 (5.58) ***	77.51 (4.81) ***	68.45 (5.03) ***	72.49 (4.81) ***
% *C*	22.19 (3.44) ***	20.44 (3.11) ***	23.46 (3.44) ***	23.83 (3.11) ***
% *E*	9.56 (2.36) ***	2.05 (1.91)	8.09 (2.36) **	3.68 (1.91) *

	2. School SES			
	Low-SES family (*N* = 14,420)		High-SES family (*N* = 10,042)	
	Boys (*N* = 7,218)	Girls (*N* = 7,202)	Boys (*N* = 5,318)	Girls (*N* = 4,724)
Mean	5.60 (0.02) ***	6.29 (0.02) ***	7.23 (0.02) ***	7.93 (0.03) ***
Variance	4.77 (0.07) ***	4.57 (0.06) ***	4.16 (0.07) ***	3.65 (0.06) ***
% *A*	79.60 (4.98) ***	81.85 (4.54) ***	64.63 (6.56) ***	82.22 (5.93) ***
% *C*	14.55 (3.15) ***	15.72 (2.95) ***	21.17 (3.98) ***	14.00 (3.80) ***
% *E*	5.85 (2.04) **	2.43 (1.79)	14.21 (2.83) ***	3.78 (2.38)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs.

Table D2.6. ACE decomposition for boys and girls by the intersection between family SES and school SES (maximum sibling spacing of two years).

	Low-SES school							
	Low-SES family (*N* = 9,046)				High-SES family (*N* = 2,739)			
	Boys (*N* = 4,537)		Girls (*N* = 4,509)		Boys (*N* = 1,445)		Girls (*N* = 1,294)	
Mean	5.36 (0.03)	***	6.07 (0.03)	***	6.69 (0.05)	***	7.53 (0.05)	***
Variance	4.77 (0.08)	***	4.56 (0.08)	***	4.61 (0.14)	***	4.08 (0.13)	***
% *A*	79.18 (6.60)	***	84.95 (5.90)	***	52.06 (13.47)	***	83.41 (10.71)	***
% *C*	13.37 (4.12)	*	13.01 (3.81)	**	28.43 (7.93)	***	15.00 (6.93)	*
% *E*	7.45 (2.75)	**	2.04 (2.34)		19.51 (6.00)	**	1.60 (4.24)	

	High-SES school							
	Low-SES family (*N* = 5,374)				High-SES family (*N* = 7,303)			
	Boys (*N* = 2,681)		Girls (*N* = 2,693)		Boys (*N* = 3,873)		Girls (*N* = 3,430)	
Mean	6.01 (0.04)	***	6.66 (0.04)	***	7.43 (0.03)	***	8.08 (0.03)	***
Variance	4.51 (0.10)	***	4.38 (0.10)	***	3.84 (0.07)	***	3.40 (0.07)	***
% *A*	84.14 (7.93)	***	81.26 (7.42)	***	73.07 (7.76)	***	83.93 (7.28)	***
% *C*	12.12 (5.11)	*	15.78 (4.84)	**	14.56 (4.77)	**	11.30 (4.64)	
% *E*	3.73 (3.17)		2.97 (2.91)		12.38 (3.28)	***	4.77 (2.95)	

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs

Table D2.7 ACE decomposition for boys and girls by family SES and school SES, controlled for cohort fixed effects.

	1. Family SES			
	Low-SES family (*N* = 21,734)		High-SES family (*N* = 15,559)	
	Boys (*N* = 11,104)	Girls (*N* = 10,630)	Boys (*N* = 8,329)	Girls (*N* = 7,230)
Mean	-0.98 (0.02) ***	-0.29 (0.02) ***	0.58 (0.02) ***	1.28 (0.02) ***
Variance	4.62 (0.05) ***	4.43 (0.05) ***	3.90 (0.05) ***	3.40 (0.05) ***
% *A*	91.55 (4.72) ***	91.57 (4.26) ***	69.68 (6.37) ***	93.49 (5.67) ***
% *C*	4.72 (2.84)	7.73 (2.65) **	15.57 (3.67) ***	4.05 (3.46)
% *E*	3.74 (2.04)	0.70 (1.78)	14.74 (2.88) ***	2.46 (2.42)

	2. School SES			
	Low-SES school (*N* = 17,373)		High-SES school (*N* = 19,920)	
	Boys (*N* = 8,966)	Girls (*N* = 8,407)	Boys (*N* = 10,467)	Girls (*N* = 9,453)
Mean	-0.90 (0.02) ***	-0.18 (0.02 ***	0.19 (0.02) ***	0.81 (0.02) ***
Variance	4.90 (0.06) ***	4.65 (0.06) ***	4.37 (0.05) ***	4.11 (0.05) ***
% *A*	76.67 (5.35) ***	87.89 (4.55) ***	78.03 (4.79) ***	80.82 (4.35) ***
% *C*	14.98 (3.26) ***	12.00 (2.84) ***	15.07 (2.85) ***	16.57 (2.68) ***
% *E*	8.34 (2.37) ***	0.11 (1.90)	6.90 (2.10) ***	2.62 (1.84)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs.

Table D2.8 *ACE* decomposition for boys and girls by the intersection between family SES and school SES, controlled for cohort fixed effects.

	Low-SES school							
	Low-SES family (*N* = 13,374)				High-SES family (*N* = 3,999)			
	Boys (*N* = 6,845)		Girls (*N* = 6,529)		Boys (*N* = 2,121)		Girls (*N* = 1,878)	
Mean	-1.20 (0.02)	***	-0.48 (0.02)	***	0.08 (0.04)	***	0.90 (0.04)	***
Variance	4.66 (0.06)	***	4.48 (0.06)	***	4.41 (0.11)	***	3.77 (0.10)	***
% *A*	89.72 (6.26)	***	94.08 (5.49)	***	53.77 (13.14)	***	99.34 (10.43)	***
% *C*	4.75 (3.74)		5.80 (3.41)		25.77 (7.43)	**	1.56 (6.51)	
% *E*	5.53 (2.74)	*	0.13 (2.30)		20.46 (6.06)	**	-0.89 (4.33)	

	High-SES school							
	Low-SES family (*N* = 8,360)				High-SES family (*N* = 11,560)			
	Boys (*N* = 4,259)		Girls (*N* = 4,101)		Boys (*N* = 6,208)		Girls(*N* = 5,352)	
Mean	-0.62 (0.03)	***	0.01 (0.03)	***	0.75 (0.02)	***	1.42 (0.02)	***
Variance	4.36 (0.08)	***	4.18 (0.08)	***	3.61 (0.05)	***	3.20 (0.05)	***
% *A*	99.04 (7.41)	***	91.34 (6.93)	***	79.35 (7.43)	***	93.43 (6.88)	***
% *C*	0.10 (4.54)		7.21 (4.31)		8.13 (4.32)		2.78 (4.17)	
% *E*	0.87 (3.15)		1.45 (2.89)		12.51 (3.33)	***	3.78 (2.97)	

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs.

Table D2.9 ACE decomposition for boys and girls by school SES (terciles).

	School SES					
	Low-SES school (*N* = 8,454)		Middle-SES school (*N* = 9,694)		High-SES school (*N* = 13,213)	
	Boys (*N* = 4,358)	Girls (*N* = 4,096)	Boys (*N* = 5,045)	Girls (N = 4,649)	Boys (*N* = 6,945)	Girls (*N* = 6,268)
Mean	5.58 (0.03) ***	6.24 (0.03) ***	6.26 (0.03) ***	6.92 (0.03) ***	7.08 (0.02) ***	7.71 (0.02) ***
Variance	4.95 (0.09) ***	4.75 (0.09) ***	4.82 (0.08) ***	4.48 (0.08) ***	4.30 (0.06) ***	3.96 (0.06) ***
% *A*	80.09 (6.91) ***	90.07 (5.95) ***	82.51 (6.29) ***	82.16 (5.64) ***	73.00 (5.74) ***	71.48 (5.43) ***
% *C*	13.89 (4.22) **	11.86 (3.84) **	12.37 (3.89) **	16.26 (3.63) ***	18.70 (3.43) ***	22.26 (3.31) ***
% *E*	6.02 (2.96) *	-1.93 (2.36)	5.11 (2.65)	1.58 (2.26)	8.30 (2.51) **	6.27 (2.33) **

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs.

Table D2.10 ACE decomposition for boys and girls by the intersection between family SES and school SES (terciles).

	Low-SES school							
	Low-SES family (*N* = 6,992)				High-SES family (*N* = 1,462)			
	Boys (*N* = 3,550)		Girls (*N* = 3,442)		Boys (*N* = 808)		Girls (*N* = 654)	
Mean	5.35 (0.03)	***	6.04 (0.03)	***	6.58 (0.07)	***	7.30 (0.07)	***
Variance	4.75 (0.09)	***	4.64 (0.09)	***	4.60 (0.19)	***	3.97 (0.18)	***
% *A*	96.98 (7.73)	***	92.06 (6.89)	***	26.49 (19.34)		110.53 (13.69)	***
% *C*	1.75 (4.81)		8.61 (4.42)		44.91 (10.96)	***	11.25 (9.26)	
% *E*	1.27 (3.23)		-0.67 (2.76)		28.60 (8.91)	**	-11.45 (5.00)	*

	Middle-SES school							
	Low-SES family (*N* = 6,097)				High-SES family (*N* = 3,597)			
	Boys (*N* = 3,136)		Girls (*N* = 2,961)		Boys (*N* = 1,909)		Girls (*N* = 1,688)	
Mean	5.73 (0.03)	***	6.44 (0.03)	***	7.12 (0.04)	***	7.76 (0.04)	***
Variance	4.56 (0.09)	***	4.34 (0.09)	***	4.06 (0.10)	***	3.63 (0.10)	***
% *A*	91.14 (8.42)	***	97.21 (7.56)	***	93.21 (11.84)	***	76.86 (10.54)	***
% *C*	3.93 (5.25)		4.03 (4.89)		0.85 (7.17)		15.69 (6.62)	*
% *E*	4.93 (3.50)		-1.24 (2.99)		5.94 (5.10)		7.45 (4.35)	

	High-SES school							
	Low-SES family (*N* = 4,726)				High-SES family (*N* = 8,487)			
	Boys (*N* = 2,403)		Girls (*N* = 2,323)		Boys (*N* = 4,542)		Girls (*N* = 3,945)	
Mean	6.17 (0.04)	***	6.83 (0.04)	***	7.56 (0.02)	***	8.22 (0.02)	***
Variance	4.42 (0.10)	***	4.08 (0.10)	***	3.57 (0.06)	***	3.17 (0.06)	***
% *A*	90.36 (9.53)	***	78.03 (9.33)	***	78.55 (8.27)	***	85.90 (7.94)	***
% *C*	6.86 (5.84)		15.19 (5.74)	**	9.31 (4.88)		8.16 (4.79)	
% *E*	2.78 (4.05)		6.78 (3.95)		12.14 (3.66)	**	5.94 (3.44)	

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Estimates with standard errors in parentheses. *A*, *C*, and *E* refer to genetic, shared-environmental, and non-shared environmental variance, respectively. *N* refers to the number of twin/sibling pairs.

E. Supplements to Chapter 4

E.1. Comparing the NCO twin sample to the full sample

For the external validity of this study, it is important that the (development in) dispersion of twins' educational performance is similar to that of pupils in the general population. I first assess whether the twin sample and the full sample differ with respect to the covariates that are included in this study. Compared to the full sample, twins have parents with a higher educational level and income and attend schools with a higher SES (see Table E1.1). However, although statistically significant, these differences are small. Moreover, the small differences between the twin sample and full sample do not result in differences in (the development of) dispersion. Table E1.2 shows the comparison of initial performance, learning growth, and their association between the twin sample and the full NCO-LVS sample. There are no substantial differences. Twins have on average lower initial reading and mathematics performance and higher growth rates than pupils in the total sample, but the differences are small (see Table E1.2). More importantly, the variances of initial performance (i.e., initial dispersion) and growth are either not statistically significantly different between the twin and full sample or only slightly different (see Table E1.2 and Figure E1.1). Also, the development of dispersion during the primary school career is highly similar for twins versus the general pupil population (see Figure E1.2).

Table E1.1 Differences in the family and school variables between the full sample and twin sample.

	Full sample			Twin sample			Difference (M)	
Variable	***N***	**M**	**SD**	***N***	**M**	**SD**	**Cohen's *d***	
ISCED father	259,662	4.40	2.08	7,718	4.56	2.07	-0.08	***
ISCED mother	297,423	4.36	2.07	8,353	4.53	2.07	-0.09	***
Income father	345,386	73.78	23.19	10,287	75.08	22.85	-0.06	***
Income mother	336,019	49.06	25.68	9,596	50.87	25.99	-0.07	***
SES school	373,063	0.00	0.86	10,692	0.08	0.85	-0.10	***
School quality	377,178	0.48	0.26	10,733	0.49	0.27	-0.04	

Note: ***p<.001 (two-tailed test). Samples excluding outliers on the intercept and slope.

Table E1.2. Differences between the full sample and twin sample in the means and variances of the intercept and slope as estimated in a latent growth model.

	Full sample			Twin sample			Difference (M)		Difference (SD)[a]	
	N	**M**	**SD**	***N***	**M**	**SD**	**Cohen's *d***		***F***	
Reading										
Intercept	291,860	123.52	20.90	8,602	122.32	20.81	0.06	***	1.01	
Slope	291,860	18.75	1.59	8,602	18.80	1.58	-0.03	**	1.02	
Math										
Intercept	354,319	132.40	24.61	10,495	130.09	24.22	0.09	***	1.03	*
Slope	354,319	32.04	2.69	10,495	32.08	2.63	-0.02		1.05	**

Note: **p*<.05, ***p*<.01, ****p*<.001 (two-tailed test). [a] Variance ratio test is used to test if the variance between the full sample and twin sample differs. This is the case if the variance ratio, $SD_{(full\ sample)}/SD_{(twin\ sample)}$, is statistically significantly different from 1.

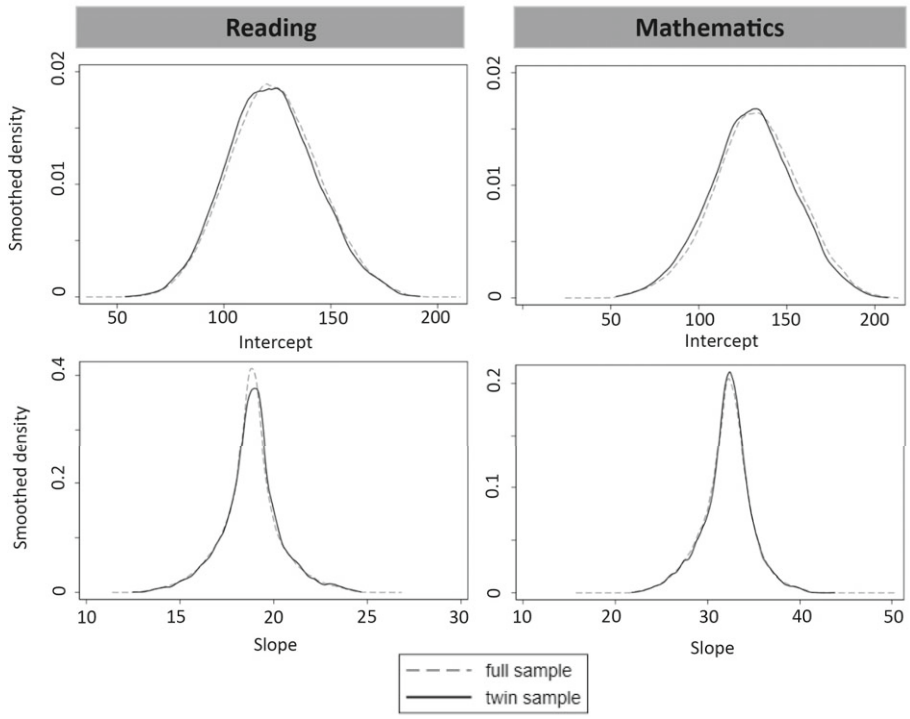


Figure E1.1 The distribution of the intercept and slope of reading (left) and mathematics (right) for the full sample and twin sample.

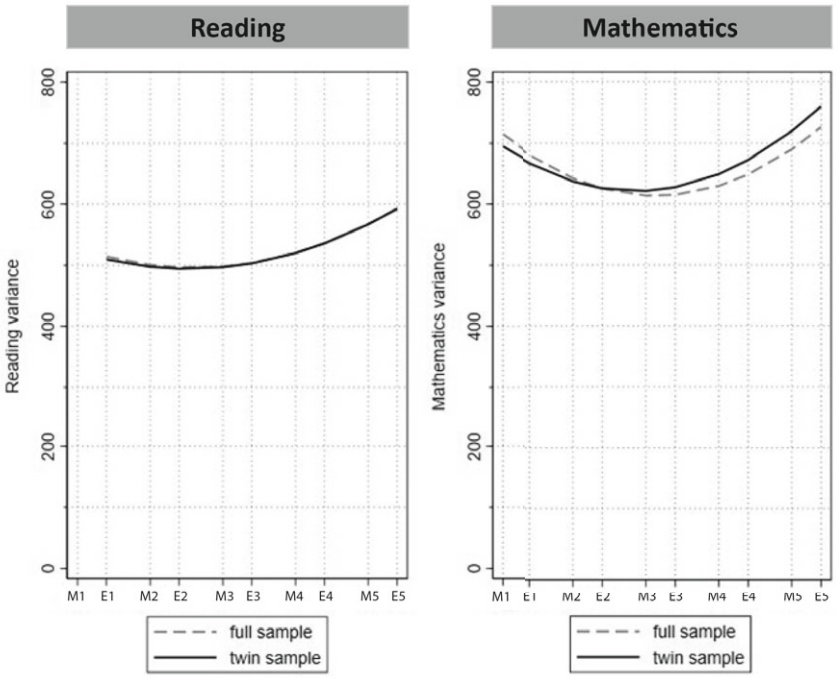


Figure E1.2 The total variance in reading (left) and mathematics (right) performance during primary school for the full sample and twin sample.

Note: Controlled for sex. Based on the variance of the intercept and slope, and their covariance.

E.2. Robustness checks

E.2.1. Exclusion of pairs of whom one or both twins repeated a grade

I performed the analyses while excluding the twin pairs where one or both twins repeated a grade during the observed primary school period (i.e., Grade 1-5). This leads to twin samples of N_{pairs}= 4,905 for reading and N_{pairs} = 5,165 for mathematics. The results of the biometric latent growth models for reading and mathematics are presented in Table E2.1, and the results of the models including the environmental covariates are shown in Table E2.2. The results are similar to those of the main analyses.

E.2.2. No exclusion of outliers on the intercept and slope

I performed the analyses while twin pairs with outliers on the estimated intercept and/or slope (defined as M ± 3 SD) are not excluded. This increases the sample size to N_{pairs}=5,591 for reading and N_{pairs}=5,722 for mathematics. If outliers are not excluded, there is more variance in the intercept and slope. However, the decomposition into genetic, shared environmental, and non-shared environmental components remains similar. Also, the covariances are similar to those reported in the main analyses. Only the small amount of initial shared environmental differences that are present slightly increases in a statistically significant way (p = .040) when outliers are not excluded (see Table E2.3). When the school and family variables are included (see Table E2.4), this explains more variance in the reading and mathematics slopes compared to the main analyses. However, the contribution of genetic and shared environmental variance remains similar regardless of whether the environmental explanatory variables were included. This confirms our finding that our school and family measures do not account for significant shared environmental influences on reading and mathematics.

E.2.3. Alternative operationalizations school quality

I perform three robustness checks with alternative operationalizations of the school quality measure. All results are controlled for sex, school SES, and parental SES. First, I use an alternative school quality factor score where the dimensions with low loadings are excluded (see Appendix E3 for the measurement model). The effects on the intercept are presented in Figure E2.1 (Model 1). The explained variance in the reading and mathematics intercepts (11% and 8%, respectively) and reading and mathematics slopes (4% and 1%, respectively) are similar to the main results. Second, I used the school climate and school resources factors instead of one overall school quality factor (see Figure E2.1, Model 2). The explained variance in the reading and mathematics intercepts are 11% and 9%, respectively. For the reading and mathematics slope, this is 4% and 2%, respectively. Lastly, I investigated the effects of all the nine underlying school dimensions separately (see Figure E2.2). The explained variance in the reading and mathematics intercepts are somewhat lower, 9% and 5%. The measures explain 4-5% of the variance in the reading slope and 1% in the mathematics slope.

E.2.4. Different genetic relatedness of SS twins

To assess how sensitive the results are to the assumed genetic relatedness among SS twins (rSS_G), I conducted the analyses with different values (i.e., .70 and .80 instead of .75). The rSS_G affect the estimated genetic and environmental variances; the higher the rSS_G, the less genetic variance and the more environmental variance. However, despite the value of rSS_G that is used, the conclusions remain the same. The development of dispersion over the school period is similar (see Figures E2.3 and E2.4). For reading, initial differences decrease (rather weakly) over the primary school career and this is related to the small decrease in initial genetic variance in reading. Concerning mathematics performance, initial differences likewise decrease over time, and this is related to a decrease in shared and non-shared environmental variance. The only difference is that the decrease in shared environmental variance in initial mathematics performance is no longer statistically significant (p = .056) when an assumed genetic relatedness of .70 is used (see Table E2.5). When using an rSS_G of .80, some new shared environmental variance is coming into play in mathematics performance (see Figure E2.4), but these new shared environmental influences unique to the slope are (like the main results) not statistically significant (see Table E2.6). Also, the results for the models including the explanatory family and school variables are substantially similar regardless of the assumed genetic relatedness that is used (see Table E2.7).

Table E2.1 Estimates of the *A, C, E* variance and covariance components of the intercept and slope for reading (N_{pairs}=4,905) and mathematics (N_{pairs}=5,165), using the sample excluding pairs where one or both twins repeated a grade during Grade 1-5.

	Reading				Mathematics			
	Total	**A**	**C**	**E**	**Total**	**A**	**C**	**E**
Intercept mean	126.15 ***				126.51 ***			
Slope mean	18.40 ***				31.80 ***			
Unstandardized								
Intercept	523.22 ***	456.36 ***	33.63	33.23 *	704.24 ***	519.80 ***	91.82	92.62 ***
Slope	12.19 ***	10.08 ***	0.19	1.92	21.06 ***	7.67	8.23 *	5.16 **
common with intercept	0.86	0.45	0.19	0.22	10.45	0.04	7.42	3.00
unique to slope	11.33 ***	9.63 ***	0.00	1.70	10.60	7.63	0.81	2.17
$COV_{I,S}$	-14.44 ***	-14.31 *	2.55	-2.68	-38.27 ***	4.48	-26.10 *	-16.66 **
Standardized								
Intercept		0.87 ***	0.06	0.06 *		0.74 ***	0.13	0.13 ***
Slope		0.83 ***	0.02	0.16		0.36	0.39 *	0.25 **
common with intercept		0.04	0.02	0.02		0.00	0.35	0.14
unique to slope		0.79 ***	0.00	0.14		0.36	0.04	0.10
$COV_{I,S}$		0.73 *	0.13	0.14		0.10	0.55 *	0.35 **

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Controlled for sex. *A* = genetic, *C* = shared environment, *E* = non-shared environment.

Table E2.2 Estimates of the biometric growth models including covariates for reading (N_{pairs}=4,905) and mathematics (N_{pairs}=5,165), using the sample excluding pairs where one or both twins repeated a grade during Grade 1-5.

	Reading		Mathematics	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Mean				
Intercept mean	126.15 ***	123.83 ***	126.51 ***	123.62 ***
$\beta_{school\ quality}$		0.02		0.05 **
$\beta_{school\ SES}$		0.04 *		0.03 *
$\beta_{parental\ SES}$		0.11 **		0.03 *
Slope mean	18.40 ***	17.66 ***	31.80 ***	31.52 ***
$\beta_{school\ quality}$		0.06		-0.04
$\beta_{school\ SES}$		0.31 ***		0.03
$\beta_{parental\ SES}$		0.10 **		0.05
Variance				
Intercept variance	523.22 ***	470.04 ***	704.24 ***	646.77 ***
SV_A	0.87 ***	0.91 ***	0.74 ***	0.74 ***
SV_C	0.06	0.00	0.13	0.10
SV_E	0.06	0.09 ***	0.13 ***	0.16 ***
Slope variance	12.19 ***	11.75 ***	21.06 ***	20.85 ***
SV_A	0.83 ***	0.83 ***	0.36	0.36
SV_C	0.02	0.00	0.39 *	0.39 *
SV_E	0.16	0.18	0.25 **	0.25 **

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Controlled for sex. SV_A, SV_C, SV_E refer to standardized genetic, shared environmental, and non-shared environmental variance, respectively. Estimates of school quality, school SES, and parental SES are standardized regression coefficients (beta's) on the intercept and slope.

Table E2.3 Estimates of the *A, C, E* variance and covariance components of the intercept and slope for reading (N_{pairs}=5,591) and mathematics (N_{pairs}=5,722), using the sample including outliers on the intercept and slope

	Reading				Mathematics			
	Total	**A**	**C**	**E**	**Total**	**A**	**C**	**E**
Unstandardized								
Intercept	544.67 ***	458.82 ***	44.72	41.13 **	765.39 ***	541.29 ***	126.11 *	97.99 ***
Slope	18.74 ***	12.74 ***	1.57	4.44 **	27.20 ***	15.83 **	6.59	4.78 *
common with intercept	3.53	1.39	1.57	0.58	6.76	0.34	4.21	2.20
unique to slope	15.21 ***	11.36 ***	0.00	3.85 **	20.44 ***	15.49 **	2.38	2.58
$COV_{I,S}$	-21.74 ***	-25.22 **	8.37 *	-4.89	-51.30 ***	-13.56	-23.05 *	-14.69 **
Standardized								
Intercept		0.84 ***	0.08	0.08 **		0.71 ***	0.17 *	0.13 ***
Slope		0.68 ***	0.08	0.24 **		0.58 **	0.24	0.18 *
common with intercept		0.07	0.08	0.03		0.01	0.16	0.08
unique to slope		0.61 ***	0.00	0.21 **		0.57 **	0.09	0.10
$COV_{I,S}$		0.66 **	0.22 *	0.13		0.26	0.45 *	0.29 **

Note. *p<.05, **p<.01, ***p<.001 (two-tailed test). Controlled for sex. *A* = genetic, *C* = shared environment, *E* = non-shared environment.

Table E2.4 Estimates of the biometric growth models including covariates for reading (N_{pairs}=5,591) and mathematics (N_{pairs}=5,722), using the sample including outliers on the intercept and slope.

	Reading		Mathematics	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Mean				
Intercept mean	125.33 ***	122.65 ***	124.77 ***	122.07 ***
$\beta_{school\ quality}$		0.02		0.05 **
$\beta_{school\ SES}$		0.05 **		0.03 *
$\beta_{parental\ SES}$		0.30 ***		0.27 ***
Slope mean	18.56 ***	17.73 ***	32.06 ***	31.64 ***
$\beta_{school\ quality}$		0.03		-0.04
$\beta_{school\ SES}$		0.10 ***		0.05
$\beta_{parental\ SES}$		0.14 ***		0.06 *
Variance				
Intercept variance	544.67 ***	491.08 ***	765.39 ***	703.64 ***
SV_A	0.84 ***	0.89 ***	0.71 ***	0.75 ***
SV_C	0.08	0.01	0.17 *	0.11
SV_E	0.08 **	0.10 ***	0.13 ***	0.15 ***
Slope variance	18.74 ***	15.58 ***	0.58 **	0.57 **
SV_A	0.68 ***	0.73 ***	0.24	0.25
SV_C	0.08	0.03	0.18 *	0.18 **
SV_E	0.24 **	0.24 **	27.20 ***	17.85 ***

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Controlled for sex. SV_A, SV_C, SV_E refer to standardized genetic, shared environmental, and non-shared environmental variance, respectively. Estimates of school quality, school SES, and parental SES are standardized regression coefficients (beta's) on the intercept and slope.

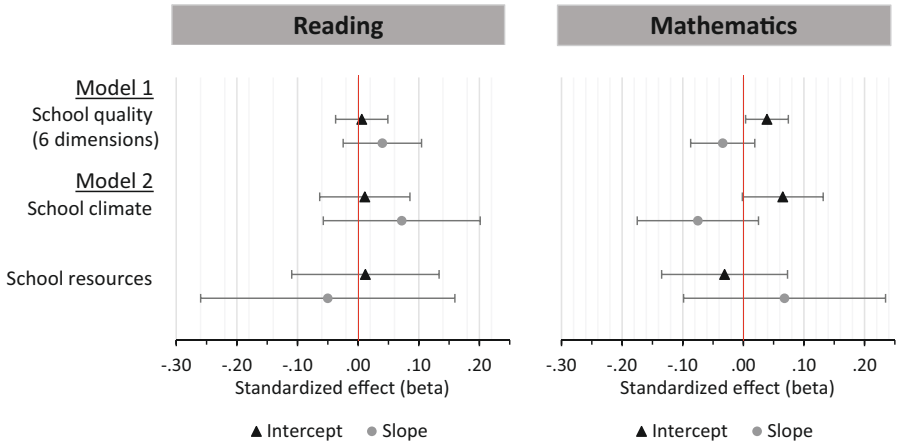


Figure E2.1 Standardized effects of school quality measures on the intercept and slope of reading and mathematics performance, including 95% confidence intervals.

Note: Controlled for sex, school SES, and parental SES.

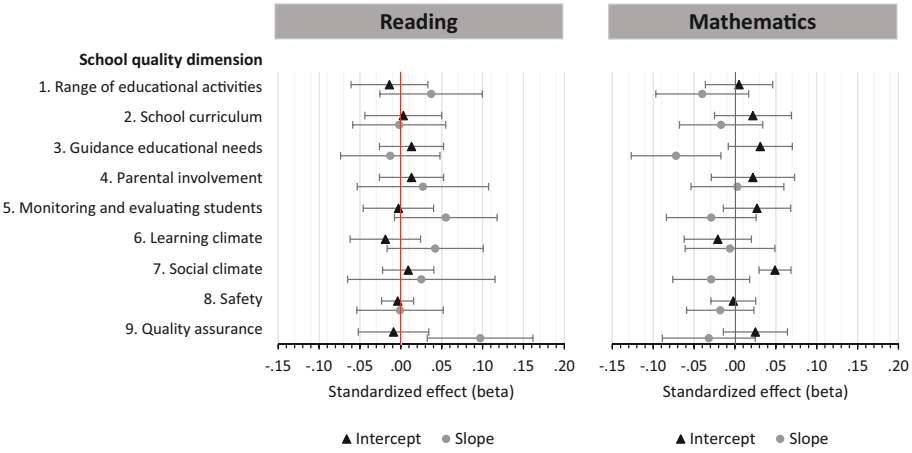


Figure E2.2 Standardized effects of school quality dimensions on the intercept and slope of reading and mathematics performance, including 95% confidence intervals.

Note: Controlled for sex, school SES, and parental SES. The results reflect nine different model (each model includes one school quality dimension and the control variables).

Table E2.5 Estimates of the *A, C, E* variance and covariance components of the intercept and slope for reading (N_{pairs}=5,304) and mathematics (N_{pairs}=5,576), using an average genetic relatedness of SS twins of .70.

	Reading				Mathematics			
	Total	**A**	**C**	**E**	**Total**	**A**	**C**	**E**
Unstandardized								
Intercept	512.31 ***	480.87 ***	25.87	5.58	694.66 ***	572.77 ***	69.55	52.34 *
Slope	12.32 ***	10.54 ***	0.11	1.67	21.59 ***	10.61	6.63	4.35 *
common with intercept	2.07	0.34	0.10	1.63	10.99 *	0.01	6.63	4.35 *
unique to slope	10.25 ***	10.20 ***	0.00	0.04	10.60 *	10.60 *	0.00	0.00
$COV_{I,S}$	-14.10 ***	-12.71 **	1.63 *	-3.02	-38.74 ***	-2.18	-21.47	-15.09 **
Standardized								
Intercept		0.94 ***	0.05	0.01		0.83 ***	0.10	0.08 *
Slope		0.86 ***	0.01	0.14		0.49	0.31	0.20 *
common with intercept		0.03	0.01	0.13		0.00	0.31	0.20 *
unique to slope		0.83 ***	0.00	0.00		0.49 *	0.00	0.00
$COV_{I,S}$		0.73 **	0.09	0.17		0.06	0.55	0.39 **

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Controlled for sex. *A* = genetic, *C* = shared environment, *E* = non-shared environment.

Table E2.6 Estimates of the *A, C, E* variance and covariance components of the intercept and slope for reading (N_{pairs}=5,304) and mathematics (N_{pairs}=5,576), using an average genetic relatedness of SS twins of .80.

	Reading				Mathematics			
	Total	**A**	**C**	**E**	**Total**	**A**	**C**	**E**
Unstandardized								
Intercept	512.77 ***	367.54 ***	73.17 *	72.06 ***	694.21 ***	406.88 ***	145.91 **	141.42 ***
Slope	12.32 ***	9.47 ***	0.18	2.68 **	21.58 ***	7.68	7.96 **	5.95 ***
common with intercept	0.89	0.55	0.18	0.16	6.94	0.06	4.81	2.07 *
unique to slope	11.44 ***	8.91 ***	0.00	2.52 **	14.65 ***	7.62	3.14	3.88 ***
$COV_{I,S}$	-14.07 ***	-14.24 *	3.58	-3.42	-38.81 ***	4.78	-26.50 *	-17.09 ***
Standardized								
Intercept		0.72 ***	0.14 *	0.14 ***		0.59 ***	0.21 **	0.20 ***
Slope		0.77 ***	0.01	0.22 **		0.36	0.37 **	0.28 ***
common with intercept		0.05	0.01	0.01		0.00	0.22	0.10 *
unique to slope		0.72 ***	0.00	0.21 **		0.35	0.15	0.18 ***
$COV_{I,S}$		0.67 *	0.17	0.16		0.08	0.63 *	0.29 ***

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Controlled for sex. *A* = genetic, *C* = shared environment, *E* = non-shared environment.

Table E2.7 Estimates of the biometric growth models including covariates for reading (N_{pairs}=5,304) and mathematics (N_{pairs}=5,576), using an average genetic relatedness of SS twins of .70 and .80.

	Reading (rSS_G = .70)				Reading (rSS_G = .80)			
	Model 1		Model 2[a]		Model 1		Model 2	
Intercept								
$\beta_{school\ quality}$							0.02	
$\beta_{school\ SES}$							0.04	**
$\beta_{parental\ SES}$							0.31	***
SV_A	0.94	***			0.72	***	0.80	***
SV_C	0.05				0.14	*	0.04	
SV_E	0.01				0.14	***	0.16	***
Total variance	512.31	***			512.77	***	458.49	***
Slope								
$\beta_{school\ quality}$							0.04	
$\beta_{school\ SES}$							0.11	***
$\beta_{parental\ SES}$							0.11	***
SV_A	0.86	***			0.77	***	0.76	***
SV_C	0.01				0.01		0.00	
SV_E	0.14				0.22	**	0.24	**
Total variance	12.32	***			12.32	***	11.85	***

	Mathematics (rSS_G = .70)				Mathematics (rSS_G = .80)			
	Model 1		Model 2		Model 1		Model 2	
Intercept								
$\beta_{school\ quality}$			0.05	**			0.05	**
$\beta_{school\ SES}$			0.04	*			0.04	*
$\beta_{parental\ SES}$			0.27	***			0.27	***
SV_A	0.83	***	0.80	***	0.59	***	0.63	***
SV_C	0.10		0.09		0.21	**	0.15	*
SV_E	0.08	*	0.11	**	0.20	***	0.23	***
Total variance	694.66	***	638.27	***	694.21	***	635.99	***
Slope								
$\beta_{school\ quality}$			-0.04				-0.04	
$\beta_{school\ SES}$			0.04				0.04	
$\beta_{parental\ SES}$			0.06	*			0.06	*
SV_A	0.49		0.50	*	0.36		0.35	
SV_C	0.31		0.30		0.37	**	0.37	**
SV_E	0.20	*	0.20	*	0.28	***	0.28	***
Total variance	21.59	***	21.33	***	21.58	***	21.32	***

Note: *p<.05, **p<.01, ***p<.001 (two-tailed test). Controlled for sex. SV_A, SV_C, SV_E refer to standardized genetic, shared environmental, and non-shared environmental variance, respectively. Estimates of school quality, school SES, and parental SES are standardized regression coefficients (beta's) on the intercept and slope.

[a] Model does not converge

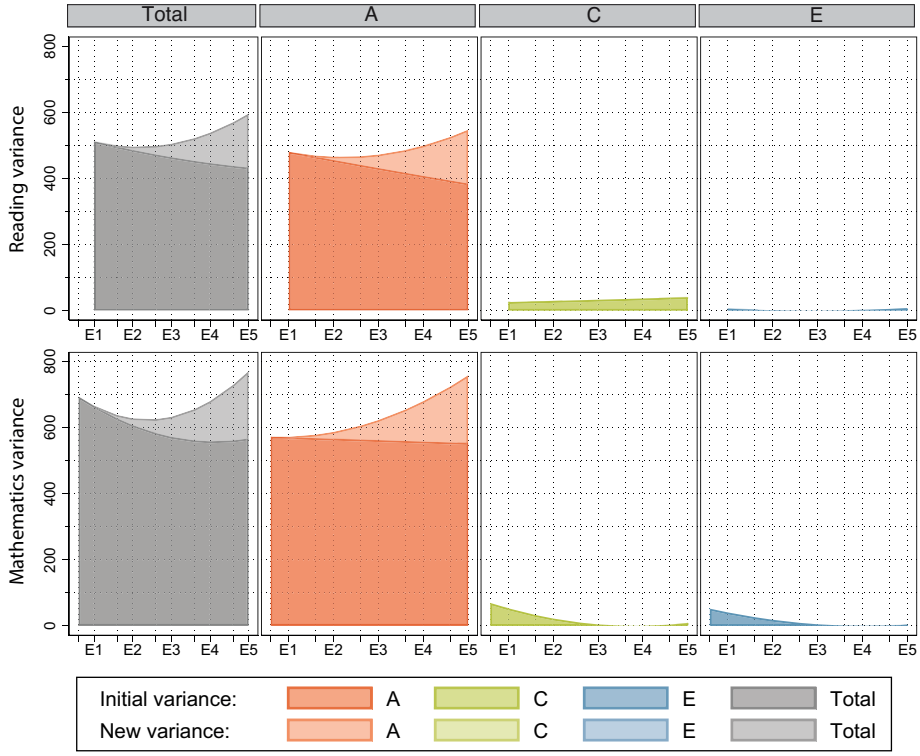


Figure E2.3 The development of the total, genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variance in reading and mathematics performance, using rSS_G = .70.

Note: Controlled for sex. Based on the variance in the intercept and slope and their covariance, reported in Table E2.5.

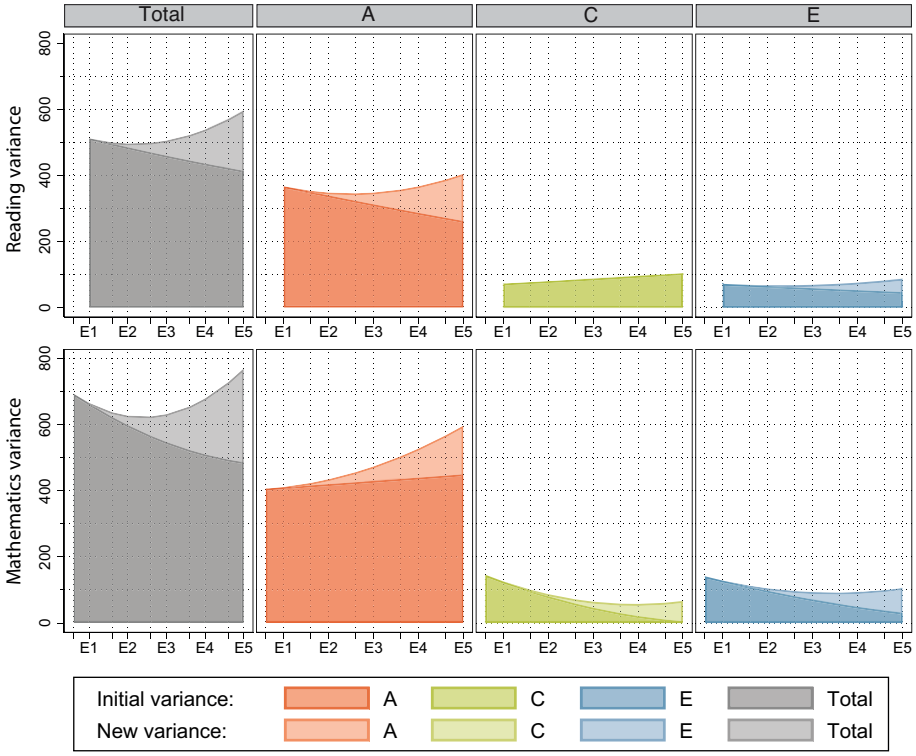


Figure E2.4 The development of the total, genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) variance in reading and mathematics performance, using rSS_G = .80.

Note: Controlled for sex. Based on the variance in the intercept and slope and their covariance, reported in Table E2.6.

E.3. Measurement models

E.3.1. Measurement model socioeconomic status (SES)

Table E3.1 Descriptive statistics and factor loadings for parental SES (*N* = 373,956).

	N	**M**	**SD**	**Factor loadings SES**	
ISCED father	260,228	4.39	2.08	.69	(<.01)
ISCED mother	298,072	4.35	2.07	.82	(<.01)
Income (percentile) father	346,141	73.76	23.21	.50	(<.01)
Income (percentile) mother	336,737	49.05	25.68	.62	(<.01)

Note: All factor loadings are statistically significant (p <.001, two-tailed test). Standard error in parentheses. Standardized effects.

E.3.2. Measurement model school quality

The same measurement model as in Chapter 3 is used. See Appendix C.1. for the details.

E.4. Model fit latent growth models

Table E4.1 Model fit of the biometric latent growth models with different residual specifications for reading (N_{pairs} = 5,304) and mathematics (N_{pairs} = 5,576).

	Model fit			Model comparison			
	Loglikelihood	Freely estimated parameters	AIC	Comparison	Chi²	Δdf	p
Reading							
1. Uncorrelated	-190,003.44	24	380,054.89	-	-	-	-
2. Correlated (*)	-189,547.50	42	379,179.00	1	911.88	18	<.001
3. Correlated, same over time	-189,695.81	25	379,441.62	2	296.62	17	<.001
4. Correlated, same for SS/OS	-189,567.80	33	379,201.60	2	40.60	9	<.001
Mathematics							
1. Uncorrelated	-335,829.34	25	671,708.68	-	-	-	-
2. Correlated (*)	-333,002.15	45	666,094.30	1	4125.42	20	<.001
3. Correlated, same over time	-334,177.15	26	668,406.30	2	1531.96	19	<.001
4. Correlated, same for SS/OS	-333,041.60	35	666,153.20	2	88.02	10	<.001

Note: Controlled for sex. Model indicated by (*) is the preferred model.

Table E4.2 Model fit of the latent growth models with linear and quadratic slope for reading (N_{pairs} = 5,304) and mathematics (N_{pairs} = 5,576).

	Model fit			Model comparison		
	Loglikelihood	Freely estimated parameters	AIC	Chi²	Δdf	p
Reading						
1. Linear	-189,547.50	42	379,179.00	-	-	-
2. Quadratic (*)	-189,403.10	47	378,900.20	288.80	5	<.001
Mathematics						
1. Linear	-270,703.44	45	541,496.89	-	-	-
2. Quadratic (*)	-267,605.12	50	535,310.23	6196.65	5	<.001

Note: Controlled for sex. Model indicated by (*) is the preferred model.

E.5. Supplementary tables and figures

Table E5.1. Path estimates of the biometric linear latent growth model for reading (N_{pairs}=5,304) and mathematics (N_{pairs}=5,576).

	Reading		Mathematics	
	estimate	**s.e.**	**estimate**	**s.e.**
Intercept	125.57 ***	(0.50)	125.93 ***	(0.51)
Slope	18.36 ***	(0.14)	31.78 ***	(0.16)
a11	-20.79 ***	(1.14)	22.10 ***	(1.69)
c11	6.51 *	(2.52)	10.26 ***	(2.53)
e11	6.18 ***	(1.34)	10.04 ***	(1.24)
a21	0.65 *	(0.32)	0.26	(0.72)
c21	0.39	(0.58)	-2.63 **	(0.86)
e21	-0.52	(0.51)	-1.75 ***	(0.44)
a22	3.09 ***	(0.24)	-3.02 ***	(0.83)
c22	0.00	(0.00)	-0.52	(3.96)
e22	1.39 ***	(0.32)	-1.46 **	(0.50)
Loglikelihood	-188635.13		-269,569.86	
Freely estimated parameters	48		51	
AIC	377366.27		539,241.71	

Note. *p<.05, **p<.01, ***p<.001 (two-tailed test). *a* = genetic, *c* = shared environment, *e* = non-shared environment, AIC = Akaike Information Criterion. The paths *a11*, *c11*, and *e11* reflect the genetic and environmental effects on the intercept, and *a21*, *c21*, and *e21* show to what extent these same influences also predict the slope. The paths *a22*, *c22*, and *e22* show the effects that are independent of the intercept and thus uniquely influence the slope.

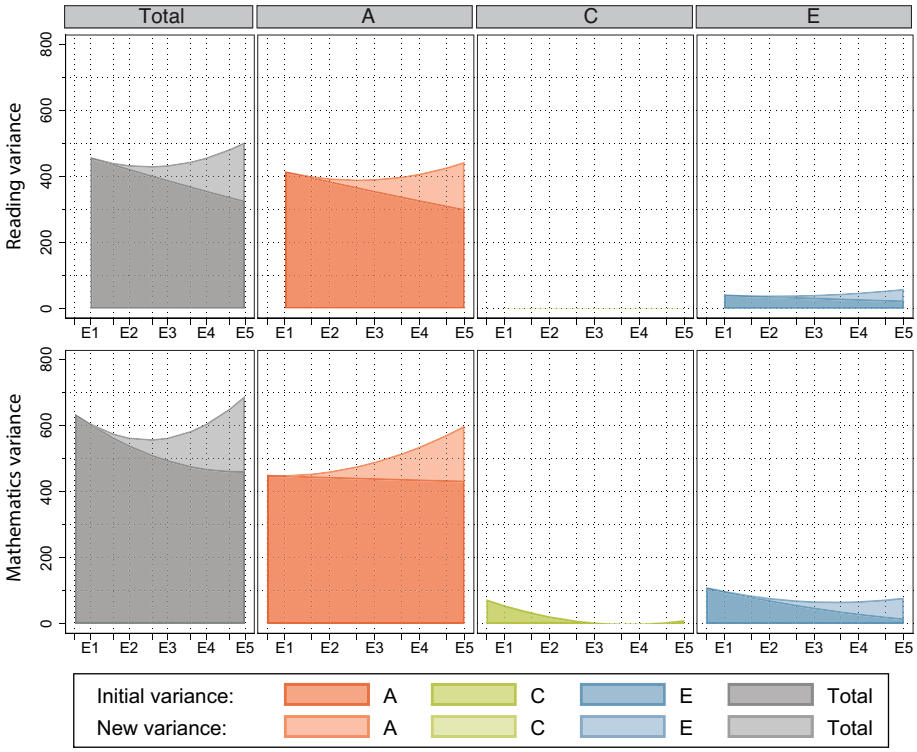


Figure E5.1 The development of the total, genetic (A), shared environmental (C), and non-shared environmental (E) variance in reading and mathematics performance.

Note: Controlled for sex, school quality, school SES, and parental SES.

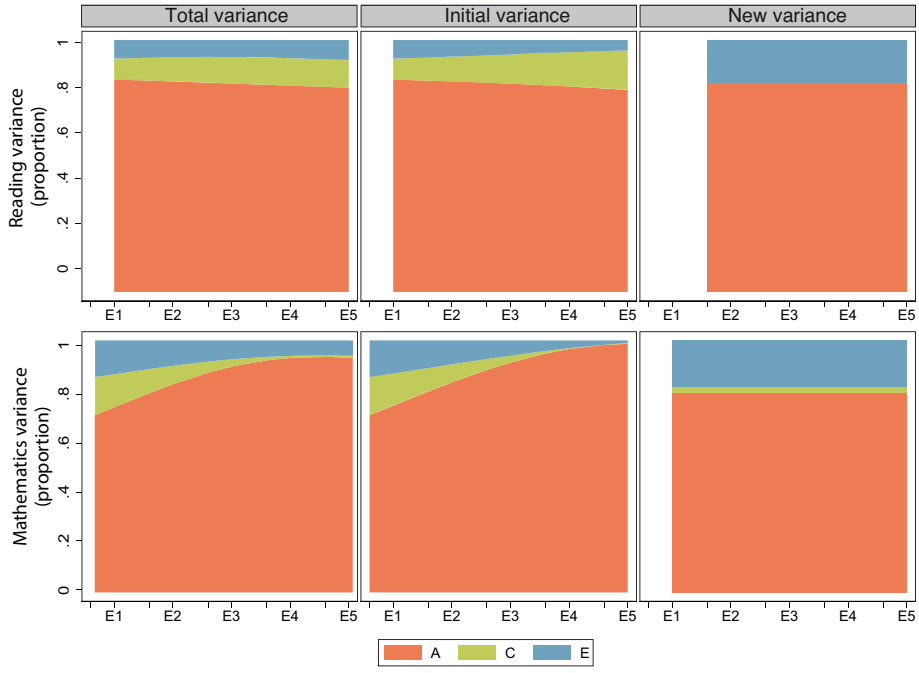


Figure E5.2 The development the relative contributions of genetic (A), shared environmental (C), and non-shared environmental (E) variance to the total variance, initial variance, and new variance in reading and mathematics performance.

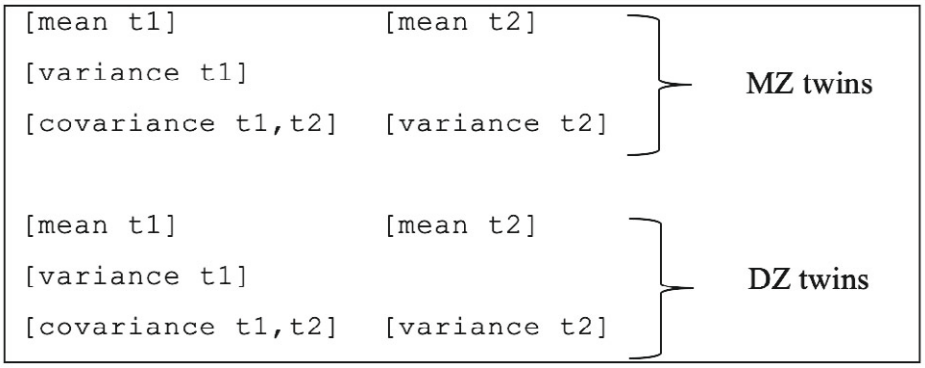
Note: Controlled for sex.

DATA AND CODE AVAILABILITY

Usage of summary data

All chapters rely on data that are not publicly available but may be accessed by researchers under certain conditions. In addition to the raw (individual level) data, results can be partly reproduced using summary data which can be publicly shared. Mplus requires a summary data input file including the means, variances, and covariances of all variables. In case of multi-group analysis, as is the case with twin models, these data have to be provided by group. In addition, the number of observations needs to be included in the analysis syntax. Example input data is presented in Box A (in .txt format). The summary data files for each chapter can be found on my OSF profile (https://osf.io/xch24/).

Box A. Summary data input structure for Mplus.



Note: t1 = twin 1, t2 = twin 2.

In Box B, an example is provided how summary data can be used to obtain the ACE decomposition in Mplus. The results show this for the univariate twin model, but extensions are possible such as bivariate models or moderation models with a multigroup setup (i.e., to test for non-parametric gene-environment interactions). Unfortunately, analysis with a continuous linear moderator requires raw data in Mplus.

Box B. Example syntax using summary data in Mplus.

```
TITLE:     Univariate ACE decomposition
DATA:
      FILE = summdata.txt ;
      TYPE = means covariance ;
      NGROUPS = 2 ;
      NOBS = 3061 18673 ;
VARIABLE:
      NAMES = v1 v2 ;
MODEL:
! Constrain means and variances to be the same for both groups
[v1 v2] (m) ;        ! Constrain means and variances to be the same
 v1 v2  (v) ;        ! for both twins of both zygosity groups

MODEL G1:
v1 with v2 (cmz) ;   ! Covariance MZ twins (or SS twins)

MODEL G2:
v1 with v2 (cdz) ;   ! Covariance DZ twins (or OS twins)

MODEL CONSTRAINT:
    new(a c e x y z) ;
    a=x*x ;
    c=y*y ;
    e=z*z ;
    v=a+c+e ;
    cmz= 1*a+c ;      ! Change 1 to rSSg (e.g., .75) for SS twins
    cdz=.5*a+c ;
```

Data and code for the empirical chapters

Chapter 2

Individual level data is provided by the Netherlands Twin Register. These data may be accessed upon reasonable request and after approval of the data access committee. For more information, see https://ntr-data-request.psy.vu.nl/index.html.

Code for replication using the individual level data can be found at OSF: https://osf.io/mxzpe/. With slight modifications (see Box B) the same code can be applied to the summary data, except the modelling of linear moderation. The summary data can also be found at OSF.

Chapter 3

All results are based on own calculations using non-public micro data from Statistics Netherlands (CBS). These data are under certain conditions accessible for scientific research. Further information and the procedure to request the data can be found on the websites of CBS (https://www.cbs.nl/en-gb/onze-diensten/customised-services-microdata/microdata-conducting-your-own-research). School quality data were obtained via a user agreement with the Dutch Inspectorate of Education. Quality data are partly publicly available (from school year 2015/2016 onward) via https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/onderwijsdata/kwaliteitsindicatoren.

Code for replication using the individual level data can be found at OSF: https://osf.io/xsgdt/. With slight modifications (see Box B) the same code can be applied to the summary data, except the modelling of linear moderation. The summary data can also be found at OSF.

Chapter 4

The administrative register data can be accessed via the Research Services at Statistics Denmark but requires formal affiliation with a Danish research institution. Interested researchers are welcome to reach out to Kristian Bernt Karlson (kbk@soc.ku.dk) to learn more about this opportunity and how the different registers have been recoded and combined in order to produce the data file used in this chapter.

Code for replication using the individual level data can be found at OSF: https://osf.io/unm9b/. With slight modifications (see Box B) the same code can be applied to the summary data. The summary data can also be found at OSF.

Chapter 5

All results are based on own calculations using non-public micro data from CBS and the Netherlands Cohort Study on Education (NCO). Data are accessible for statistical and scientific research under certain conditions. Further information and the procedure to request the data can be found on the websites of CBS (https://www.cbs.nl/en-gb/onze-diensten/customised-services-microdata/microdata-conducting-your-own-research) and NCO (https://www.nationaalcohortonderzoek.nl/onderzoek). School quality data were obtained via a user agreement with the Dutch Inspectorate of Education. Quality data are partly publicly available (from school year 2015/2016 onward) via https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/onderwijsdata/kwaliteitsindicatoren.

Code for replication using the individual level data can be found at OSF: https://osf.io/kc9g3/. With slight modifications (see Box B) the same code can be applied to the summary data. The summary data can also be found at OSF.

NEDERLANDSE SAMENVATTING

Summary in Dutch

Onderwijskwaliteit en ongelijkheid: De wisselwerking tussen scholen, gezinnen en genen

Achtergrond

In veel landen staan de kwaliteit en gelijkheid van onderwijs onder druk. Te veel kinderen verlaten de basisschool zonder voldoende basisvaardigheden (Unicef, 2020; United Nations, 2019). Daarnaast zijn er indicaties dat de onderwijsprestaties van kinderen steeds meer afhankelijk zijn geworden van de sociaaleconomische status (SES) van hun ouders en van de school die ze bezoeken (Chmielewski, 2019; Inspectie van het Onderwijs, 2017; Unicef Office of Research, 2018). Recente ontwikkelingen, zoals de COVID-19-pandemie en het groeiende tekort aan leraren en schoolleiders, brengen verdere uitdagingen met zich mee voor onderwijskwaliteit en -ongelijkheid. Het bestuderen van de rol van scholen is cruciaal om onderwijsongelijkheid te begrijpen en mogelijk te verminderen. Sociologisch onderzoek heeft zich hier al decennia op gericht (Coleman et al., 1966). Desondanks blijft veel onduidelijk over hoe scholen en onderwijsongelijkheid met elkaar samenhangen (Downey & Condron, 2016).

Twee redenen dragen bij aan deze onduidelijkheid. De eerste reden heeft betrekking op de conceptuele vraag wat onderwijsongelijkheid inhoudt. Ongelijkheid betekent verschil, maar welke verschillen maken deel uit van onderwijsongelijkheid en worden als zodanig geproblematiseerd? Hierover bestaan verschillende perspectieven. Men kan zich richten op de totale prestatieverschillen, die de spreiding (of prestatiekloof) tussen de slechtst en best presterende leerlingen omvat. Dit type ongelijkheid wordt in dit proefschrift aangeduid als 'spreiding' (*inequality as dispersion*) (Van de Werfhorst & Mijs, 2010). Men kan zich ook richten op specifieke oorzaken van verschillen. Een vaak bestudeerde oorzaak die ten grondslag ligt aan prestatieverschillen is de sociaaleconomische achtergrond van kinderen. SES-verschillen in prestaties wordt vaak aangeduid als 'kansenongelijkheid' (*inequality of opportunity*) of 'sociale ongelijkheid' (Strello et al., 2021; Strietholt, 2014; Van de Werfhorst & Mijs, 2010). Kinderen hebben zelf ook weinig controle over de schoolomgeving waarin zij zich bevinden, net zoals ze ook geen controle hebben over hun gezinsachtergrond. Daarom kunnen prestatieverschillen tussen scholen ook gezien worden als problematisch. Maar het is onduidelijk of schoolverschillen nog steeds als problematisch gezien worden als ze een compenserende rol spelen en bijvoorbeeld SES-verschillen doen verminderen.

Bij de conceptuele vraag wat onderwijsongelijkheid inhoudt, is daarnaast de rol van genetische verschillen in onderwijsprestaties, soms ook wel 'genetische ongelijkheid genoemd', niet eenduidig. Terwijl sociale verschillen (bijvoorbeeld SES-verschillen) in onderwijsprestaties vaak worden gezien als oneerlijk, is er minder consensus over of de spreiding als gevolg van genetische verschillen problematisch is. Enerzijds weerspiegelen genetische verschillen variatie in aangeboren potentie, wat vaak wordt gezien als een legitieme oorzaak van prestatieverschillen (Dias Pereira, 2021; Tannock, 2008). Als er geen sociale barrières zijn die de expressie van aangeboren potentie belemmeren, zal een groter deel van de verschillen in onderwijsprestaties worden verklaard door genetische verschillen. Daarom kan de genetische bijdrage aan prestatieverschillen worden gezien als een indicatie van kansengelijkheid (Nielsen, 2006). Anderzijds is het zo dat individuen geen controle hebben over hun aangeboren potentie, net zoals zij geen controle hebben over hun

sociale achtergrond. Geluk in de genetische loterij kan daarom ook worden gezien als oneerlijk en in strijd met het idee van gelijke kansen (Harden, 2021).

De tweede reden waarom de rol van scholen in onderwijsongelijkheid nog onduidelijk is, is dat weinig studies alle drie de aspecten onderzoeken: scholen, gezinnen en genen. Zelfs als er consensus is over welke invloeden op onderwijsprestaties deel uitmaken van onderwijsongelijkheid, blijft het nog een empirische vraag hoe scholen, gezinnen en genen samen een rol spelen. Het niet gelijktijdig bestuderen van deze factoren is problematisch om twee redenen. Ten eerste kunnen de invloeden met elkaar worden verward omdat ze met elkaar correleren. Bijvoorbeeld, ouders die zelf goed presteerden op school deden dit mede vanwege hun genetische aanleg. Deze ouders geven deze genen deels door aan hun kinderen, maar ze zorgen ook vaker voor een stimulerende leeromgeving thuis en kiezen vaker voor scholen van hogere kwaliteit voor hun kinderen. Ten tweede kunnen de verschillende factoren met elkaar interacteren. Scholen van hoge kwaliteit kunnen bijvoorbeeld verschillen in prestaties op basis van gezinsachtergrond verminderen (i.e., compensatie), maar tegelijkertijd genetische verschillen tussen kinderen vergroten (i.e., multiplicatie). Eerdere studies vertellen slechts een deel van het verhaal omdat ze de interactie met ofwel genetische ofwel omgevingsfactoren negeren.

Zowel de sociologie als gedragsgenetica bieden verschillende theorieën voor multiplicatie- en compensatie-effecten. *Multiplicatie* vergroot prestatieverschillen, bijvoorbeeld omdat de school- en thuisomgeving elkaars invloed versterken. Kinderen uit hoge-SES-gezinnen profiteren mogelijk meer van scholen van hoge kwaliteit, omdat ze de school beginnen met meer vaardigheden en academische voorbereiding van huis uit (Hanselman, 2018; Heckman, 2000). Ook vanuit het idee van culturele reproductie kan worden verwacht dat hoge-SES-kinderen bevoordeeld worden door leerkrachten en zich meer thuis voelen op school, wat leidt tot betere onderwijsprestaties (Bourdieu, 1986; Bourdieu & Passeron, 1977; De Graaf et al., 2000). Genetische en omgevingsinvloeden kunnen elkaar eveneens versterken. Bronfenbrenners bio-ecologisch model stelt dat stabiele, hulpbronrijke omgevingen met continue wederkerige interacties genetische potentie tot uiting brengen (Bronfenbrenner, 1992; Bronfenbrenner & Ceci, 1994). Hieruit volgt de voorspelling dat genetische potentie voor betere onderwijsprestaties meer tot uiting komt in hoge-SES-gezinnen en hogekwaliteitscholen.

Bij *compensatie* worden prestatieverschillen juist kleiner. De thuis- en schoolomgeving kunnen elkaar compenseren, omdat de leermogelijkheden in hoge-SES-gezinnen overlappen met die in hogekwaliteitscholen (Chiu & Khoo, 2005; Hanselman, 2018). Kinderen uit lage-SES-milieus profiteren mogelijk meer van goede scholen (zoals het academische klimaat en de hogere niveaus van motivaties en aspiraties), omdat ze die minder vaak thuis ervaren (Coleman et al., 1966; Rumberger & Palardy, 2005). Andersom kunnen hoge SES-ouders mogelijk beter compenseren voor een lagere schoolkwaliteit, bijvoorbeeld door extra begeleiding en bijles (Hanselman, 2018). Een hoge-SES-thuisomgeving of -schoolomgeving zouden ook kunnen compenseren voor genetische invloeden die anders tot lagere prestaties leiden. Volgens het kwetsbaarheid-stressmodel komt een genetische kwetsbaarheid (zoals leerproblemen) eerder tot uiting bij meer risicofactoren en stressoren in de omgeving (Rende & Plomin, 1992; Shanahan & Hofer, 2005). Positieve omgevingskenmerken, zoals vaker aanwezig in hoge-SES-gezinnen en hogekwaliteitscholen, compenseren voor genetische kwetsbaarheden of risico's (see Shanahan

& Hofer, 2005). Dit komt overeen met het compenserend voordeel (*compensatory advantage*) mechanisme, dat stelt dat eerdere negatieve uitkomsten of capaciteiten (bijvoorbeeld lagere cognitieve capaciteiten of een slechte gezondheid) minder negatieve consequenties hebben voor de onderwijsprestaties van kinderen uit hoge-SES-gezinnen (Bernardi, 2014).

In dit proefschrift stel ik de vraag: *Hoe draagt de wisselwerking tussen scholen, gezinnen en genen bij aan onderwijsongelijkheid*? Terwijl sociologisch onderzoek de invloeden van scholen, gezinnen en hun interactie op onderwijsresultaten bestudeert, wordt de rol van genetische verschillen tussen leerlingen vaak buiten beschouwing gelaten. Vanuit de gedragsgenetica worden genetische verschillen en omgevingsverschillen in onderwijsresultaten onderzocht. Gedragsgenetici bestuderen onder andere in hoeverre deze genetische verschillen en omgevingsverschillen afhangen van de thuisomgeving (zoals ouderlijke SES), maar besteden minder aandacht aan de schoolomgeving. Bovendien blijken deze studies vaak voorbij te gaan aan de complexiteit en diversiteit van omgevingsomstandigheden (Baier, 2019). Hoewel gedragsgenetische studies onderzoeken hoe genetische invloeden afhankelijk zijn van zowel de gezins- als de schoolsituatie, wordt er zelden rekening gehouden met dat kinderen tegelijkertijd in zowel gezinnen als op scholen zijn ingebed. Hierdoor blijft het grotendeels onbekend hoe scholen, gezinnen en genen samen functioneren.

Met dit proefschrift beoog ik beter te begrijpen hoe de wisselwerking tussen schoolkenmerken, SES van de ouders en genen onderwijsprestaties beïnvloedt en wat de implicaties hiervan zijn voor onderwijsongelijkheid. Hoewel mijn focus ligt op het theoretische en empirische deel van deze vraag, bespreek ik ook het conceptuele deel aangezien de definitie van onderwijsongelijkheid gevolgen heeft voor de interpretatie van de resultaten. Ik onderzoek de wisselwerking tussen scholen, families en genen door sociologische theorieën te combineren met theorieën en methoden uit de gedragsgenetica.

Vanuit de gedragsgenetica wordt het onderzoeken van de invloed van genen en omgeving op onderwijsprestaties vaak gedaan door de gelijkenis in prestaties tussen identieke tweelingen (die 100% van hun genen delen) en niet-identieke tweelingen (die ~50% van hun genen delen) te vergelijken. Als identieke tweelingen vergelijkbaardere onderwijsprestaties hebben dan niet-identieke tweelingen, wijst dit erop dat onderwijsprestaties genetisch beïnvloed worden. Op basis van dit idee kunnen verschillen in onderwijsprestaties worden opgesplitst in drie componenten: een genetisch component, een gedeelde omgevingscomponent (dat alle omgevingsinvloeden omvat die tweelingen op elkaar doen lijken) en een niet-gedeelde omgevingscomponent (omgevingsinvloeden die tweelingen van elkaar doen verschillen, inclusief meetfouten). Ik pas variaties van de methode toe op de onderwijsprestaties van kinderen, waaronder gestandaardiseerde testresultaten en cijfergemiddelden. Hiervoor werk ik met gegevens van het Nederlands Tweelingen Register (Hoofdstuk 2), Nederlandse registerdata van het Centraal Bureau voor de Statistiek (CBS) aangevuld met data van de Inspectie voor het Onderwijs (Hoofdstuk 3 en 5), en Deense registerdata (Hoofdstuk 4).

Samenvatting van de resultaten per hoofdstuk

Hoofdstuk 2: Klasomgeving vermindert sociale ongelijkheid

In hoofdstuk 2 onderzoek ik of de thuis- en schoolklasomgeving elkaars invloed versterken of verzwakken bij het verklaren van onderwijsprestaties. Ik maak gebruik van het feit dat sommige tweelingen in dezelfde klas zitten terwijl andere in aparte klassen zitten. Op deze manier kan ik de bijdrage van de klasomgeving onderzoeken naast die van de andere omgevings- en genetische bronnen van verschillen. Gemiddeld kan slechts een klein deel (2%) van de prestatieverschillen in de score op de Cito-eindtoets verklaard worden door de klasomgeving. Voor kinderen wiens ouders hoogstens basisonderwijs hebben voltooid, is de klasomgeving echter verantwoordelijk voor bijna 8% van de variantie in onderwijsprestaties. Dit daalt naar 1% voor kinderen van ouders met een postdoctorale opleiding. De klasomgeving (bijvoorbeeld kwaliteit van de leraar, pedagogisch klimaat, invloed van klasgenoten) is dus belangrijker voor de prestaties van kinderen uit lagere-SES-milieus. Dit suggereert dat gunstige klascontexten ongelijkheid met betrekking tot sociaaleconomische achtergrond kunnen compenseren.

Hoofdstuk 3: Schoolkwaliteit vermindert onderwijsongelijkheid niet

In hoofdstuk 3 onderzoek ik in hoeverre genetische en omgevingsverschillen in onderwijsprestaties afhankelijk zijn van schoolkwaliteit. De resultaten tonen aan dat kinderen op scholen met een hogere kwaliteit, gemeten aan de hand van vele indicatoren gerelateerd aan schoolmiddelen en schoolklimaat, enigszins hogere Cito-scores hebben, maar dat de effectgrootte klein is. Daarentegen hangt kansenongelijkheid niet af van schoolkwaliteit. Ook de spreiding in prestaties hangt niet af van schoolkwaliteit zodra rekening wordt gehouden met het feit dat hogeropgeleide ouders scholen van hogere kwaliteit kiezen. Niet schoolkwaliteit, maar de SES-compositie van de school en de SES van de ouders zijn gerelateerd aan spreiding. De totale spreiding en spreiding als gevolg van genetische en niet-gedeelde omgevingsverschillen zijn lager in scholen en gezinnen met een hogere SES. Het verminderen van kwaliteitsverschillen tussen scholen is daarom waarschijnlijk niet voldoende om onderwijsongelijkheid te verminderen.

Hoofdstuk 4: De wisselwerking tussen gender, gezinsachtergrond en schoolcontext

In hoofdstuk 4 worden gen-omgevingsinteracties in onderwijsprestaties onderzocht met verdere bijkomende complexiteiten. Genetische en omgevingsinvloeden op onderwijsprestaties kunnen afhangen van de SES van het gezin en de schoolpopulatie, maar de rollen van gezinnen en scholen zijn daarbij niet noodzakelijk onafhankelijk en kunnen met elkaar interacteren. Bovendien kan een gen-omgevingsinteractie in onderwijsprestaties anders werken voor jongens dan voor meisjes. Ik onderzoek de gemiddelde examencijfers van leerlingen in Denemarken, en vind een interactie tussen genen en SES van het gezin, maar geen interactie tussen genen en de SES-compositie van de school. In hoge-SES-gezinnen spelen genen een kleinere rol bij het verklaren van verschillen in onderwijsprestaties. Deze gen-SES-interactie hangt af van geslacht. In hoge-SES-gezinnen blijkt de rol van genen aanzienlijk lager voor jongens dan voor meisjes, waardoor ik alleen een gen-SES-interactie vind voor jongens. Tot slot blijkt de modererende rol van de SES van het

gezin voor jongens vrijwel alleen op te gaan voor kinderen die naar lage-SES-scholen gaan. Een mogelijke interpretatie van deze bevindingen is dat als jongens een hoger genetische risico voor lagere onderwijsprestaties hebben dit wordt gecompenseerd in hoge-SES-gezinnen, vooral wanneer ze een lage-SES-school bezoeken. De bevinding dat de interactie tussen genen en SES van het gezin meer uitgesproken is in lage-SES-scholen komt overeen met het idee dat gezin- en schoolinvloeden elkaar compenseren.

Hoofdstuk 5: De ontwikkeling van ongelijkheid tijdens de basisschool

In hoofdstuk 5 bestudeer ik de ontwikkeling van spreiding in prestaties gedurende de basisschoolperiode (groep 3 t/m 7) in Nederland. Ik onderzoek (1) of genetische en omgevingsverschillen in taal- en rekenprestaties worden gereproduceerd, gemultipliceerd of gecompenseerd gedurende de basisschoolperiode, en (2) in hoeverre prestatieverschillen kunnen worden verklaard door schoolkwaliteit, SES-compositie van de school en SES van het gezin. De resultaten tonen allereerst aan dat de spreiding in prestaties aan het begin van formeel onderwijs wordt gecompenseerd gedurende de basisschoolperiode. Zowel voor lees- als rekenprestaties spelen genetische verschillen een grote rol bij de initiële spreiding in prestaties. Voor lezen nemen deze genetische verschillen af terwijl omgevingsverschillen grotendeels worden gereproduceerd gedurende de schoolloopbaan. Het omgekeerde is het geval voor rekenen; de afname in spreiding wordt veroorzaakt door afnemende omgevingsverschillen, terwijl genetische verschillen worden gereproduceerd. Ten tweede komen in de loop van de tijd nieuwe invloeden naar voren die zorgen voor prestatieverschillen. Dit zijn voornamelijk nieuwe genetische invloeden, wat resulteert in een toename van de totale spreiding in prestaties over de tijd. Ten derde verklaren gemeten schoolkenmerken de spreiding niet, wat suggereert dat kwaliteitsverschillen tussen scholen waarschijnlijk een beperkte rol spelen.

Vier conclusies

In dit proefschrift heb ik vier empirische studies uitgevoerd om antwoord te geven op de vraag: *Hoe draagt de wisselwerking tussen scholen, gezinnen en genen bij aan onderwijsongelijkheid?* Hieruit kunnen vier overkoepelende conclusies worden getrokken.

1. Eerder compensatie dan multiplicatie

Scholen, gezinnen en genen blijken elkaar eerder te compenseren dan multipliceren. Des te hoger de SES van de ouders, des te minder de klasomgeving ertoe doet en des te minder genetische verschillen er zijn met betrekking tot onderwijsprestaties. Ook voor school SES geldt dat naarmate een groter aandeel van de leerlingen op een school een hogere SES heeft, omgevingsverschillen en genetische verschillen vaak kleiner zijn. Deze resultaten laten zien dat de school- en klasomgeving de potentie hebben om gedeelde omgevings- en genetische verschillen in onderwijsprestaties te verminderen. Over het algemeen zijn deze compenserende gen-omgevingsinteracties met betrekking tot gezinnen sterker dan die met betrekking tot scholen.

2. Schoolkwaliteit speelt geen rol, de klasomgeving en de SES-compositie van de school wel

De schoolomgeving speelt een rol in onderwijsongelijkheid, maar de kwaliteit van scholen is niet zo belangrijk als vaak wordt beweerd (Inspectie van het Onderwijs, 2017; Jennings et al., 2015). De mate van kansenongelijkheid en spreiding blijkt niet af te hangen van schoolkwaliteit. Dit betekent echter niet dat scholen onbelangrijk zijn voor onderwijsongelijkheid. Een groter aandeel leerlingen met een hoge-SES op school hangt samen met minder spreiding van onderwijsprestaties, vooral minder genetische spreiding. Bovendien blijkt de klasomgeving een rol te spelen, aangezien lage-SES kinderen meer profiteren van de klasomgeving. Verschillen in de SES-compositie van de school en de klasomgeving zijn dus op een compenserende manier gerelateerd aan verschillende aspecten van onderwijsongelijkheid.

3. SES van het gezin speelt een rol, maar anders dan vaak gedacht

Er wordt vaak beweerd dat een aanzienlijk deel van de verschillen in onderwijsprestaties voortkomt uit de verschillende opvoedingssituaties die samenhangen met ouders' SES (Sirin, 2005). De invloed van SES op onderwijsprestaties wordt meestal gezien als uniform, waarbij SES op dezelfde manier van invloed is op kinderen binnen hetzelfde gezin (cf. Freese, 2008). Dit proefschrift toont aan dat dit slechts een klein deel van het verhaal is. Ten eerste blijkt de gedeelde omgevingsvariantie in onderwijsprestaties relatief klein is, wat impliceert dat gezins-SES minder een exogene omgevingsinvloed is dan gedacht. Ten tweede tonen de resultaten aan dat genetische invloeden kleiner zijn in gezinnen met een hogere SES. Dit kan worden geïnterpreteerd als hoge-SES-ouders die compenseren voor genetische risico's op lagere onderwijsprestaties van hun kinderen. Niet-gedeelde omgevingsinvloeden op de prestaties van kinderen zijn ook afhankelijk van de SES van het gezin. Een deel van de SES-invloed komt dus voort uit verschillende reacties van ouders op kinderen uit hetzelfde gezin. Ten slotte is er veel variatie in de onderwijsprestaties van kinderen uit lage-SES-gezinnen, meer dan voor kinderen uit hoge-SES-gezinnen. Dat er veel verschillen zijn binnen groepen wordt vaak over het hoofd gezien als de focus ligt op verschillen in prestaties tussen SES-groepen.

4. Belangrijk om scholen, gezinnen en genen samen te bestuderen

Het gelijktijdig bestuderen van scholen, gezinnen en genen is van belang om hun invloeden niet verkeerd te interpreteren en hun rol in onderwijsongelijkheid beter te begrijpen. Ten eerste zijn deze factoren gecorreleerd, waardoor de invloed van een factor die niet bestudeerd wordt onterecht toegewezen kan worden aan een factor die wel wordt bestudeerd. Zo zitten kinderen uit hetzelfde sociaaleconomische milieu vaak bij elkaar op school. Als alleen de schoolomgeving of alleen de thuisomgeving wordt bestudeerd, worden deze invloeden (deels) door elkaar gehaald. Dit is bijvoorbeeld te zien in hoofdstuk 3. Daar lijkt de schoolkwaliteit en -compositie eerst belangrijk, maar zodra de SES van de ouders wordt meegenomen, wordt de invloed van de schoolomgeving grotendeels (en voor schoolkwaliteit zelfs volledig) verklaard door de SES van de ouders. Ten tweede zijn de effecten van scholen, gezinnen en genen onderling afhankelijk van elkaar: de invloed van de ene factor hangt af van een andere factor. Dit blijkt uit de verschillende gen-omgevingsinteracties die ik in de empirische hoofdstukken van dit proefschrift heb gevonden.

De klasomgeving verklaart bijvoorbeeld maar een heel klein deel van de prestatieverschillen in het algemeen. Maar als ik kijk hoe de invloed van de klas verschilt naar gelang de SES van het gezin, dan zie ik dat de klasomgeving vooral belangrijk is voor kinderen met een lage-SES-achtergrond. Door rekening te houden met de interacties tussen scholen, gezinnen en genen, krijgen we dus een beter inzicht in hun rol bij het verklaren van onderwijsprestaties en -ongelijkheid.

Beperkingen en toekomstig onderzoek

Hoewel dit proefschrift waardevolle kennis biedt over de wisselwerking tussen scholen, gezinnen en genen in onderwijsongelijkheid, zijn er nog uitdagingen en vragen voor toekomstig onderzoek. Allereerst wijzen de bevindingen van het proefschrift op een beperkte rol van schoolkwaliteit, wat mogelijk deels te maken heeft met meetbeperkingen. Schoolkwaliteit wordt gemeten op basis van indicatoren die worden gebruikt door de Onderwijsinspectie. De inspectie richt zich op het naleven van wettelijke normen (Inspectie van het Onderwijs, 2018), waardoor hun oordelen waarschijnlijk niet alle kwaliteitsverschillen tussen scholen reflecteren. Bovendien bevatten deze data geen gedetailleerde informatie van de meest directe invloeden van scholen op het leren van kinderen, namelijk de invloed van leraren en de leeromgeving die zij bieden (Raudenbush, 2008). Om de rol van scholen diepgaander te onderzoeken, zouden toekomstige studies verschillende metingen en methoden kunnen gebruiken zoals observaties van de vaardigheden en inzet van leraren bij het uitvoeren van specifieke onderwijstaken (Burgess et al., 2022).

Ten tweede is er een alternatieve verklaring voor de afnemende spreiding bij toenemende ouderlijke SES, namelijk een zogenoemde plaffondeffect. Sommige metingen van onderwijsprestaties, zoals de *Cito*-toets, hebben een bepaald maximum en zijn niet moeilijk genoeg voor sommige kinderen, waardoor ze niet goed kunnen laten zien hoeveel ze weten of kunnen. Dit kan leiden tot een schijninteractie tussen SES en (genetische en omgevings-) variantie in onderwijsprestaties (zie Rohrer & Adams, 2021). De variatie in prestaties bij kinderen uit hoge-SES-gezinnen is dan niet lager vanwege substantiële invloeden (zoals de compensatie van genetische risico's), maar omdat hoge-SES kinderen vaker de hoogste score halen en er door de testeigenschappen onvoldoende onderscheid gemaakt worden tussen de prestatieniveaus van de hoogstscorende leerlingen. Hoewel dit een mogelijk scenario is, beargumenteer ik in onder andere hoofdstuk 1 dat de afnemende variantie bij toenemende SES ten minste gedeeltelijk substantiële compensatiemechanismen weerspiegelt. Om deze mechanismen beter te begrijpen, is het nodig om meer onderzoek te doen door mogelijke tussenliggende factoren op te nemen in het tweelingmodel, zoals ADHD of dyslexie (zie Ruks, 2022; Tucker-Drob & Harden, 2012b, 2012c).

Ten derde is de wisselwerking tussen scholen, gezinnen en genen complex en reflecteren de toepassingen van het tweelingmodel in dit proefschrift deze complexiteit slechts gedeeltelijk. De correlaties tussen genen en de gezins- en schoolsituatie zijn niet allemaal gemodelleerd. Een implicatie hiervan is dat dit proefschrift geen causaal bewijs levert voor de omgevingsinvloed van gezinnen en scholen. Een andere implicatie is dat de rol van genetische verschillen niet moet worden overgeïnterpreteerd, omdat het waarschijnlijk is dat veel genetische invloeden

gemedieerd worden door de omgeving. Uitbreidingen van het tweelingmodel kunnen hier meer inzicht in verschaffen.

Tot slot richtte ik me op de rol van genetische en gedeelde omgevingsvariantie en formuleerde ik geen verwachtingen over de rol van de niet-gedeelde omgeving. Echter blijkt de niet-gedeelde omgeving een belangrijke bron van prestatieverschillen, vaak zelfs belangrijker dan gedeelde omgevingsfactoren. Bovendien blijken niet-gedeelde omgevingsinvloeden afhankelijk te zijn van de gezin- en schoolomgeving. Niet-gedeelde omgevingsinvloeden kunnen invloeden van gezinnen en scholen omvatten. Om hier meer inzicht in te krijgen, zouden specifieke metingen op individueel niveau kunnen worden meegenomen, bijvoorbeeld aspecten van de ouder-kind relatie of individuele ervaringen en percepties van de gezins- en schoolsituatie.

Beleidsreflecties

Uit dit proefschrift komen verschillende beleidsreflecties voort. Ten eerste zou bij het ontwikkelen en evalueren van beleid duidelijker moeten worden gemaakt hoe onderwijsongelijkheid wordt gedefinieerd en welke bronnen van ongelijkheid moeten worden verminderd. Onderwijsongelijkheid kan bijvoorbeeld verwijzen naar de totale spreiding in onderwijsprestaties, verschillen in prestaties naar gezinsachtergrond en genetische verschillen in prestaties. Niet al deze vormen van ongelijkheid worden altijd als problematisch en oneerlijk beschouwd. Voor effectief onderwijsbeleid is het van belang om duidelijk te hebben over welk specifieke type ongelijkheid het gaat en om de impact van het beleid op andere bronnen van ongelijkheid in ogenschouw te nemen. Anders bestaat het risico dat beleidsmaatregelen onbedoeld genetische en/of gezinsverschillen vergroten of verkleinen, of de dieperliggende oorzaken missen. 'One size fits all'-beleid, zoals het verlengen van onderwijstijd of universele kinderopvang voor alle leerlingen, kan resulteren in een stijging van de gemiddelde prestaties, maar lijkt weinig effectief te zijn voor het verminderen van spreiding en kansenongelijkheid (Asbury & Wai, 2020). Voor het effectief aanpakken van ongelijkheid hebben specifieke maatregelen gericht op bijvoorbeeld leerlingen met lage prestaties of een lage-SES-achtergrond de voorkeur. Uit dit proefschrift blijkt namelijk dat er meer variantie is, vooral meer genetische variantie, in lage-SES-gezinnen en -scholen. Dit suggereert dat het de moeite waard is om te focussen op het verminderen van ongelijkheid binnen groepen, met name binnen de lage-SES-groep.

Een tweede punt betreft het vergroten van het bewustzijn omtrent genetische verschillen tussen leerlingen en de wisselwerking tussen genen en de omgeving. Onderwijsbeleid houdt vaak geen rekening met genetica, ondanks sterk bewijs, inclusief dit proefschrift, dat genen en omgeving gezamenlijk invloed uitoefenen op onderwijsresultaten (Asbury & Wai, 2020). Hoewel de precieze mechanismen nog nader onderzoek vereisen, kan het erkennen van de rol van genen in onderwijsprestaties leiden tot andere perspectieven op onderwijsongelijkheid en het daaraan gekoppelde beleid. Een te sterke nadruk op omgevingsfactoren leidt tot de situatie waarin ouders en leraren verantwoordelijk worden gehouden voor de lage onderwijsprestaties van sommige kinderen (Asbury & Plomin, 2014). Het besef dat genen ook een rol spelen, kan de verwachtingen en verantwoordelijkheden die aan leraren en scholen worden gesteld om onderwijsongelijkheid

te verminderen wat matigen. Bovendien kan het de focus verleggen naar interventies die de belangrijkste bronnen van onderwijsongelijkheid aanpakken en daarmee waarschijnlijk effectiever zijn. Hierbij valt te denken aan interventies die helpen voorkomen dat genetische risico's tot uiting komen in de leerprestaties van met name de lage-SES-groep (aangezien de hoge-SES-groep zelf al beter in staat is dit soort risico's te compenseren).

Tot slot zouden beleidsmaatregelen verder moeten kijken dan verschillen in schoolkwaliteit en zich meer moeten richten op SES-verschillen tussen scholen en op verschillen binnen scholen. Dit zijn namelijk de schoolaspecten die het belangrijkst blijken te zijn. Hoewel de SES-compositie van de school en de klasomgeving bescheiden effecten hebben, blijft de schoolomgeving een aantrekkelijke context voor interventies om prestaties te verhogen en ongelijkheid te verminderen (Downey & Condron, 2016). Via scholen kan een diverse groep kinderen worden bereikt, terwijl het moeilijker is om in te grijpen in de gezinsomgeving. Wat de beoordeling van schoolkwaliteit betreft laat dit proefschrift zien dat het belangrijk is om de onderwijskwaliteit binnen scholen (de kwaliteit in elke klas) te waarborgen en bevorderen. Meer diepgaande evaluaties van de onderwijskwaliteit, zoals recentelijk gedaan in de Monitor Leskwaliteit van de Onderwijsinspectie, zijn een belangrijke stap in deze richting (Inspectie van het Onderwijs, 2023b). De variatie in de kwaliteit van het onderwijs dat gegeven wordt binnen dezelfde school is waarschijnlijk de belangrijkste factor die bijdraagt aan verschillen tussen klassen (cf. Byrne et al., 2010). Het verbeteren van leskwaliteit komt daarom ten goede aan onderwijsresultaten, waarschijnlijk vooral voor kwetsbare leerlingen. In hoeverre dit wordt gerealiseerd, hangt af van gelijke toegang tot goede leraren. Gezien het lerarentekort, dat zich ook nog eens concentreert op scholen die vanwege hun leerlingpopulatie het meest gebaat zijn bij goede leerkrachten, vraagt dit om voortdurende aandacht in de komende jaren.

REFERENCES

Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher, 36*(7), 369–387. https://doi.org/10.3102/0013189x07308739

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72*(2), 167–180. https://doi.org/10.1177/000312240707200202

Anderson, C. S. (1982). The search for school climate: A review of the research. *Review of Educational Research, 52*(3), 368–420. https://doi.org/10.2307/1170423

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–277). Erlbaum.

Armor, D. J., Marks, G. N., & Malatinszky, A. (2018). The impact of school SES on student achievement: Evidence from U.S. statewide achievement data. *Educational Evaluation and Policy Analysis, 40*(4), 613–630. https://doi.org/10.3102/0162373718787917

Asbury, K., Almeida, D., Hibel, J., Harlaar, N., & Plomin, R. (2008). Clones in the classroom: A daily diary study of the nonshared environmental relationship between monozygotic twin differences in school experience and achievement. *Twin Research and Human Genetics, 11*(6), 586–595. https://doi.org/10.1375/twin.11.6.586

Asbury, K., & Plomin, R. (2014*). G is for genes: The impact of genetics on eduation and achievemen*t. John Wiley & Sons, Inc.

Asbury, K., Wachs, T. D., & Plomin, R. (2005). Environmental moderators of genetic influence on verbal and nonverbal abilities in early childhood. *Intelligence, 33*(6), 643–661. https://doi.org/10.1016/j.intell.2005.03.008

Asbury, K., & Wai, J. (2020). Viewing education policy through a genetic lens. *Journal of School Choice, 14*(2), 301–315. https://doi.org/10.1080/15582159.2019.1705008

Atlay, C., Tieben, N., Hillmert, S., & Fauth, B. (2019). Instructional quality and achievement inequality: How effective is teaching in closing the social achievement gap? *Learning and Instruction, 63*(101211), 1–10. https://doi.org/10.1016/j.learninstruc.2019.05.008

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2019). Family disadvantage and the gender gap in behavioral and educational outcomes. *American Economic Journal: Applied Economics, 11*(3), 338–381. https://doi.org/10.1257/app.20170571

Baier, T. (2019). *What have genes got to do with it? How social and genetic influences contribute to differences in educational success within the family*. [Doctoral dissertation, Bielefeld University]. Deutschen Nationalbibliothek.

Baier, T., & Lang, V. (2019). The social stratification of environmental and genetic influences on education: New evidence using a register-based twin sample. *Sociological Science, 6,* 143–171. https://doi.org/10.15195/V6.A6

Baier, T., Lang, V., Grätz, M., Barclay, K. J., Conley, D. C., Dawes, C. T., Laidley, T., & Lyngstad, T. H. (2022). Genetic Influences on Educational Achievement in Cross-National Perspective. *European Sociological Review, 38*(6), 959–974. https://doi.org/10.1093/esr/jcac014

Baye, A., & Monseur, C. (2016). Gender differences in variability and extreme scores in an international context. *Large-Scale Assessments in Education, 4*(1), 1-16. https://doi.org/10.1186/s40536-015-0015-x

Behrman, J. R., & Taubman, P. (1989). Is schooling "mostly in the genes"? Nature-nurture decomposition using data on relatives. *Journal of Political Economy, 97*(6), 1425–1446. https://doi.org/10.1086/261661

Berkowitz, R., Glickman, H., Benbenishty, R., Ben-Artzi, E., & Astor, R. A. (2015). Compensating, mediating, and moderating effects of school climate on academic achievement gaps in Israel. *Teachers College Record, 117*(070308), 1–34. https://doi.org/10.1177/016146811511700703

Bernardi, F. (2014). Compensatory advantage as a mechanism of educational inequality: A regression discontinuity based on month of birth. *Sociology of Education, 87*(2), 74–88. https://doi.org/10.1177/0038040714524258

Bernardi, F., & Grätz, M. (2015). Making up for an unlucky month of birth in school: Causal evidence on the compensatory advantage of family background in England. *Sociological Science, 2*, 235–251. https://doi.org/10.15195/v2.a12

Bertrand, M., & Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics, 5*(1), 32–64. https://doi.org/10.1257/app.5.1.32

Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behaviour, 7*, 365-385. https://doi.org/10.1038/s41562-022-01506-4

Blatchford, P., & Russell, A. (2020). *Rethinking class size: The complex story of impact on teaching and learning*. UCL Press.

Blau, P. M., & Duncan, O. D. (1967). *The American occupational structure*. Wiley.

Bolhaar, J., & Scheer, B. (2019). *Verschillen in leerresultaten basisscholen*. Centraal Planbureau. https://www.cpb.nl/sites/default/files/omnidownload/CPB-Notitie-19feb2019-Verschillen-in-leerresultaten-tussen-basisscholen.pdf

Borghans, L., Golsteyn, B. H. H., & Zölitz, U. (2015a). Parental preferences for primary school characteristics. *B.E. Journal of Economic Analysis and Policy, 15*(1), 85–117. https://doi.org/10.1515/bejeap-2014-0032

Borghans, L., Golsteyn, B. H. H., & Zölitz, U. (2015b). School quality and the development of cognitive skills between age four and six. *PLoS ONE, 10*(7), e0129700. https://doi.org/10.1371/journal.pone.0129700

Borman, G. D., & Kimball, S. M. (2005). Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *Elementary School Journal, 106*(1), 3–20. https://doi.org/10.1086/496904

Bosker, R. J., & Scheerens, J. (1994). Alternative models of school effectiveness put to the test. *International Journal of Educational Research, 21*(2), 159–180. https://doi.org/10.1016/0883-0355(94)90030-2

Bouchard, T., Lykken, D., McGue, M., Segal, N., & Tellegen, A. (1990). Sources of human psychological differences: the Minnesota Study of Twins Reared Apart. *Science, 250*(4978), 223–228. https://doi.org/10.1126/science.2218526

Boudon, R. (1974). *Education, Opportunity, and Social Inequality. Changing Prospects in Western Society*. Wiley.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research in the sociology of education*. Greenwood Press.

Bourdieu, P., & Passeron, J.-C. (1977). *Reproduction in education, society and culture*. Sage.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*(1), 371–399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials towards a formal rational action theory. *Rationality and Society, 9*(3), 275–305. https://doi.org/10.1177/104346397009003002

Breen, R., & Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology, 31*(1), 223–243. https://doi.org/10.1146/annurev.soc.31.041304.122232

Briley, D. A., & Tucker-Drob, E. M. (2013). explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science, 24*(9), 1704–1713. https://doi.org/10.1177/0956797613478618

Bronfenbrenner, U. (1992). Ecological systems theory. In R. Vasta (Ed.), *Six theories of child development: Revised formulations and current issues* (pp. 187–249). Jessica Kingsley.

Bronfenbrenner, U., & Ceci, S. J. (1994). nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review, 101*(4), 568–586.

Brown, T. A. (2015). CFA models with single indicators. In *Confirmatory factor analysis for applied research* (2nd ed., pp. 122–127). The Guilford Press.

Buchmann, C., & DiPrete, T. A. (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review, 71*(4), 515–541. https://doi.org/10.1177/000312240607100401

Buchmann, C., DiPrete, T. A., & McDaniel, A. (2008). Gender inequalities in education. *Annual Review of Sociology, 34,* 319–337. https://doi.org/10.1146/annurev.soc.34.040507.134719

Bueno, D. (2019). Genetics and learning: How the genes influence educational attainment. *Frontiers in Psychology, 10*(1622), 1–10. https://doi.org/10.3389/fpsyg.2019.01622

Bulder, E., Van den Langenberg, K., Van der Ploeg, S., Karssen, M., Vaessen, A., Leest, B., & Beurskens, K. (2020). *Alle scholen toegankelijk: een open deur? Toelatingsbeleid en toegankelijkheid in het funderend onderwijs*. Kohnstamm Instituut. https://kohnstamminstituut.nl/wp-content/uploads/2021/01/rapport-toelatingsbeleid-en-toegankelijkheid-4.pdf

Burgess, S., rawal, shenila, & Taylor, E. S. (2022). *Teachers' use of class time and student achievement* (NBER Working paper no. w30686). National Bureau of Economic Research. https://doi.org/10.2139/ssrn.4288519

Byrne, B., Coventry, W. L., Olson, R. K., Wadsworth, S. J., Samuelsson, S., Petrill, S. A., Willcutt, E. G., & Corley, R. (2010). "Teacher effects" in early literacy development: Evidence from a study of twins. *Journal of Educational Psychology, 102*(1), 32–42. https://doi.org/10.1037/a0017288

Byrne, B., Samuelsson, S., Wadsworth, S., Hulslander, J., Corley, R., Defries, J. C., Quain, P., Willcutt, E. G., & Olson, R. K. (2007). Longitudinal twin study of early literacy development: Preschool through Grade 1. *Reading and Writing, 20*(1–2), 77–102. https://doi.org/10.1007/s11145-006-9019-9

Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research, 90*(5), 269–277. https://doi.org/10.1080/00220671.1997.10544583

Carey, G. (2003). Quantitative genetics I: genetic variation. In *Human Genetics for the Social Sciences* (pp. 285–309). SAGE Publications Inc. https://doi.org/https://doi.org/10.4135/9781452229591

Caro, D. H., Lenkeit, J., Lehmann, R., & Schwippert, K. (2009). The role of academic achievement growth in school track recommendations. *Studies in Educational Evaluation, 35*(4), 183–192. https://doi.org/10.1016/J.STUEDUC.2009.12.002

Centola, D., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology, 113*(3), 702–734. https://doi.org/10.1086/521848

Centraal Bureau voor Statistiek. (1995). *In de Gemeentelijke Basisadministratie Persoonsgegevens (GBA) ingeschreven personen en hun juridische ouders* [data set]. ODISSEI Portal, V1. https://doi.org/https://doi.org/10.57934/0b01e410801f9401

Centraal Bureau voor Statistiek. (2011). *Inkomen van personen (revisie 2017)* [data set]. ODISSEI Portal, V1. https://doi.org/https://doi.org/10.57934/0b01e41080372fbd

Centraal Bureau voor Statistiek. (2014). *Inschrijvingen van leerlingen in het basisonderwijs* [data set]. ODISSEI Portal, V1. https://doi.org/https://doi.org/10.57934/0b01e410803b8794

Centraal Bureau voor Statistiek. (2018). *Kenmerken van deelnemers aan de Eindtoets Basisonderwijs van Cito* [data set]. ODISSEI Portal, V1. https://doi.org/https://doi.org/10.57934/0b01e4108071ba40

Centraal Bureau voor Statistiek. (2019). *Hoogst behaald en hoogst gevolgd opleidingsniveau en opleidingsrichting van de bevolking in Nederland* [data set]. ODISSEI Portal, V1. https://doi.org/https://doi.org/10.57934/0b01e410806c9615

Centraal Bureau voor Statistiek. (2021). *Persoonskenmerken van alle in de Gemeentelijke Basis Administratie (GBA) ingeschreven personen, gecoördineerd* [data set]. ODISSEI Portal, V1. https://doi.org/https://doi.org/10.57934/0b01e4108071ba40

Cheesman, R., Borgen, N. T., Lyngstad, T. H., Eilertsen, E. M., Ayorech, Z., Torvik, F. A., Andreassen, O. A., Zachrisson, H. D., & Ystrom, E. (2022). How genes, schools, and residential areas shape achievement: A population-wide gene-environment interaction study. *Npj Science of Learning, 7*(29), 1–9. https://doi.org/10.1038/s41539-022-00145-8

Chiu, M. M., & Khoo, L. (2005). Effects of resources, inequality, and privilege bias on achievement: country, school, and student level analyses. *American Educational Research Journal, 42*(4), 575–603. https://doi.org/10.3102/00028312042004575

Chmielewski, A. K. (2019). The global increase in the socioeconomic achievement gap, 1964 to 2015. *American Sociological Review, 84*(3), 517–544. https://doi.org/10.1177/0003122419847165

Christopher, M. E., Hulslander, J., Byrne, B., Samuelsson, S., Keenan, J. M., Pennington, B., Defries, J. C., Wadsworth, S. J., Willcutt, E. G., & Olson, R. K. (2013). Modeling the etiology of individual differences in early reading development: Evidence for strong genetic influences. *Scientific Studies of Reading, 17*(5), 350–368. https://doi.org/10.1080/10888438.2012.729119

Coenen, J., Cornelisz, I., Groot, W., Maassen van den Brink, H., & Van Klaveren, C. (2018). Teacher characteristics and their effects on student test scores: A systematic review. *Journal of Economic Surveys, 32*(3), 848–877. https://doi.org/10.1111/joes.12210

Cohen, J., McCabe, E. M., Michelli, N. M., & Pickeral, T. (2009). School climate: Research, policy, teacher education and practice. *Teachers College Record, 111*(1), 180–213. https://doi.org/10.1177/016146810911100108

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfield, F. D., & York, R. L. (1966). *Equality of educational opportunity.* U.S. Government Printing Office.

Collins, F. S., & Mansoura, M. K. (2001). The human genome project: Revealing the shared inheritance of all humankind. *Cancer, 91*(1), 221–225. https://doi.org/10.1002/1097-0142(20010101)91:1+<221::aid-cncr8>3.0.co;2-9

Condron, D. J. (2007). Stratification and educational sorting: Explaining ascriptive inequalities in early childhood reading group placement. *Social Problems, 54*(1), 1533–8533. https://doi.org/10.1525/sp.2007.54.1.139

Condron, D. J. (2008). An early start: Skill grouping and unequal reading gains in the elementary years. *The Sociological Quarterly, 49*(2), 363–394. https://doi.org/10.1111/j.1533-8525.2008.00119.x

Condron, D. J., Downey, D. B., & Kuhfeld, M. (2021). Schools as refractors: Comparing summertime and school-year skill inequality trajectories. *Sociology of Education, 94*(4), 316–340. https://doi.org/10.1177/00380407211041542

Conley, D. (2008). Bringing sibling differences in: Enlarging our understanding of the transmission of advantage in families. In A. Lareau & D. Conley (Eds.), *Social Class: How Does It Work?* (pp. 179–200). Russell Sage Foundation.

Conley, D., & Fletcher, J. (2017). The genome factor: What the social genomics revolution reveals about ourselves, our history, and the future. Princeton University Press.

Conley, D., Rauscher, E., Dawes, C., Magnusson, P. K. E., & Siegal, M. L. (2013). Heritability and the equal environments assumption: Evidence from multiple samples of misclassified twins. *Behavior Genetics, 43*(5), 415–426. https://doi.org/10.1007/s10519-013-9602-1

Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., Underwood, P., & Morrison, F. J. (2009). Individualizing student instruction precisely: Effects of child × instruction interactions on first graders' literacy development. *Child Development, 80*(1), 77–100. https://doi.org/10.1111/J.1467-8624.2008.01247.X

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66*(3), 227–268. https://doi.org/10.3102/00346543066003227

Crosnoe, R. (2009). Low-income students and the socioeconomic composition of public high schools. *American Sociological Review, 74*(5), 709–730. https://doi.org/10.1177/000312240907400502

De Graaf, N. D., De Graaf, P. M., & Kraaykamp, G. (2000). Parental cultural capital and educational attainment in the Netherlands: A refinement of the cultural capital perspective. *Sociology of Education, 73*(2), 92–111. https://doi.org/10.2307/2673239

De Wolf, I. (2023). Kwaliteit basisonderwijs staat onder druk. ESB, 108(4819), 104–107. https://esb.nu/wp-content/uploads/2023/03/104-107_deWolf3.pdf

De Zeeuw, E. L., & Boomsma, D. I. (2017). Country-by-genotype-by-environment interaction in childhood academic achievement. *Proceedings of the National Academy of Sciences of the United States of America, 114*(51), 13318–13320. https://doi.org/10.1073/pnas.1718938115

De Zeeuw, E. L., De Geus, E. J. C., & Boomsma, D. I. (2015). Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement. *Trends in Neuroscience and Education, 4*(3), 69–76. https://doi.org/10.1016/j.tine.2015.06.001

De Zeeuw, E. L., Kan, K.-J., Van Beijsterveldt, C. E. M., Mbarek, H., Hottenga, J.-J., Davies, G. E., Neale, M. C., Dolan, C. V., & Boomsma, D. I. (2019). The moderating role of SES on genetic differences in educational achievement in the Netherlands. *Npj Science of Learning, 4*(13), 1–8. https://doi.org/10.1038/s41539-019-0052-2

De Zeeuw, E. L., Van Beijsterveldt, C. E. M., De Geus, E. J. C., & Boomsma, D. I. (2012). Twin specific risk factors in primary school achievements. *Twin Research and Human Genetics, 15*(1), 107–115. https://doi.org/10.1375/twin.15.1.107

De Zeeuw, E. L., Van Beijsterveldt, C. E. M., Glasner, T. J., De Geus, E. J. C., & Boomsma, D. I. (2016). Arithmetic, reading and writing performance has a strong genetic component: A study in primary school children. *Learning and Individual Differences, 47*, 156–166. https://doi.org/10.1016/j.lindif.2016.01.009

Demange, P. A., Malanchini, M., Mallard, T. T., Biroli, P., Cox, S. R., Grotzinger, A. D., Tucker-Drob, E. M., Abdellaoui, A., Arseneault, L., van Bergen, E., Boomsma, D. I., Caspi, A., Corcoran, D. L., Domingue, B. W., Harris, K. M., Ip, H. F., Mitchell, C., Moffitt, T. E., Poulton, R., … Nivard, M. G. (2021). Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nature Genetics, 53*(1), 35–44. https://doi.org/10.1038/s41588-020-00754-2

Dias Pereira, R. (2021). Genetic advantage and equality of opportunity in education: Two definitions and an empirical illustration. *BioRxiv*. https://doi.org/10.2139/ssrn.4116304

Diewald, M., Baier, T., Schulz, W., & Schunck, R. (2015). status attainment and social mobility. How can genetics contribute to an understanding of their causes? *KZfSS Kölner Zeitschrift Für Soziologie Und Sozialpsychologie, 67*, 371–395. https://doi.org/10.1007/s11577-015-0317-6

DiPrete, T. A., & Eirich, G. M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual Review of Sociology, 32*(1), 271–297. https://doi.org/10.1146/annurev.soc.32.061604.123127

Domingue, B. W., Trejo, S., Armstrong-Carter, E., & Tucker-Drob, E. M. (2020). Interactions between polygenic scores and environments: Methodological and conceptual challenges. *Sociological Science, 7*, 465–486. https://doi.org/10.15195/V7.A19

Downey, D. B., & Condron, D. J. (2016). Fifty Years since the Coleman report: Rethinking the relationship between schools and inequality. *Sociology of Education, 89*(3), 207–220. https://doi.org/10.1177/0038040716651676

Downey, D. B., Kuhfeld, M., & van Hek, M. (2022). Schools as a relatively standardizing institution: The case of gender gaps in cognitive skills. *Sociology of Education, 95*(2), 89–109. https://doi.org/10.1177/00380407211070319

Downey, D. B., Quinn, D. M., & Alcaraz, M. (2019). The distribution of school quality: Do schools serving mostly white and high-SES children produce the most learning? *Sociology of Education, 92*(4), 386–403. https://doi.org/10.1177/0038040719870683

Downey, D. B., & Vogt Yuan, A. S. (2005). Sex differences in school performance during high school: Puzzling patterns and possible explanations. *The Sociological Quarterly, 46*(2), 299–321. https://doi.org/10.1111/j.1533-8525.2005.00014.x

Downey, D. B., von Hippel, P. T., & Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review, 69*(5), 613–635. http://journals.sagepub.com/doi/pdf/10.1177/000312240406900501

Dronkers, J., & Robert, P. (2008). Differences in scholastic achievement of public, private government-dependent, and private independent schools: A cross-national analysis. *Educational Policy, 22*(4), 541–577. https://doi.org/10.1177/0895904807307065

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genetics*, 9(3), 1003348. https://doi.org/10.1371/journal.pgen.1003348

Dumas, J. E., Nissley, J., Nordstrom, A., Smith, E. P., Prinz, R. J., & Levine, D. W. (2005). Home chaos: Sociodemographic, parenting, interactional, and child correlates. *Journal of Clinical Child and Adolescent Psychology, 34*(1), 93–104. https://doi.org/10.1207/s15374424jccp3401_9

Dumont, H., & Ready, D. D. (2020). Do schools reduce or exacerbate inequality? How the associations between student achievement and achievement growth influence our understanding of the role of schooling. *American Educational Research Journal, 57*(2), 728–774. https://doi.org/10.3102/0002831219868182

Eaves, L. J., & Jinks, J. L. (1972). Insignificance of evidence for differences in heritability of IQ between races and social classes. *Nature, 240,* 84–88. https://doi.org/10.1038/240084a0

Eckland, B. K. (1967). Genetics and sociology: A recosideration. *American Sociological Review, 32*(2), 173–194. https://about.jstor.org/terms

Eifler, E. F., Starr, A., & Riemann, R. (2019). The genetic and environmental effects on school grades in late childhood and adolescence. *PLoS ONE, 14*(12), 1–17. https://doi.org/10.1371/journal.pone.0225946

Elder, G. H. J., & Shanahan, M. J. (2006). The lifecourse and child development. In R. M. Lerner (Ed.), *Handbook of Child Psychology, Volume 1: Theoretical Models of Human Development* (6th ed., pp. 665–715). John Wiley & Sons, Inc.

Engelhardt, L. E., Church, J. A., Harden, K. P., & Tucker-Drob, E. M. (2019). Accounting for the shared environment in cognitive abilities and academic achievement with measured socioecological contexts. *Developmental Science, 22*(1), e12699. https://doi.org/10.1111/desc.12699

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences of the United States of America*, 118(17), 1–7. https://doi.org/10.1073/pnas.2022376118/-/DCSupplemental.y

Erola, J., & Kilpi-Jakonen, E. (2017a). Compensation and other forms of accumulation in intergenerational social inequality. In J. Erola & E. Kilpi-Jakonen (Eds.), *Social inequality across the generations. The role of compensation and multiplication in resource accumulation* (pp. 3–24). Edward Elgar Publishing.

Erola, J., & Kilpi-Jakonen, E. (Eds.). (2017b). *Social inequality across the generations. The role of compensation and multiplication in resource accumulation*. Edward Elgar Publishing.

Erola, J., Lehti, H., Baier, T., & Karhula, A. (2021). Socioeconomic background and gene–environment interplay in social stratification across the early life course. *European Sociological Review, 38*(1), 1–17. https://doi.org/10.1093/esr/jcab026

Evans, D. M., Gillespie, N. A., & Martin, N. G. (2002). Biometrical genetics. *Biological Psychology, 61*, 33–51. https://doi.org/10.1038/hdy.1972.69

Evans, D. M., & Martin, N. G. (2000). The validity of twin studies. *GeneScreen, 1*(2), 77–79. https://doi.org/10.1046/j.1466-9218.2000.00027.x

Fagerness, J. A., & Nyholt, D. R. (2008). Basics of DNA and genotyping. In B. M. Neale, M. A. R. Ferreira, S. E. Medland, & D. Posthuma (Eds.), *Statistical genetics: gene mapping through linkage and association* (pp. 5–15). Taylor & Francis.

Fan, X. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 70*(1), 27–61. https://doi.org/10.1023/A:1009048817385

Felson, J. (2014). What can we learn from twin studies? A comprehensive evaluation of the equal environments assumption. *Social Science Research, 43*, 184–199. https://doi.org/10.1016/j.ssresearch.2013.10.004

Ferreira, F. H. G., & Gignoux, J. (2014). The Measurement of educational inequality: Achievement and opportunity. *World Bank Economic Review, 28*(2), 210–246. https://doi.org/10.1093/wber/lht004

Figlio, D. N., Freese, J., Karbownik, K., & Roth, J. (2017). Socioeconomic status and genetic influences on cognitive development. *Proceedings of the National Academy of Sciences of the United States of America, 114*(51), 13441–13446. https://doi.org/10.1073/pnas.1708491114

Fletcher, J. M., & Conley, D. (2013). The challenge of causal inference in gene-environment interaction research: Leveraging research designs from the social sciences. American *Journal of Public Health, 103*(1), 42–45. https://doi.org/10.2105/AJPH.2013.301290

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal , individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3–17. http://dx.doi.org/10.1037/0022-0663.88.1.3

Franić, S., Dolan, C. V., Borsboom, D., & Boomsma, D. I. (2012). Structural Equation Modeling in Genetics. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 617–635). Guilford Press.

Freese, J. (2008). Genetics and the social science explanation of individual outcomes. *American Journal of Sociology, 114*(1), 1–35. https://doi.org/10.1086/592208

Freese, J., & Jao, Y. H. (2017). Shared environment estimates for educational attainment: A puzzle and possible solutions. *Journal of Personality, 85*(1), 79–89. https://doi.org/10.1111/jopy.12226

Freese, J., & Powell, B. (1999). Sociobiology, status, and parental investment in sons and daughters: Testing the Trivers-Willard hypothesis. *American Journal of Sociology, 104*(6), 1704–1743. https://doi.org/10.1086/210221

Friend, A., Defries, J. C., & Olson, R. K. (2008). Parental education moderates genetic influences on reading disability. *Psychological Science, 19*(11), 1124–1130. https://doi.org/10.1111/j.1467-9280.2008.02213.x

Gensowski, M., Landersø, R., Bleses, D., Dale, P., Højen, A., & Justice, L. (2020). *Public and parental investments and children's skill formation* (Working paper no. 155). The ROCKWOOL Foundation Research Unit. https://en.rockwoolfonden.dk/publications/public-and-parental-investments-and-childrens-skill-formation/

Ghirardi, G., Baier, T., Kleinert, C., & Triventi, M. (2023). Is early formal childcare an equalizer? How attending childcare and education centres affects children's cognitive and socio-emotional skills in Germany. *European Sociological Review, 39*(5), 692–707. https://doi.org/10.1093/esr/jcac048

Ghirardi, G., & Bernardi, F. (2023). Re-theorizing sociogenomics research on the gene-by-family socioeconomic status (GxSES) interaction for educational attainment. *SocArXiv.* https://doi.org/https://doi.org/10.31235/osf.io/2xny7

Glasner, T. J., van Beijsterveldt, C. E. M., Willemsen, G., & Boomsma, D. I. (2013). Meerlinggeboorten in Nederland. *Nederlands Tijdschrift Voor Geneeskunde, 157*(A5962), 1–5.

Gottschling, J., Hahn, E., Beam, C. R., Spinath, F. M., Carroll, S., & Turkheimer, E. (2019). Socioeconomic status amplifies genetic effects in middle childhood in a large German twin sample. *Intelligence, 72*, 20–27. https://doi.org/10.1016/j.intell.2018.11.006

Grasby, K. L., & Coventry, W. L. (2016). Longitudinal stability and growth in literacy and numeracy in Australian school students. *Behavior Genetics, 46*, 649–664. https://doi.org/10.1007/s10519-016-9796-0

Grasby, K. L., Little, C. W., Byrne, B., Coventry, W. L., Olson, R. K., Larsen, S., & Samuelsson, S. (2020). Estimating classroom-level influences on literacy and numeracy: A twin study. *Journal of Educational Psychology, 112*(6), 1154–1166. https://doi.org/10.1037/edu0000418.supp

Grätz, M. (2018). Competition in the family: inequality between siblings and the intergenerational transmission of educational advantage. *Sociological Science, 5*, 246–269. https://doi.org/10.15195/v5.a11

Grätz, M., & Torche, F. (2016). Compensation or reinforcement? The stratification of parental responses to children's early ability. *Demography, 53*(6), 1883–1904. https://doi.org/10.1007/s13524-016-0527-1

Gray, H., Lyth, A., McKenna, C., Stothard, S., Tymms, P., & Copping, L. (2019). Sex differences in variability across nations in reading, mathematics and science: a meta-analytic extension of Baye and Monseur (2016). *Large-Scale Assessments in Education, 7*(2), 1-29. https://doi.org/10.1186/s40536-019-0070-9

Greenwald, R., Hedges, L. V, & Laine, R. D. (1996). The effect of school resources on student achievement. *American Educational Research Association, 66*(3), 361–396.

https://doi.org/10.3102/00346543066003361

Gremmen, M. C., van den Berg, Y. H. M., Steglich, C., Veenstra, R., & Dijkstra, J. K. (2018). The importance of near-seated peers for elementary students' academic engagement and achievement. *Journal of Applied Developmental Psychology, 57*, 42–52. https://doi.org/10.1016/j.appdev.2018.04.004

Grubb, W. N. (2009). The money myth: School resources, outcomes, and equity. In *The money myth: School resources, outcomes, and equity*. Russell Sage Foundation. https://doi.org/10.1177/0094306110373238u

Guo, G., & Stearns, E. (2002). The social influences on the realization of genetic potential for intellectual development. *Social Forces, 80*(3), 881–910. https://doi.org/10.1353/sof.2002.0007

Gustafsson, J. E., Nilsen, T., & Hansen, K. Y. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation, 57*, 16–30. https://doi.org/10.1016/j.stueduc.2016.09.004

Gutiérrez, G. (2023). Is it socioeconomic or academic? Disentangling sources of peer effects on student achievement. *British Journal of Sociology of Education, 44*(1), 144–163. https://doi.org/10.1080/01425692.2022.2137465

Hadley, E. B., Newman, K. M., & Kim, E. S. (2023). Identifying levers for improvement: examining proximal processes and contextual influences on preschool language development. *Early Education and Development, 34*(1), 181–207. https://doi.org/10.1080/10409289.2021.1979835

Haelermans, C., Huijgen, T., Jacobs, M., Levels, M., Van Der Velden, R., Van Vugt, L., & Van Wetten, S. (2020). Using data to advance educational research, policy, and practice: Design, content, and research potential of the Netherlands cohort study on education. *European Sociological Review, 36*(4), 643–662. https://doi.org/10.1093/esr/jcaa027

Haelermans, C., Jacobs, M., & Van Vugt, L. (2021). *Factsheets leergroei na anderhalf jaar COVID-19-crisis. Technische Toelichting* (Version 1.0). Nationaal Cohortonderzoek Onderwijs. https://www.nationaalcohortonderzoek.nl/sites/nco/files/media-files/TechnischeToelichting_v1.0.docx.pdf

Haelermans, C., Korthals, R., Jacobs, M., de Leeuw, S., Vermeulen, S., van Vugt, L., Aarts, B., Prokic-Breuer, T., van der Velden, R., van Wetten, S., & de Wolf, I. (2022). Sharp increase in inequality in education in times of the COVID-19-pandemic. *PLoS ONE, 17*(2). https://doi.org/10.1371/journal.pone.0261114

Haelermans, C., van Wetten, S., Abbink, H., Baumann, S., Hendrickx, S., Hendrikse, A., Jacobs, M., Meijer, R., & van Vugt, L. (2022). *Verschillen tussen scholen blijven bestaan* (NCO Factsheet No. 2022-9). Nationaal Cohortonderzoek Onderwijs. https://www.nationaalcohortonderzoek.nl/sites/nco/files/media-files/factsheet2022_09_nco_schoolkenmerken_def.pdf

Hallinan, M. T. (1988). Equality of educational opportunity. *Annual Review of Sociology, 14*, 249–268. https://doi.org/10.1111/j.1741-5446.1976.tb00708.x

Hanselman, P. (2018). Do school learning opportunities compound or compensate for background inequalities? Evidence from the case of assignment to effective teachers. *Sociology of Education, 91*(2), 132–158. https://doi.org/10.1177/0038040718761127

Hanselman, P., & Fiel, J. E. (2017). School opportunity hoarding? Racial segregation and access to high growth schools. *Social Forces, 95*(3), 1077–1104. https://doi.org/10.1093/sf/sow088

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics, 18*(5), 527–544. https://doi.org/10.1002/jae.741

Hanushek, E. A., Piopiunik, M., & Wiederhold, S. (2019). The value of smarter teachers: International evidence on teacher cognitive skills and student performance. *Journal of Human Resources, 54*(4), 857–899. https://doi.org/10.3368/jhr.54.4.0317.8619R1

Hanushek, E. A., & Rivkin, S. G. (2006). Teacher Quality. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. 2, pp. 1051–1078). Elsevier. https://doi.org/10.1016/S1574-0692(06)02018-6

Hanushek, E. A., & Wößmann, L. (2006). Does early tracking affect educational inequality and performance? Differences-in-differences evidence across countries. *Economic Journal, 116*(115), C63–C76. https://doi.org/10.1111/j.1468-0297.2006.01076.x

Harden, K. P. (2021). *The genetic lottery: Why DNA matters for social equality*. Princeton University Press.

Harden, K. P., Domingue, B. W., Belsky, D. W., Boardman, J. D., Crosnoe, R., Malanchini, M., Nivard, M., Tucker-Drob, E. M., & Harris, K. M. (2020). Genetic associations with mathematics tracking and persistence in secondary school. *Npj Science of Learning, 5*(1), 1-8. https://doi.org/10.1038/s41539-020-0060-2

Hart, S. A., Little, C., & van Bergen, E. (2021). Nurture might be nature: cautionary tales and proposed solutions. *Npj Science of Learning, 6*(2), 1–12. https://doi.org/10.1038/s41539-020-00079-z

Hart, S. A., Logan, J. A. R., Soden-Hensler, B., Kershaw, S., Taylor, J., & Schatschneider, C. (2013). Exploring how nature and nurture affect the development of reading: An analysis of the Florida Twin Project on reading. *Developmental Psychology, 49*(10), 1971–1981. https://doi.org/10.1037/a0031348

Hart, S. A., Soden, B., Johnson, W., Schatschneider, C., & Taylor, J. (2013). Expanding the environment: Gene × school-level SES interaction on reading comprehension. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 54*(10), 1047–1055. https://doi.org/10.1111/jcpp.12083

Haughbrook, R., Hart, S. A., Schatschneider, C., & Taylor, J. (2017). Genetic and environmental influences on early literacy skills across school grade contexts. *Developmental Science, 20*(5), 1-12. https://doi.org/10.1111/desc.12434

Haworth, C. M. A., Asbury, K., Dale, P. S., & Plomin, R. (2011). Added value measures in education show genetic as well as environmental influence. *PLoS ONE, 6*(2), 16006. https://doi.org/10.1371/journal.pone.0016006

Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics, 54*(1), 1–48. https://doi.org/10.1006/reec.1999.0225

Herd, P., Freese, J., Sicinski, K., Domingue, B. W., Mullan Harris, K., Wei, C., & Hauser, R. M. (2019). Genes, gender inequality, and educational attainment. *American Sociological Review, 84*(6), 1069–1098. https://doi.org/10.1177/0003122419886550

Hivert, V., Sidorenko, J., Rohart, F., Goddard, M. E., Yang, J., Wray, N. R., Yengo, L., & Visscher, P. M. (2021). Estimation of non-additive genetic variance in human complex traits from a large sample of unrelated individuals. *American Journal of Human Genetics, 108*(5), 786–798. https://doi.org/10.1016/j.ajhg.2021.02.014

Hofflinger, A., & von Hippel, P. T. (2020). Does achievement rise fastest with school choice, school resources, or family resources? Chile from 2002 to 2013. *Sociology of Education, 93*(2), 132–152. https://doi.org/10.1177/0038040719899358

Hop, M., Scheltens, F., Janssen, J., & Engelen, R. (2016). *Wetenschappelijke verantwoording. Rekenen-basisbewerkingen voor groep 3 tot en met 8 van het primair onderwijs.* Cito. https://cito.nl/media/2wvd1y1r/103-cito_lvs-rekenen_basisbewerkingen-gr-3-tm-8_wet-verantwoording.pdf

Horvat, E. M., Weininger, E. B., & Lareau, A. (2003). From social ties to social capital: Class differences in the relations between schools and parent networks. *American Educational Research Journal, 40*(2), 319–351. https://doi.org/10.3102/00028312040002319

Hunter, M. D., Garrison, S. M., Burt, S. A., Rodgers, J. L., & Hunter, M. D. (2021). The analytic identification of variance component models common to behavior Genetics. *Behavior Genetics, 51*(4), 425–437. https://doi.org/10.1007/s10519-021-10055-x

Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 38*(3), 499–534. https://doi.org/10.3102/00028312038003499

Inspectie der Rijksfinanciën. (2020). *Fundament op orde: kwalitatief goed onderwijs met kansen voor iedereen.* Inspectie der Rijksfinanciën. https://www.rijksoverheid.nl/documenten/rapporten/2020/04/20/bmh-1-kwalitatief-goed-onderwijs-met-kansen-voor-iedereen

Inspectie van het Onderwijs. (n.d.). *Toezicht op scholen*. Retrieved July 12, 2023, from https://www.onderwijsinspectie.nl/onderwerpen/vierjaarlijks-onderzoek/toezicht-op-scholen

Inspectie van het Onderwijs. (2016). *De staat van het onderwijs: Onderwijsverslag 2014-2015.* Inspectie van het Onderwijs. https://www.onderwijsinspectie.nl/documenten/publicaties/2016/04/13/staat-van-het-onderwijs-2014-2015

Inspectie van het Onderwijs. (2017). *De staat van het onderwijs: Onderwijsverslag 2015-2016*. Inspectie van het Onderwijs. https://www.onderwijsinspectie.nl/documenten/rapporten/2017/04/12/staat-van-het-onderwijs-2015-2016

Inspectie van het Onderwijs. (2018). *De staat van het onderwijs 2018: Onderwijsverslag over 2016/2017.* Inspectie van het Onderwijs. https://www.onderwijsinspectie.nl/documenten/rapporten/2018/04/11/rapport-de-staat-van-het-onderwijs

Inspectie van het Onderwijs. (2019). *Technisch rapport - De staat van het primair onderwijs 2019*. Inspectie van het Onderwijs. https://www.onderwijsinspectie.nl/documenten/rapporten/2019/04/10/technisch-rapport-primair-onderwijs

Inspectie van het Onderwijs. (2020). *Peil.Taal en rekenen einde basisonderwijs 2018 - 2019*. Inspectie van het Onderwijs. https://www.onderwijsinspectie.nl/documenten/themarapporten/2020/04/22/peil.taal-en-rekenen-2018-2019

Inspectie van het Onderwijs. (2021a). *De staat van het onderwijs 2021.* Inspectie van het Onderwijs. https://www.onderwijsinspectie.nl/documenten/rapporten/2021/04/14/de-staat-van-het-onderwijs-2021

Inspectie van het Onderwijs. (2021b). *Inspection framework 2021: For the supervision of preschool education and primary education*. Inspectie van het Onderwijs. https://english.onderwijsinspectie.nl/inspection/documents/publications/2021/12/13/inspection-framework-primary-education-2021

Inspectie van het Onderwijs. (2023a). *De staat van het onderwijs 2023.* Inspectie van het Onderwijs. https://www.onderwijsinspectie.nl/documenten/rapporten/2023/05/10/rapport-de-staat-van-het-onderwijs-2023

Inspectie van het Onderwijs. (2023b). *Leskwaliteit in beeld: kenmerken van een effectieve les in de praktijk* (Factsheet basisonderwijs – maart 2023). Inspectie van het Onderwijs. https://www.onderwijsinspectie.nl/documenten/publicaties/2023/03/06/factsheet-monitor-leskwaliteit-po

ISLED. (2014). *Overview of the International Standard Level of Education [ISLED].* Retrieved March 20, 2021, from http://www.harryganze boom.nl/isled

Jencks, C., & Mayer, S. E. (1990). The social consequences of growing up in a poor neighborhood. In National Research Council (Ed.), *Inner-City Poverty in the United States* (pp. 111–186). The National Academies Press.

Jennings, J. L., Deming, D., Jencks, C., Lopuch, M., & Schueler, B. E. (2015). Do Differences in school quality matter more than we thought? New evidence on educational opportunity in the twenty-first century. *Sociology of Education, 88*(1), 56–82. https://doi.org/10.1177/0038040714562006

Jensen, V. M., & Rasmussen, A. W. (2011). Danish education registers. *Scandinavian Journal of Public Health, 39*(7), 91–94. https://doi.org/10.1177/1403494810394715

Johnson, W. (2007). Genetic and environmental influences on behavior: Capturing all the interplay. *Psychological Review, 114*(2), 423–440. https://doi.org/10.1037/0033-295X.114.2.423

Jones, L., & De Gioia, K. (2010). The same or separate? An exploration of teachers' perceptions of the classroom assignment of twins in prior to school and kindergarten to Year Two school settings. *Journal of Early Childhood Research, 8*(3), 239–253. https://doi.org/10.1177/1476718X10366772

Jonsson, J. O., & Treuter, G. (2019). Likvärdighet och skolkvalitet: socioekonomiskt ursprung och invandrarbakgrund. In R. Erikson & L. Unemo (Eds.), *Lika för alla? En ESO-antologi om skolans likvärdighe*t (pp. 49–86). Finansdepartementet. https://eso.expertgrupp.se/rapporter/2019_1-lika-for-alla/

Joseph, J. (2014). *The trouble with twin studies: A reassessment of twin research in the social and behavioral sciences.* Routledge. https://doi.org/10.4324/9781315748382

Kahlenberg, R. D. (2001). The significance of the socioeconomic makeup of schools. In *All together now: The case for economic integration of the public schools* (pp. 47–76). Brookings Institution.

Kendler, K. S., & Baker, J. H. (2007). Genetic influences on measures of the environment: A systematic review. *Psychological Medicine, 37*(5), 615–626. https://doi.org/10.1017/S0033291706009524

Kievit, R. A., Logan, J. A., & Hart, S. A. (2021). From the trajectory of heritability to the heritability of trajectories. *Behavioral and Brain Sciences, 45*(E165), 1–3. https://doi.org/10.1017/S0140525X21001643

Knigge, A., Maas, I., Stienstra, K., de Zeeuw, E. L., & Boomsma, D. I. (2022). Delayed tracking and inequality of opportunity: Gene-environment interactions in educational attainment. *Npj Science of Learning, 7*(6), 1-13. https://doi.org/10.1038/s41539-022-00122-1

Knopik, V. S., Neiderhiser, J. M., DeFries, J. C., & Plomin, R. (2016). *Behavioral Genetics* (7th ed.). Worth Publishers.

Kovas, Y., Haworth, C. M. A., Dale, P. S., & Plomin, R. (2007). The genetic and environmental origins of learning abilities and disabilities in the early school years. *Monographs of the Society for Research in Child Development, 72*, 1–158. https://doi.org/10.1111/j.1540-5834.2007.00439.x

Kovas, Y., Voronin, I., Kaydalov, A., Malykh, S. B., Dale, P. S., & Plomin, R. (2013). Literacy and numeracy are more heritable than intelligence in primary school. *Psychological Science, 24*(10), 2048–2056. https://doi.org/10.1177/0956797613486982

Krapohl, E., Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J.-B., Asbury, K., Harlaar, N., Kovas, Y., Dale, P. S., & Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences, 111*(42), 15273–15278. https://doi.org/10.1073/pnas.1408777111

Kwiatkowska-White, B., Kirby, J. R., & Lee, E. A. (2016). A longitudinal study of reading comprehension achievement from grades 3 to 10: Investigating models of stability, cumulative growth, and compensation. *Journal of Psychoeducational Assessment, 34*(2), 153–165. https://doi.org/10.1177/0734282915593188

Kyriakides, L., Creemers, B. P. M., & Charalambous, E. (2019). Searching for differential teacher and school effectiveness in terms of student socioeconomic status and gender: implications for promoting equity. *School Effectiveness and School Improvement, 30*(3), 286–308. https://doi.org/10.1080/09243453.2018.1511603

Lareau, A. (2003). *Unequal childhoods: Class, Race, and family life.* University of California Press.

Larsen, S. A., Byrne, B., Little, C. W., Coventry, W. L., Ho, C. S., Olson, R. K., & Stevenson, A. (2019). Identical genes, unique environments: A qualitative exploration of persistent monozygotic-twin discordance in literacy and numeracy. *Frontiers in Education, 4*(21), 1-14. https://doi.org/10.3389/FEDUC.2019.00021

Lazear, E. P. (2002). Educational production. *The Quarterly Journal of Economics, CXVII*(1), 1–37. https://doi.org/10.1162/00335530360535135

Legewie, J., & DiPrete, T. A. (2012). School context and the gender gap in educational achievement. *American Sociological Review, 77*(3), 463–485. https://doi.org/10.1177/0003122412440802

Lei, Z., & Lundberg, S. (2020). Vulnerable boys: Short-term and long-term gender differences in the impacts of adolescent disadvantage. *Journal of Economic Behavior and Organization, 178*, 424–448. https://doi.org/10.1016/j.jebo.2020.07.020

Ligthart, L., van Beijsterveldt, C. E. M., Kevenaar, S. T., de Zeeuw, E., van Bergen, E., Bruins, S., Pool, R., Helmer, Q., van Dongen, J., Hottenga, J.-J., van't Ent, D., Dolan, C. V, Davies, G. E., Ehli, E. A., Bartels, M., Willemsen, G., de Geus, E. J., & Boomsma, D. I. (2019). The Netherlands Twin Register: Longitudinal research based on twin and twin-family designs. *Twin Research and Human Genetics, 22*(6), 1–14. https://doi.org/10.1017/thg.2019.93

Lin, M. J. (2020). The social and genetic inheritance of educational attainment: Genes, parental education, and educational expansion. *Social Science Research, 86*(102387), 1-17. https://doi.org/10.1016/j.ssresearch.2019.102387

Loehlin, J. C., Paige, A. K., Ae, H., Turkheimer, E., Harden, K. P., & Turkheimer, E. (2009). The effect of assumptions about parental assortative mating and genotype-income correlation on estimates of genotype-environment interaction in the national merit twin study. *Behavior Genetics, 39*(2), 165–169. https://doi.org/10.1007/s10519-008-9253-9

Logan, J. A. R., Hart, S. A., Cutting, L., Deater-Deckard, K., Schatschneider, C., & Petrill, S. (2013). Reading development in young children: Genetic and environmental influences. *Child Development, 84*(6), 2131–2144. https://doi.org/10.1111/cdev.12104

Lyu, X., & Garrison, S. M. (2023). Effects of Genetic Relatedness of Kin Pairs on Univariate ACE Model Performance. *Twin Research and Human Genetics* (forthcoming) https://doi.org/https://doi.org/10.21203/rs.3.rs-2277334/v1

Ma, X. (2003). Sense of belonging to school: Can schools make a difference? *The Journal of Educational Research, 96*(6), 340–349. https://doi.org/10.1080/00220670309596617

Malanchini, M., Allegrini, A. G., Nivard, M. G., Biroli, P., Rimfeld, K., Cheesman, R., von Stumm, S., Demange, P. A., van Bergen, E., Grotzinger, A. D., Raffington, L., De la Fuente, J., Pingault, J.-B., Harden, K. P., Tucker-Drob, E. M., & Plomin, R. (2023). Genetic contributions of noncognitive skills to academic development. *bioRxiv.* https://doi.org/https://doi.org/10.1101/2023.04.03.535380

Marks, G. N. (2021). Should value-added school effects models include student- and school-level covariates? Evidence from Australian population assessment data. British *Educational Research Journal, 47*(1), 181–204. https://doi.org/10.1002/berj.3684

Marks, G. N., O'Connell, M., & O'Connell, M. (2023). The importance of parental ability for cognitive ability and student achievement: Implications for social stratification theory and practice. *Research in Social Stratification and Mobility, 83*(100762), 1-14. https://doi.org/10.1016/j.rssm.2023.100762

McAdams, T. A., Neiderhiser, J. M., Rijsdijk, F. V., Narusyte, J., Lichtenstein, P., & Eley, T. C. (2014). Accounting for genetic and environmental confounds in associations between parent and child characteristics: A systematic review of children-of-twins studies. *Psychological Bulletin, 140*(4), 1138–1173. https://doi.org/10.1037/a0036416

McArdle, J. J., Prescott, C. A., Hamagami, F., & Horn, J. L. (1998). A contemporary method for developmental-genetic analyses of age changes in intellectual abilities. *Developmental Neuropsychology, 14*(1), 69–114. https://doi.org/10.1080/87565649809540701

Medland, S. E., & Hatemi, P. K. (2009). Political science, biometric theory, and twin studies: A methodological introduction. *Political Analysis, 17*(2), 191–214. https://doi.org/10.1093/pan/mpn016

Meece, J. L., Glienke, B. B., & Burg, S. (2006). Gender and motivation. *Journal of School Psychology, 44*(5), 351–373. https://doi.org/10.1016/j.jsp.2006.04.004

Meyer, K. (2016). Why should we demand equality of educational opportunity? *Theory and Research in Education, 14*(3), 333–347. https://doi.org/10.1177/1477878516676709

Mills, M. C., & Tropf, F. C. (2020). Sociology, genetics, and the coming of age of sociogenomics. *Annual Review of Sociology, 46,* 553–581. https://doi.org/10.1146/annurev-soc-121919-054756

Ministerie van Onderwijs Cultuur en Wetenschap. (2022). *Staatscourant 18 juli 2022* [letter of government] (No. 19058; Issue 19058). Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. https://zoek.officielebekendmakingen.nl/staatscourant/stcrt-2022-19058.pdf

Ministry of Children and Education. (2023). *7-Point grading scale*. https://eng.uvm.dk/primary-and-lower-secondary-education/the-folkeskole/7-point-grading-scale#:~:text=The 7-point grading scale,(00 and -3).

Mönkediek, B. (2021). Trait-specific testing of the equal environment assumption: The case of school grades and upper secondary school attendance. *Journal of Family Research, 33*(1), 115–147. https://doi.org/10.20377/jfr-381

Mönkediek, B., Diewald, M., & Lang, V. (2023). Does social origin modify the heritability of cognitive ability? A close look at the relevance of different parental resources. *Research in Social Stratification and Mobility, 86*(100824), 1-18. https://doi.org/10.1016/j.rssm.2023.100824

Mönkediek, B., Schulz, W., Eichhorn, H., & Diewald, M. (2020). Is there something special about twin families? A comparison of parenting styles in twin and non-twin families. *Social Science Research, 90*(102441), 1-17. https://doi.org/10.1016/j.ssresearch.2020.102441

Montt, G. (2011). Cross-national differences in educational achievement inequality. *Sociology of Education, 84*(1), 49–68. https://doi.org/10.1177/0038040710392717

Morris, T. T., Davies, N. M., Dorling, D., Richmond, R. C., & Smith, G. D. (2018). Testing the validity of value-added measures of educational progress with genetic data. *British Educational Research Journal, 44*(5), 725–747. https://doi.org/10.1002/berj.3466

Murayama, K., Pekrun, R., Lichtenfeld, S., & vom Hofe, R. (2013). Predicting long-term growth in students' mathematics achievement: The unique contributions of motivation and cognitive strategies. *Child Development, 84*(4), 1475–1490. https://doi.org/10.1111/cdev.12036

Muthén, B. O., & Khoo, S.-T. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences, 10*(2), 73–101. https://doi.org/10.1016/S1041-6080(99)80135-6

Naayer, H. M., Spithoff, M., Osinga, M., Klitzing, N., Korpershoek, H., & Opdenakker, M.-C. (2016). *De overgang van primair naar voortgezet onderwijs in internationaal perspectief.* GION onderwijs/onderzoek. https://www.nro.nl/wp-content/uploads/2015/09/NRO_OPRO_De-overgang-van-po-naar-vo-in-internationaal-perspectief_Opdenakker_projectleider.pdf

NCO. (2022). *Codeboek data LVS-NCO-data* (Versie 1.2). Nationaal Cohortenonderzoek Onderwijs.

Neale, M. C., & Maes, H. M. (2004). *Methodology for genetic studies of twins and families*. Kluwer Academic Publishers. https://doi.org/10.1136/jmg.30.9.800-a

Nielsen, F. (2006). Achievement and ascription in educational attainment: Genetic and environmental influences on adolescent schooling. *Social Forces, 85*(1), 193–216. https://doi.org/10.1353/sof.2006.0135

Nielsen, F. (2016). The status-achievement process: insights from genetics. *Frontiers in Sociology, 1*(9), 1–15. https://doi.org/10.3389/fsoc.2016.00009

Nielsen, F., & Roos, J. M. (2015). Genetics of educational attainment and the persistence of privilege at the turn of the 21st century. *Social Forces, 94*(2), 535–561. https://doi.org/10.1093/sf/sov080

Noord, J. Van, Spruyt, B., Kuppens, T., & Spears, R. (2021). In the shadow of the schooled society: Feelings of misrecognition and the education ladder. *Social Problems, 70*(3), 831–848. https://doi.org/10.1093/socpro/spab034

Nusche, D., Braun, H., Halász, G., & Santiago, P. (2014). *OECD Reviews of Evaluation and Assessment in Education: Netherlands 2014*. OECD Publishing. https://doi.org/http://dx.doi.org/10.1787/9789264211940-en

NVOM. (2019). Samen of apart? U zult het zelf moeten ontdekken. *Meerlingen Magazine, 4*. https://www.nvom.nl/assets/uploads/2022/10/schoolspecial-2019-web.pdf

OECD. (2014). *Education policy outlook Netherlands*. OECD Publishing. www.oecd.org/edu/policyoutlook.htm

OECD. (2016). *Netherlands 2016: Foundations for the future. reviews of national policies for education*. OECD Publishing. https://doi.org/10.1787/9789264257658-en

OECD. (2017). *Educational opportunity for all: Overcoming inequality throughout the life course.* OECD Publishing. http://dx.doi.org/10.1787/9789264287457-en

OECD. (2018). *Measuring improvements in learning outcomes. Best practices to assess the value-added of schools*. OECD Publishing. https://doi.org/10.1787/9789264050259-en

OECD. (2019). PISA 2018 results (Volume II): Where all students can succeed. OECD Publishing. https://doi.org/10.1787/19963777

Opdenakker, M.-C. (2021). Need-supportive and need-thwarting teacher behavior: Their importance to boys' and girls' academic engagement and procrastination behavior. *Frontiers in Psychology, 21*(12), 1-17. https://doi.org/10.3389/FPSYG.2021.628064

Opdenakker, M.-C., & Van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education*? British Educational Research Journal, 33*(2), 179–206. https://doi.org/10.1080/01411920701208233

Opdenakker, M. C., & Damme, J. (2001). Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research Journal, 27*(4), 407–432. https://doi.org/10.1080/01411920120071434

Parcel, T. L., & Dufur, M. J. (2001). Capital at home and at school: Effects on student achievement. *Social Forces, 79*(3), 881–911. https://doi.org/10.1353/sof.2001.0021

Paris, J. (1999). *Genetics and psychopathology: Predisposition–stress interactions.* American Psychiatric Press.

Passaretta, G., & Skopek, J. (Eds.). (2018). *Roots and development of achievement gaps. A longitudinal assessment in selected European countries*. ISOTIS Report. http://www.isotis.org/wp-content/uploads/2019/01/ISOTIS_D1.3-Roots-and-Development-of-Achievement-Gaps.pdf

Passaretta, G., & Skopek, J. (2021). Does schooling decrease socioeconomic inequality in early achievement? A differential exposure approach. *American Sociological Review, 86*(6), 1017–1042. https://doi.org/10.1177/00031224211049188

Pennington, B. F., Mcgrath, L. M., Rosenberg, J., Barnard, H., Smith, S. D., Munroe, H. B., Willcutt, E. G., Friend, A., Defries, J. C., & Olson, R. K. (2009). Gene × environment interactions in reading disability and attention-deficit/hyperactivity disorder. *Developmental Psychology, 45*(1), 77–89. https://doi.org/10.1037/a0014549

Petrill, S. A., Hart, S. A., Harlaar, N., Logan, J., Justice, L. M., Schatschneider, C., Thompson, L., Dethorne, L. S., Deater-Deckard, K., & Cutting, L. (2010). Genetic and environmental influences on the growth of early reading skills. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 51*(6), 660–667. https://doi.org/10.1111/j.1469-7610.2009.02204.x

Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). individual differences in reading development: A review of 25 years of empirical research on Matthew effects in reading. *Review of Educational Research, 84*(2), 203–244. https://doi.org/10.3102/0034654313509492

Plomin, R. (1986). Multivariate analysis and developmental behavioral genetics: Developmental change as well as continuity. *Behavior Genetics, 16*(1), 25–43. https://doi.org/10.1007/BF01065477

Plomin, R., Asbury, K., & Dunn, J. (2001). Why are children in the same family so different? Nonshared environment a decade later. *The Canadian Journal of Psychiatry, 46*(3), 225–233. http://journals.sagepub.com/doi/pdf/10.1177/070674370104600302

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84*(2), 309–322. https://doi.org/10.1037/0033-2909.84.2.309

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2008). *Behavioral Genetics.* Worth Publishers.

Plomin, R., & Spinath, F. M. (2004). Intelligence: Genetics, genes, and genomics. *Journal of Personality and Social Psychology, 86*(1), 112–129. https://doi.org/10.1037/0022-3514.86.1.112

Pluess, M., & Belsky, J. (2013). Vantage sensitivity: Individual differences in response to positive experiences. *Psychological Bulletin, 139*(4), 901–916. https://doi.org/10.1037/a0030196

Pokropek, A., & Sikora, J. (2015). Heritability, family, school and academic achievement in adolescence. *Social Science Research, 53*, 73–88. https://doi.org/10.1016/j.ssresearch.2015.05.005

Polderman, T. J. C., Bartels, M., Verhulst, F. C., Huizink, A. C., Van Beijsterveldt, C. E. M., & Boomsma, D. I. (2010). No effect of classroom sharing on educational achievement in twins: A prospective, longitudinal cohort study. *Journal of Epidemiology and Community Health, 64*(1), 36–40. https://doi.org/10.1136/jech.2009.091629

Posthuma, D., & Boomsma, D. I. (2000). A note on the statistical power in extended twin designs. *Behavior Genetics, 30*(2), 147–158. https://doi.org/10.1023/A:1001959306025

Preedy, P. (1999). Meeting the educational needs of preschool and primary aged twins and higher multiples. In A. C. Sandbank (Ed.), *Twin and Triplet Psychology* (pp. 70–99). Routledge.

Purcell, S. (2002). Variance components models for gene–environment interaction in twin analysis. *Twin Research, 5*(06), 554–571. https://doi.org/10.1375/twin.5.6.554

Quadlin, N. (2019). Sibling achievement, sibling gender, and beliefs about parental investment: Evidence from a national survey experiment. *Social Forces, 97*(4), 1603–1630. https://doi.org/10.1093/sf/soy095

Raabe, I. J., & Wölfer, R. (2019). What is going on around you: Peer milieus and educational aspirations. *European Sociological Review, 35*(1), 1–14. https://doi.org/10.1093/esr/jcy048

Raudenbush, S. W. (2008). Advancing educational policy by advancing research on instruction. *American Educational Research Journal, 45*(1), 206–230. https://doi.org/10.3102/0002831207312905

Ready, D. D. (2013). Associations between student achievement and student learning: implications for value-added school accountability models. *Educational Policy, 27*(1), 92–120. https://doi.org/10.1177/0895904811429289

Reardon, S. F. (2019). Educational opportunity in early and middle childhood: Using full population administrative data to study variation by place and age. *RSF: The Russell Sage Foundation Journal of the Social Sciences, 5*(2), 40–68. https://doi.org/10.7758/RSF.2019.5.2.03

Reardon, S. F., & Owens, A. (2014). 60 years after brown: Trends and consequences of school segregation. *Annual Review of Sociology, 40*, 199–218. https://doi.org/10.1146/annurev-soc-071913-043152

Rende, R., & Plomin, R. (1992). Diathesis-stress models of psychopathology: A quantitative genetic perspective. *Applied and Preventive Psychology, 1*(4), 177–182. https://doi.org/10.1016/S0962-1849(05)80123-4

Reynolds, C. A., Finkel, D., Gatz, M., & Pedersen, N. L. (2002). Sources of influence on rate of cognitive change over time in Swedish twins: An application of latent growth models. *Experimental Aging Research, 28*(4), 407–433. https://doi.org/10.1080/03610730290103104

Rietveld, M. J. H., Posthuma, D., Dolan, C. V, & Boomsma, D. I. (2003). ADHD: Sibling interaction or dominance: An evaluation of statistical power. *Behavior Genetics, 33*(3), 247–254. https://doi.org/10.1023/A:1023490307170

Rijsdijk, F. V, & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Brief Bioinform, 3*(2), 119–133. https://doi.org/10.1093/bib/3.2.119

Ritzen, J. M. M., Van Dommelen, J., & De Vijlder, F. J. (1997). School finance and school choice in the Netherlands. *Pergamon Economics of Education Review, 16*(3), 329–335.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458. https://doi.org/10.1002/polq.12145

Rjosk, C. (2022). Dispersion of student achievement and classroom composition. In T. Nilsen, A. Stancel-Piatak, & J.-E. Gustafsson (Eds.), *International handbook of comparative large-scale studies in education: perspectives, methods and findings* (pp. 1399–1431). Springer. https://doi.org/10.1007/978-3-030-88178-8_47

Robert, P. (2010). Social origin, school choice, and student performance. *Educational Research and Evaluation, 16*(2), 107–129. https://doi.org/10.1080/13803611.2010.484972

Robin, M., & Casati, I. (1994). Are twins different from singletons during early childhood? *Early Development and Parenting, 3*(4), 211–221. https://doi.org/10.1002/edp.2430030404

Rodgers, J. L., Rowe, D. C., & May, K. (1994). DF analysis of NLSY IQ/achievement data: Nonshared environmental influences. *Intelligence, 19*(2), 157–177. https://doi.org/10.1016/0160-2896(94)90011-6

Roemer, J. E. 1998. *Equality of Opportunity.* Harvard University Press.

Rohrer, J. M., & Adams, D. (2021). Precise answers to vague questions: Issues with interactions. *Advances in Methods and Practices in Psychological Science, 4*(2), 1–19. https://doi.org/10.1177/25152459211007368

Rowe, D. C., Jacobson, K. C., & Van den Oord, E. J. C. G. (1999). Genetic and environmental influences on vocabulary IQ: Parental education level as moderator. *Child Development, 70*(5), 1151–1162. https://doi.org/10.1111/1467-8624.00084

Ruks, M. (2022). Investigating the mechanisms of G × SES interactions for education. *Research in Social Stratification and Mobility, 81*(100730), 1–19. https://doi.org/10.1016/J.RSSM.2022.100730

Rumberger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record, 107*(9), 1999–2045. https://doi.org/10.1111/j.1467-9620.2005.00583.x

Sameroff, A. J. (2009). *The transactional model of development: How children and contexts shape each other.* American Psychological Association. http://www.apa.org/pubs/books/4316113.aspx

Scarr-Salapatek, S. (1971). Race, social class, and IQ. *Science, 174*(4016), 1285–1295. https://doi.org/10.1126/science.174.4016.1285

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54*(2), 424–435.

Scheeren, L., Das, M., & Liefbroer, A. C. (2017). Intergenerational transmission of educational attainment in adoptive families in the Netherlands. *Research in Social Stratification and Mobility, 48*, 10–19. https://doi.org/10.1016/j.rssm.2016.12.002

Scheerens, J. (2000). *Improving school effectiveness*. International Institute for Educational Planning.

Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness* (1st ed.). Pergamon. https://doi.org/10.1023/A:1003534107087

Scheerens, J., Ehren, M., Sleegers, P., & De Leeuw, R. (2012). *OECD Review on evaluation and assessment frameworks for improving school outcomes: Country background report for the Netherlands*. OECD Publishing. https://www.oecd.org/education/school/NLD_CBR_Evaluation_and_Assessment.pdf

Schmitz, L., & Conley, D. (2017). Modeling gene-environment interactions with quasi-natural experiments. *Journal of Personality, 85*(1), 10–21. https://doi.org/10.1111/JOPY.12227

Scholten, L., & Wolbers, M. H. J. (2018). High quality education: Bridging the socioeconomic achievement gap? *Mens en Maatschappij, 93*(4), 375–406. https://doi.org/10.5117/MEM2018.4.004.SCHO

Schröder, H., & Ganzeboom, H. B. G. (2014). Measuring and modelling level of education in european societies. *European Sociological Review, 30*(1), 119–136. https://doi.org/10.1093/esr/jct026

Schwabe, I., Janss, L., & van den Berg, S. M. (2017). Can we validate the results of twin studies? A census-based study on the heritability of educational achievement. *Frontiers in Genetics, 8*(160), 1–8. https://doi.org/10.3389/fgene.2017.00160

Seabrook, J. A., & Avison, W. R. (2010). Genotype-environment interaction and sociology: Contributions and complexities. *Social Science and Medicine, 70*(9), 1277–1284. https://doi.org/10.1016/j.socscimed.2010.01.016

Shanahan, M. J., & Hofer, S. M. (2005). Social context in gene–environment interactions: Retrospect and prospect. *The Journals of Gerontology: Series B, 60*(Special Issue), 65–76. https://doi.org/10.1093/geronb/60.special_issue_1.65

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417–453. https://doi.org/10.3102/00346543075003417

Smith-Woolley, E., Pingault, J.-B., Selzam, S., Rimfeld, K., Krapohl, E., von Stumm, S., Asbury, K., Dale, P. S., Young, T., Allen, R., Kovas, Y., & Plomin, R. (2018). Differences in exam performance between pupils attending selective and non-selective schools mirror the genetic differences between them. *Npj Science of Learning, 3*(3), 3. https://doi.org/10.1038/s41539-018-0019-8

Sociaal-Economische Raad. (2021). *Gelijke kansen in het onderwijs: Structureel investeren in kansengelijkheid voor iedereen*. Sociaal-Economische Raad. https://www.ser.nl/-/media/ser/downloads/adviezen/2021/gelijke-kansen-in-onderwijs.pdf

Sørenson, A. B., & Hallinan, M. T. (2016). A reconceptualization of school effects. *Sociology of Education, 50*(4), 273–289.

Spinath, F. M., Price, T. S., Dale, P. S., & Plomin, R. (2004). The genetic and environmental origins of language disability and ability. *Child Development, 75*(2), 445–454. https://doi.org/10.1111/J.1467-8624.2004.00685.X

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Research in Reading Quarterly, 21*(4), 360–407. https://doi.org/10.1177/0022057409189001-204

Stienstra, K., Knigge, A., Maas, I., Zeeuw, E. L. De, & Boomsma, D. I. (2023). Are classrooms equalizers or amplifiers of inequality? A genetically informative investigation of educational performance. *European Sociological Review, 39*(5), 708–723. https://doi.org/10.1093/esr/jcac054

Stienstra, K., Maas, I., Knigge, A., & Schulz, W. (2021). Resource compensation or multiplication? The interplay between cognitive ability and social origin in explaining educational attainment. *European Sociological Review, 37*(2), 186–200. https://doi.org/10.1093/esr/jcaa054

Strello, A., Strietholt, R., Steinmann, I., & Siepmann, C. (2021). Early tracking and different types of inequalities in achievement: difference-in-differences evidence from 20 years of large-scale assessments. *Educational Assessment, Evaluation and Accountability, 33*(1), 139–167. https://doi.org/10.1007/s11092-020-09346-4

Strietholt, R. (2014). Studying educational inequality: Reintroducing normative notions. In R. Strietholt, W. Bos, J.-E. Gustafsson, & M. Rosén (Eds.), *Educational Policy Evaluation Through International Comparative Assessments* (pp. 51–58). Waxmann Verlag.

Swart, N. M., Gubbels, J., in 't Zandt, M., Wolbers, M. H. J., & Segers, E. (2023). *PIRLS-2021: Trends in leesprestaties, leesattitude en leesgedrag van tienjarigen uit Nederland*. Expertisecentrum Nederlands. https://www.expertisecentrumnederlands.nl/pirls-2021-resultaten/download/PIRLS-2021_Rapportage.pdf

Swift, A., & Marshall, G. (1997). Meritocratic equality of opportunity: Economic efficiency, social justice, or both? *Policy Studies, 18*(1), 35–48. https://doi.org/10.1080/01442879708423720

Sykes, B., & Kuyper, H. (2013). School segregation and the secondary-school achievements of youth in the Netherlands. *Journal of Ethnic and Migration Studies, 39*(10), 1699–1716. https://doi.org/10.1080/1369183X.2013.833707

Tannock, S. (2008). The problem of education-based discrimination. *British Journal of Sociology of Education, 29*(5), 439–449. https://doi.org/10.1080/01425690802326846

Taubman, P. (1976). The determinants of earnings: Genetics, family, and other environments; a study of white male twins. *American Economic Review, 66*(5), 858–870.

Taylor, J., Erbeli, F., Hart, S. A., & Johnson, W. (2020). Early classroom reading gains moderate shared environmental influences on reading comprehension in adolescence. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 61*(6), 689–698. https://doi.org/10.1111/jcpp.13134

Taylor, J., Roehrig, A. D., Soden Hensler, B., Connor, C. M., & Schatschneider, C. (2010). Teacher quality moderates the genetic effects on early reading. *Science, 328*(5977), 512–514. https://doi.org/10.1126/science.1186149

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research, 83*(3), 357–385. https://doi.org/10.3102/0034654313483907

Thrupp, M., Lauder, H., & Robinson, T. (2002). School composition and peer effects. *International Journal of Educational Research, 37,* 483–504. https://doi.org/10.1016/S0883-0355(03)00016-8

Timmermans, A., & van der Werf, G. (2017). Accounting for previous performance of students by means of growth curves analyses to estimate the size, stability, and consistency of school effects. *Educational Research and Evaluation, 23*(5–6), 221–246. https://doi.org/10.1080/13803611.2017.1455300

Tomesen, M., Engelen, R., & Hiddink, L. (2018). *Wetenschappelijke verantwoording. Begrijpend lezen 3.0 voor groep 7*. Cito. https://cito.nl/media/m0diuahs/3-cito-lvs-begrijpend-lezen-3-0-gr7-wet-verantwoording.pdf

Tourrette, C., Robin, M., & Josse, D. (1989). Treating twins as individuals: Maternal educative practices. *European Journal of Psychology of Education, 4*(2), 269–283. https://doi.org/10.1007/BF03172610

Trejo, S., Belsky, D. W., Boardman, J. D., Freese, J., Harris, K. M., Herd, P., Sicinski, K., & Domingue, B. W. (2018). Schools as moderators of genetic associations with life course attainments: Evidence from the WLS and add health. *Sociological Science, 5,* 513–540. https://doi.org/10.15195/V5.A22

Tucker-Drob, E. M. (2012). Preschools reduce early academic-achievement gaps: A longitudinal twin approach. *Psychological Science, 23*(3), 310–319. https://doi.org/10.1177/0956797611426728

Tucker-Drob, E. M., & Bates, T. C. (2016). Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychological Science, 27*(2), 138–149. https://doi.org/10.1177/0956797615612727

Tucker-Drob, E. M., Briley, D. A., & Harden, K. P. (2013). Genetic and environmental influences on cognition across development and context. *Current Directions in Psychological Science, 22*(5), 349–355. https://doi.org/10.1177/0963721413485087

Tucker-Drob, E. M., & Harden, K. P. (2012a). Early childhood cognitive development and parental cognitive stimulation: Evidence for reciprocal gene-environment transactions. *Developmental Science, 15*(2), 250–259. https://doi.org/10.1111/j.1467-7687.2011.01121.x

Tucker-Drob, E. M., & Harden, K. P. (2012b). Intellectual interest mediates gene x socioeconomic status interaction on adolescent academic achievement. *Child Development, 83*(2), 743–757. https://doi.org/10.1111/j.1467-8624.2011.01721.x

Tucker-Drob, E. M., & Harden, K. P. (2012c). Learning motivation mediates gene-by-socioeconomic status interaction on mathematics achievement in early childhood. *Learning and Individual Differences, 22*(1), 37–45. https://doi.org/10.1016/j.lindif.2011.11.015

Turkheimer, E., D'Onofrio, D. M., Maes, H. H., & Eaves, L. J. (2005). Analysis and interpretation of twin studies showing measures of the shared environment. *Child Development, 76*(6), 1217–1233. https://doi.org/10.1111/j.1467-8624.2005.00846.x

Turkheimer, E., & Horn, E. E. (2014). Interactions between socioeconomic status and components of variation in cognitive ability. In D. Finkel & C. A. Reynolds (Eds.), *Behavior genetics of cognition across the lifespan* (pp. 41–68). Springer.

UNESCO. (2018). *Handbook on measuring equity in education.* UNESCO Institute for Statistics. https://uis.unesco.org/sites/default/files/documents/handbook-measuring-equity-education-2018-en.pdf

Unicef. (2020). *Addressing the learning crisis: An urgent need to better finance education for the poorest children*. Unicef. https://www.unicef.org/media/63896/file/Addressing-the-learning-crisis-advocacy-brief-2020.pdf

Unicef Office of Research. (2018). *An unfair start: Inequality in children's education in rich countries* (Innocenti Report Card 15). Unicef Office of Research. https://www.unicef-irc.org/publications/995-an-unfair-start-education-inequality-children.html

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. https://sdgs.un.org/2030agenda

United Nations. (2019). *The sustainable development goals report 2019*. United Nations Department of Economic and Social Affairs. https://doi.org/10.18356/55eb9109-en

Van Beijsterveldt, C. E. M., Groen-Blokhuis, M., Hottenga, J. J., Franić, S., Hudziak, J. J., Lamb, D., Huppertz, C., De Zeeuw, E., Nivard, M., Schutte, N., Swagerman, S., Glasner, T., Van Fulpen, M., Brouwer, C., Stroet, T., Nowotny, D., Ehli, E. A., Davies, G. E., Scheet, P., … Boomsma, D. I. (2013). The young Netherlands twin register (YNTR): Longitudinal twin and family studies in over 70,000 children. *Twin Research and Human Genetics, 16*(1), 252–267. https://doi.org/10.1017/thg.2012.118

Van Boxtel, H., Engelen, R., & De Wijs, A. (2010). *Wetenschappelijke verantwoording van de Eindtoets Basisonderwijs 2010*. Cito. https://www.beteronderwijsnederland.nl/wp-content/uploads/2013/04/cito_wetenschappelijke_verantwoording_eindtoets.pdf

Van de Gaer, E., Pustjens, H., Van Damme, J., & De Munter, A. (2006). The gender gap in language achievement: The role of school-related attitudes of class groups. *Sex Roles, 55*(5–6), 397–408. https://doi.org/10.1007/s11199-006-9092-1

Van de Werfhorst, H. G. (2014). Changing societies and four tasks of schooling: Challenges for strongly differentiated educational systems. *International Review of Education, 60*(1), 123–144. https://doi.org/10.1007/s11159-014-9410-8

Van de Werfhorst, H. G., & Mijs, J. J. B. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology, 36*(1), 407–428. https://doi.org/10.1146/annurev.soc.012809.102538

Van Hek, M., Kraaykamp, G., & Pelzer, B. (2017). Do schools affect girls' and boys' reading performance differently? A multilevel study on the gendered effects of school resources and school practices. *School Effectiveness and School Improvement, 29*(1), 1–21. https://doi.org/10.1080/09243453.2017.1382540

Van Houtte, M. (2004). Gender context of the school and study culture, or how the presence of girls affects the achievement of boys. *Educational Studies, 30*(4), 409–423. https://doi.org/10.1080/0305569042000310336

Van Leeuwen, M., Van den Berg, S. M., Van Beijsterveldt, T. C. E. M., & Boomsma, D. I. (2005). Effects of twin separation in primary school. *Twin Research and Human Genetics, 8*(4), 384–391. https://doi.org/10.1375/1832427054936682

Vantieghem, W., & Van Houtte, M. (2018). Differences in study motivation within and between genders: An examination by gender typicality among early adolescents. *Youth and Society, 50*(3), 377–404. https://doi.org/10.1177/0044118X15602268

Verbakel, E., & Kalmijn, M. (2014). Assortative mating among Dutch married and cohabiting same-sex and different-sex couples. *Journal of Marriage and Family, 76*(1), 1–12. https://doi.org/10.1111/jomf.12084

Verhulst, B., & Hatemi, P. K. (2013). Gene-environment interplay in twin models. *Political Analysis, 21*(3), 368–389. https://doi.org/10.1093/pan/mpt005

Verhulst, B., Prom-Wormley, E., Keller, M., Medland, S., & Neale, M. C. (2019). Type I error rates and parameter bias in multivariate behavioral genetic models. *Behavior Genetics, 49*(1), 99–111. https://doi.org/10.1007/s10519-018-9942-y

Vitthala, S., Gelbaya, T. A., Brison, D. R., Fitzgerald, C. T., & Nardo, L. G. (2009). The risk of monozygotic twins after assisted reproductive technology: a systematic review and meta-analysis. *Human Reproduction Update, 15*(1), 45–55. https://doi.org/10.1093/HUMUPD/DMN045

Von Hippel, P. T., & Hamrock, C. (2019). Do test score gaps grow before, during, or between the school years? Measurement artifacts and what we can know in spite of them. *Sociological Science, 6*, 43–80. https://doi.org/10.15195/V6.A3

Von Hippel, P. T., Workman, J., & Downey, D. B. (2018). inequality in reading and math skills forms mainly before kindergarten: A replication, and partial correction, of "Are schools the great equalizer?" *Sociology of Education, 91*(4), 323–357. https://doi.org/10.1177/0038040718801760

Walhout, J., & Scholtus, S. (2022). *Herijking onderwijsachterstandenindicator primair onderwijs 2021*. Centraal Bureau voor de Statistiek. https://www.cbs.nl/nl-nl/longread/rapportages/2022/herijking-onderwijsachterstandenindicator-primair-onderwijs-2021

Warrington, M., Younger, M., & Williams, J. (2000). Student attitudes, image and the gender gap. *British Educational Research Journal, 26*(3), 393–407. https://doi.org/10.1080/01411920050030914

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89–122.

Webbink, D., Posthuma, D., Boomsma, D. I., de Geus, E. J. C., & Visscher, P. M. (2008). Do twins have lower cognitive ability than singletons? *Intelligence, 36*(6), 539–547. https://doi.org/10.1016/j.intell.2007.12.002

Webbink, D., Roeleveld, J., & Visscher, P. M. (2006). Identification of twin pairs from large population-based samples. *Twin Research and Human Genetics, 9*(4), 496–500. https://doi.org/10.1375/183242706778024892

Weinberg, W. (1901). Beiträge zur Physiologie und Pathologie der Mehrlingsgeburten beim Menschen. *Archiv für die gesammte Physiologie des Menschen und der Thiere, 88*(6–8), 346–430. https://doi.org/10.1007/BF01657695

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology, 102*(1), 43–53. https://doi.org/10.1037/a0016738

Wentzel, K. R., & Muenks, K. (2016). Peer influence on students' motivation, academic achievement, and social behavior. In K. R. Wentzel & G. B. Ramani (Eds.), *Handbook of social influences in school contexts: Social-emotional, motivation, and cognitive outcomes* (pp. 13–30). Routledge.

Workman, J., Von Hippel, P. T., & Merry, J. (2023). Findings on summer learning loss often fail to replicate, even in recent data. *Sociological Science, 10*(1), 251–285. https://doi.org/10.15195/v10.a8

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*(1), 57–67. https://doi.org/10.1023/a:1007999204543

Younger, M., Warrington, M., & Williams, J. (1999). The gender gap and classroom interactions: Reality and rhetoric? *British Journal of Sociology of Education, 20*(3), 325–341. https://doi.org/10.1080/01425699995290

Ziegler, A., Kuo, C. C., Eu, S. P., Gläser-Zikuda, M., Nuñez, M., Yu, H. P., & Harder, B. (2021). Equity gaps in education: Nine points toward more transparency. *Education Sciences, 11*(11), 1–12. https://doi.org/10.3390/educsci11110711

Zwier, D., & Geven, S. (2023). Knowing me, knowing you: Socio-economic status and (segregation in) peer and parental networks in primary school. *Social Networks, 74*, 127–138. https://doi.org/10.1016/j.socnet.2023.03.003

ACKNOWLEDGEMENTS

During the years of hard work on this dissertation, I have received constant support from a diverse group of people. I would like to use this opportunity to acknowledge the contribution of those who have made this academic journey possible and enjoyable.

First and foremost, I would like to express my sincere gratitude to my PhD supervisors, Ineke Maas and Antonie Knigge. I cannot thank you enough for your guidance, support, and encouragement over the many years (which have flown by for me). You have both been instrumental in shaping my research and helping me grow as a scholar and person. I have greatly enjoyed our meetings and have witnessed how you strengthen each other, providing me with some evidence for a multiplication effect after all. Ineke, I am very fortunate to have had you as a supervisor. Your critical and constructive feedback not only enhanced the quality of this dissertation but will also continue to resonate in the future. I highly appreciate the supportive and stimulating environment you created, which made me felt at home and encouraged my further development. Thank you also for the efforts in ensuring that we celebrated all the (small) successes along the way. Antonie, I find it very special and valuable that you have been my supervisor since my bachelor's thesis. Your encouragement and empathy have played an important role in fostering my confidence in my work and myself. I am in awe of how you are always able to provide such sharp and thought-provoking feedback. It often added an extra layer of complexity, challenging me to think beyond my current understanding, and I am grateful for that.

I extend my sincere gratitude to my co-authors for their contribution to this dissertation. Dorret Boomsma and Eveline de Zeeuw, thank you for enriching Chapter 2 with your expertise and collaboration. Kristian Karlson, thank you for the very fruitful collaboration on Chapter 4 and for welcoming me to Copenhagen. Thanks to you and the other people at the department I had a very pleasant and inspiring research visit.

I would like to offer my special thanks to the members of the Assessment Committee, Elsje van Bergen, Mads Jæger, Marco van Leeuwen, Daniel Oberski, and Inge de Wolf, for the time and effort to evaluate my manuscript.

Many colleagues from Utrecht University and the broader ICS community have contributed to the realization of this dissertation, being it through substantive input, by creating a pleasant working environment, or by providing space for relaxation. I am very grateful to everyone for this. This dissertation befitted greatly from the discussions and insightful feedback at the Migration and Social Stratification (MaSS) seminars and the ICS forum days, and I am thankful to all the participants and discussants at these events. I would also like to express my gratitude to Aleida, Ellen, Marjet, Marjolein, Pim, Sanne, and Sofie for their practical support.

My dearest office mates, Sanjana and Anne, I thank you for all the support, walks, chats, coffee breaks, snacks, and above all the 'gezelligheid' that we have had during and after working hours. I am very grateful for the photo album you made capturing all our great memories. Hopefully, we can print a second edition with new memories soon. Kasper and Philipp, thank you for sharing this PhD journey with me. I cherish the enjoyable dinners, outings, and everyday working moments we have had. Ece, many thanks for all your kindness, encouragement, and enthusiasm (I wish I had adopted you as a cohort member sooner). Vardan, thank you for joining our small twin research team. It has been very pleasant working with you, not the least because of your optimism and for

creating reasons for celebration with coffee and cake. Renae, as ICS-colleague and friend I truly appreciate all the conversations we have had since SaSR. I am thankful for all the other fellow PhDs, including Amina, Ana, Christian, Cloé, David, Kevin, Jan-Willem (thanks for the great company in C2.06), Jos (for the coffee and all our talks, research-related and not), Lea, Lian, Luuk, Nick, Paula (for taking such good care of my apartment when I was abroad and our talks that followed), and Rita. You all have made my PhD experience more enjoyable and memorable through the coffee and lunch breaks, walks, dinners, and other PhD-activities.

I joined the LOF dance crew at the start of my PhD, and I could not think of a better way to spend my free time. Dancing helped maintain balance, both figuratively and literally. To Charlotte and the LOFly dancers, thank you for all the trainings, shows, and great memories throughout the past years. A special thanks to Amber, Bente, José, Liv-Maritt, and Rosana who 'danced my PhD'.

Finally, I am highly indebted to my dear family and friends who provided me with their support and welcome distractions from work when needed. I would like to mention and thank some of them in particular. Melanie, we know each other since high school, and our friendship is very dear to me. Thank you for always being there. Marlijn, thank you for your support and friendship, as well as all the performances we have attended and dinners we have had. My dad Fred, mom Lonneke, and sisters Sanne and Kiki, you all played an influential role in defining the person I am today. I am deeply thankful for your support and understanding throughout my studies. Special thanks to Sanne for the great cover design. Last but not least, dearest Richard, thank you for standing by my side, as a paranymph but especially in life.

Thank you all for making this journey memorable. Your collective influence has not only shaped this dissertation but also shaped me.

ABOUT THE AUTHOR

Kim Stienstra was born in Harderwijk, the Netherlands, on 11 August 1995. She obtained her bachelor's degree in Sociology in 2016, along with the honors program of the Faculty of Social and Behavioral Sciences at Utrecht University. In 2018, she graduated cum laude from the research master Sociology and Social Research at the same university. Together with prof. dr. Ineke Maas and dr. Antonie Knigge, she received the Talent Grant from the Dutch Research Council (NWO) for the PhD project that resulted in this dissertation. She conducted the research as PhD candidate at the Interuniversity Center for Social Science Theory and Methodology and the Department of Sociology at Utrecht University. As part of her PhD, she spent three months as a visiting researcher at the Department of Sociology at the University in Copenhagen in 2021. As of September 2023, Kim has been working as a Max Weber Fellow at the European University Institute in Florence, Italy.

List of peer-reviewed publications

Stienstra, K., Knigge, A., Maas, I., de Zeeuw, E. L. & Boomsma, D. I. (2023). Are classrooms equalizers or amplifiers of inequality? A genetically informative investigation of educational performance. *European Sociological Review, 39*(5), 708-723. https://doi.org/10.1093/esr/jcac054

Stienstra, K. & Karlson, K. B. (2023). The nature-nurture of academic achievement at the intersection between gender, family background, and school context. *Social Science Research, 111*(102870), 1-14. https://doi.org/10.1016/j.ssresearch.2023.102870

Stienstra, K. & Knigge, A. (2023). Indirect pathways of multigenerational persistence: The role of uncles and assortative mating in the Netherlands, 1857-1922. *The History of the Family, 28*(1), 67-94. https://doi.org/10.1080/1081602X.2022.2084441

Knigge, A., Maas, I., Stienstra, K., de Zeeuw, E. L. & Boomsma, D. I. (2022). Who benefits from tracking and why? Gene-environment interactions in educational attainment. *npj Science of Learning, 7*(6), 1-6. https://doi.org/10.1038/s41539-022-00122-1

Nollet, M.A.Y. & Stienstra, K. (2022) Het effect van leerkrachtverwachtingen op het onderpresteren van jongens in lezen. *Mens en Maatschappij, 97*(2), 126-153. https://doi.org/10.5117/MEM2022.2.002.NOLL

Wafelaar, F. & Stienstra, K. (2021) Een multilevel onderzoek naar het effect van tablet gebruik in Nederlandse basisscholen op leesprestaties. *Mens en Maatschappij, 96*(1), 31-66. https://doi.org/10.5117/MEM2021.1.003.WAFE

Stienstra, K., Maas, I., Knigge, A. & Schulz, W. (2021) Resource compensation or multiplication? The interplay between cognitive ability and social origin in explaining educational attainment. *European Sociological Review, 37*(2), 186-200. https://doi.org/10.1093/esr/jcaa054

ICS DISSERTATION SERIES

The ICS series presents dissertations of the Interuniversity Center for Social Science Theory and Methodology. Each of these studies aims at integrating explicit theory formation with state-of-the-art empirical research or at the development of advanced methods for empirical research. The ICS was founded in 1986 as a cooperative effort of the universities of Groningen and Utrecht. Since 1992, the ICS has expanded to the University of Nijmegen and, since 2017, to the University of Amsterdam. Most of the projects are financed by the participating universities or by the Dutch Research Council. The international composition of the ICS graduate students is mirrored in the increasing international orientation of the projects and thus of the ICS series itself.

1. Kees van Liere (1990), *"Lastige leerlingen: Een empirisch onderzoek naar sociale oorzaken van probleemgedrag op basisscholen."* Amsterdam: Thesis Publishers
2. Marco van Leeuwen (1990), *"Bijstand in Amsterdam, ca. 1800-1850: Armenzorg als beheersings en overlevingsstrategie."* ICS dissertation, Utrecht
3. Ineke Maas (1990), *"Deelname aan podiumkunsten via de podia, de media en actieve beoefening: Substitutie of leereffecten?"* Amsterdam: Thesis Publishers
4. Marjolein Broese van Groenou (1991), *"Gescheiden netwerken: De relaties met vrienden en verwanten na echtscheiding."* Amsterdam: Thesis Publishers
5. Jan van den Bos (1991), *"Dutch EC policy making: A model guided approach to coordination and negotiation."* Amsterdam: Thesis Publishers
6. Karin Sanders (1991), *"Vrouwelijke pioniers: Vrouwen en mannen met een 'mannelijke' hogere beroepsopleiding aan het begin van hun loopbaan."* Amsterdam: Thesis Publishers
7. Sjerp de Vries (1991), *"Egoism, altruism, and social justice: Theory and experiments on cooperation in social dilemmas."* Amsterdam: Thesis Publishers
8. Ronald Batenburg (1991), *"Automatisering in bedrijf."* Amsterdam: Thesis Publishers
9. Rudi Wielers (1991), *"Selectie en allocatie op de arbeidsmarkt. Een uitwerking voor de informele en geïnstitutionaliseerde kinderopvang."* Amsterdam: Thesis Publishers
10. Gert Westert (1991), *"Verschillen in ziekenhuisgebruik."* ICS dissertation, Groningen
11. Hanneke Hermsen (1992), *"Votes and policy preferences: Equilibria in party systems."* Amsterdam: Thesis Publishers
12. Cora Maas (1992), *"Probleemleerlingen in het basisonderwijs."* Amsterdam: Thesis Publishers
13. Ed Boxman (1992), *"Contacten en carrière : Een empirisch theoretisch onderzoek naar de relatie tussen sociale netwerken en arbeidsmarktposities"* Amsterdam: Thesis Publishers
14. Conny Taes (1992), *"Kijken naar banen: Een onderzoek naar de inschatting van arbeidsmarktkansen bij schoolverlaters uit het middelbaar beroepsonderwijs."* Amsterdam: Thesis Publishers
15. Peter van Roozendaal (1992), *"Cabinets in multi party democracies: The effect of dominant and central parties on cabinet composition and durability."* Amsterdam: Thesis Publishers
16. Marcel van Dam (1992), *"Regio zonder regie: Verschillen in en effectiviteit van gemeentelijk arbeidsmarktbeleid."* Amsterdam: Thesis Publishers
17. Tanja van der Lippe (1993), *"Arbeidsverdeling tussen mannen en vrouwen."* Amsterdam: Thesis Publishers
18. Marc Jacobs (1993), *"Software: Kopen of kopiëren? Een sociaal wetenschappelijk onderzoek onder PC gebruikers."* Amsterdam: Thesis Publishers
19. Peter van der Meer (1993), *"Verdringing op de Nederlandse arbeidsmarkt: Sector- en sekseverschillen."* Amsterdam: Thesis Publishers
20. Gerbert Kraaykamp (1993), *"Over lezen gesproken: Een studie naar sociale differentiatie in leesgedrag."* Amsterdam: Thesis Publishers
21. Evelien Zeggelink (1993), *"Strangers into friends: The evolution of friendship networks using an individual oriented modeling approach."* Amsterdam: Thesis Publishers
22. Jaco Berveling (1994), *"Het stempel op de besluitvorming: Macht, invloed en besluitvorming op twee Amsterdamse beleidsterreinen."* Amsterdam: Thesis Publishers

23. Wim Bernasco (1994), *"Coupled careers: The effects of spouse's resources on success at work."* Amsterdam: Thesis Publishers
24. Liset van Dijk (1994), *"Choices in child care: The distribution of child care among mothers, fathers and non parental care providers."* Amsterdam: Thesis Publishers
25. Jos de Haan (1994), *"Research groups in Dutch sociology."* Amsterdam: Thesis Publishers
26. Kwasi Boahene (1995), *"Innovation adoption as a socioeconomic process: The case of the Ghanaian cocoa industry."* Amsterdam: Thesis Publishers
27. Paul Ligthart (1995), *"solidarity in economic transactions: An experimental study of framing effects in bargaining and contracting."* Amsterdam: Thesis Publishers
28. Roger Leenders (1995), *"Structure and influence: Statistical models for the dynamics of actor attributes, network structure, and their interdependence."* Amsterdam: Thesis Publishers
29. Beate Volker (1995), *"Should auld acquaintance be forgot...? Institutions of communism, the transition to capitalism and personal networks: The case of East Germany."* Amsterdam: Thesis Publishers
30. Anneke Cancrinus-Matthijsse (1995), *"Tussen hulpverlening en ondernemerschap: Beroepsuitoefening en taakopvattingen van openbare apothekers in een aantal West-Europese landen."* Amsterdam: Thesis Publishers
31. Nardi Steverink (1996), *"Zo lang mogelijk zelfstandig: Naar een verklaring van verschillen in oriëntatie ten aanzien van opname in een verzorgingstehuis onder fysiek kwetsbare ouderen."* Amsterdam: Thesis Publishers
32. Ellen Lindeman (1996), *"Participatie in vrijwilligerswerk."* Amsterdam: Thesis Publishers
33. Chris Snijders (1996), *"Trust and commitments."* Amsterdam: Thesis Publishers
34. Koos Postma (1996), *"Changing prejudice in Hungary. A study on the collapse of state socialism and its impact on prejudice against Gypsies and Jews."* Amsterdam: Thesis Publishers
35. Jooske van Busschbach (1996), *"Uit het oog, uit het hart? Stabiliteit en verandering in persoonlijke relaties."* Amsterdam: Thesis Publishers
36. René Torenvlied (1996), *"Besluiten in uitvoering: Theorieën over beleidsuitvoering modelmatig getoetst op sociale vernieuwing in drie gemeenten."* Amsterdam: Thesis Publishers
37. Andreas Flache (1996), *"The double edge of networks: An analysis of the effect of informal networks on cooperation in social dilemmas."* Amsterdam: Thesis Publishers
38. Kees van Veen (1997), *"Inside an internal labor market: Formal rules, flexibility and career lines in a Dutch manufacturing company."* Amsterdam: Thesis Publishers
39. Lucienne van Eijk (1997), *"Activity and wellbeing in the elderly."* Amsterdam: Thesis Publishers
40. Róbert Gál (1997), *"Unreliability: Contract discipline and contract governance under economic transition."* Amsterdam: Thesis Publishers
41. Anne-Geerte van de Goor (1997), *"Effects of regulation on disability duration."* ICS dissertation, Utrecht
42. Boris Blumberg (1997), *"Das Management von Technologiekooperationen: Partnersuche und Verhandlungen mit dem Partner aus Empirisch Theoretischer Perspektive."* ICS dissertation, Utrecht
43. Marijke von Bergh (1997), *"Loopbanen van oudere werknemers."* Amsterdam: Thesis Publishers
44. Anna Petra Nieboer (1997), *"Life events and well-being: A prospective study on changes in well-being of elderly people due to a serious illness event or death of the spouse."* Amsterdam: Thesis Publishers
45. Jacques Niehof (1997), *"Resources and social reproduction: The effects of cultural and material resources on educational and occupational careers in industrial nations at the end of the twentieth century."* ICS dissertation, Nijmegen
46. Ariana Need (1997), *"The kindred vote: Individual and family effects of social class and religion on electoral change in the Netherlands, 1956 1994."* ICS dissertation, Nijmegen
47. Jim Allen (1997), *"Sector composition and the effect of education on Wages: An international comparison."* Amsterdam: Thesis Publishers
48. Jack Hutten (1998), *"Workload and provision of care in general practice: An empirical study of the relation between workload of Dutch general practitioners and the content and quality of their care."* ICS dissertation, Utrecht
49. Per Kropp (1998), *"Berufserfolg im Transformationsprozeß: Eine theoretisch empirische Studie über die Gewinner und Verlierer der Wende in Ostdeutschland."* ICS dissertation, Utrecht

50. Maarten Wolbers (1998), *"Diploma-inflatie en verdringing op de arbeidsmarkt: Een studie naar ontwikkelingen in de opbrengsten van diploma's in Nederland."* ICS dissertation, Nijmegen
51. Wilma Smeenk (1998), *"Opportunity and marriage: The impact of individual resources and marriage market structure on first marriage timing and partner choice in the Netherlands."* ICS dissertation, Nijmegen
52. Marinus Spreen (1999), *"Sampling personal network structures: Statistical inference in ego graphs."* ICS dissertation, Groningen
53. Vincent Buskens (1999), *"Social networks and trust."* ICS dissertation, Utrecht
54. Susanne Rijken (1999), *"Educational expansion and status attainment: A cross-national and over-time comparison."* ICS dissertation, Utrecht
55. Mérove Gijsberts (1999), *"The legitimation of inequality in state-socialist and market societies, 1987-1996."* ICS dissertation, Utrecht
56. Gerhard van de Bunt (1999), *"Friends by choice: An actor-oriented statistical network model for friendship networks through time."* ICS dissertation, Groningen
57. Robert Thomson (1999), *"The party mandate: Election pledges and government actions in the Netherlands, 1986 1998."* Amsterdam: Thela Thesis
58. Corine Baarda (1999), *"Politieke besluiten en boeren beslissingen: Het draagvlak van het mestbeleid tot 2000."* ICS dissertation, Groningen
59. Rafael Wittek (1999), *"Interdependence and informal control in organizations."* ICS dissertation, Groningen
60. Diane Payne (1999), *"Policy Making in the European Union: An analysis of the impact of the reform of the structural funds in Ireland."* ICS dissertation, Groningen
61. René Veenstra (1999), *"Leerlingen, Klassen, Scholen: Prestaties en vorderingen van leerlingen in het voortgezet onderwijs."* Amsterdam, Thela Thesis
62. Marjolein Achterkamp (1999), *"Influence strategies in collective decision making: A comparison of two models."* ICS dissertation, Groningen
63. Peter Mühlau (2000), *"The governance of the employment relation: Arelational signaling perspective."* ICS dissertation, Groningen
64. Agnes Akkerman (2000), *"Verdeelde vakbeweging en stakingen: Concurrentie om leden."* ICS dissertation, Groningen
65. Sandra van Thiel (2000), *"Quangocratization: Trends, causes and consequences."* ICS dissertation, Utrecht
66. Sylvia Peacock-Korupp (2000), *"Mothers and the process of social stratification."* ICS dissertation, Utrecht
67. Rudi Turksema (2000), *"Supply of day care."* ICS dissertation, Utrecht
68. Bernard Nijstad (2000), *"How the group affects the mind: Effects of communication in idea generating groups."* ICS dissertation, Utrecht
69. Inge de Wolf (2000), *"Opleidingsspecialisatie en arbeidsmarktsucces van sociale wetenschappers."* ICS dissertation, Utrecht
70. Jan Kratzer (2001), *"Communication and performance: An empirical study in innovation teams."* ICS dissertation, Groningen
71. Madelon Kroneman (2001), *"Healthcare systems and hospital bed use."* ICS/NIVEL-dissertation, Utrecht
72. Herman van de Werfhorst (2001), *"Field of study and social inequality: Four types of educational resources in the process of stratification in the Netherlands."* ICS dissertation, Nijmegen
73. Tamás Bartus (2001), *"Social capital and earnings inequalities: The role of informal job search in Hungary."* ICS dissertation Groningen
74. Hester Moerbeek (2001), *"Friends and foes in the occupational career: The influence of sweet and sour social capital on the labour market."* ICS dissertation, Nijmegen
75. Marcel van Assen (2001), *"Essays on actor perspectives in exchange networks and social dilemmas."* ICS dissertation, Groningen
76. Inge Sieben (2001), *"Sibling similarities and social stratification: The impact of family background across countries and cohorts."* ICS dissertation, Nijmegen
77. Alinda van Bruggen (2001), *"Individual production of social well-being: An exploratory study."* ICS dissertation, Groningen
78. Marcel Coenders (2001), *"Nationalistic attitudes and ethnic exclusionism in a comparative perspective: An empirical study of attitudes toward the country and ethnic immigrants in 22 countries."* ICS dissertation, Nijmegen

79. Marcel Lubbers (2001), *"Exclusionistic electorates: Extreme right-wing voting in Western Europe."* ICS dissertation, Nijmegen
80. Uwe Matzat (2001), *"Social networks and cooperation in electronic communities: A theoretical-empirical analysis of academic communication and internet discussion groups."* ICS dissertation, Groningen
81. Jacques Janssen (2002), *"Do opposites attract divorce? Dimensions of mixed marriage and the risk of divorce in the Netherlands."* ICS dissertation, Nijmegen
82. Miranda Jansen (2002), *"Waardenoriëntaties en partnerrelaties: Een panelstudie naar wederzijdse invloeden."* ICS dissertation, Utrecht
83. Anne-Rigt Poortman (2002), *"Socioeconomic causes and consequences of divorce."* ICS dissertation, Utrecht
84. Alexander Gattig (2002), *"Intertemporal decision making."* ICS dissertation, Groningen
85. Gerrit Rooks (2000), *"Contract en conflict: Strategisch management van inkooptransacties."* ICS dissertation, Utrecht
86. Károly Takács (2002), *"Social networks and intergroup conflict."* ICS dissertation, Groningen
87. Thomas Gautschi (2002), *"Trust and exchange, effects of temporal embeddedness and network embeddedness on providing and dividing a surplus."* ICS dissertation, Utrecht
88. Hilde Bras (2002), *"Zeeuwse meiden: Dienen in de levensloop van vrouwen, ca. 1850 – 1950."* Aksant Academic Publishers, Amsterdam
89. Merijn Rengers (2002), *"Economic lives of artists: Studies into careers and the labour market in the cultural sector."* ICS dissertation, Utrecht
90. Annelies Kassenberg (2002), *"Wat scholieren bindt: Sociale gemeenschap in scholen."* ICS dissertation, Groningen
91. Marc Verboord (2003), *"Moet de meester dalen of de leerling klimmen? De invloed van literatuuronderwijs en ouders op het lezen van boeken tussen 1975 en 2000."* ICS dissertation, Utrecht
92. Marcel van Egmond (2003), *"Rain falls on all of us (but some manage to get more wet than others): Political context and electoral participation."* ICS dissertation, Nijmegen
93. Justine Horgan (2003), *"High-performance human resource management in Ireland and the Netherlands: Adoption and effectiveness."* ICS dissertation, Groningen
94. Corine Hoeben (2003), *"LETS' be a community: Community in local exchange trading systems."* ICS dissertation, Groningen
95. Christian Steglich (2003), *"The framing of decision situations: Automatic goal selection and rational goal pursuit."* ICS dissertation, Groningen
96. Johan van Wilsem (2003), *"Crime and context: The impact of individual, neighborhood, city and country characteristics on victimization."* ICS dissertation, Nijmegen
97. Christiaan Monden (2003), *"Education, inequality and health: The impact of partners and life course."* ICS dissertation, Nijmegen
98. Evelyn Hello (2003), *"Educational attainment and ethnic attitudes: How to explain their relationship."* ICS dissertation, Nijmegen
99. Marnix Croes en Peter Tammes (2004). *Gif laten wij niet voortbestaan: Een onderzoek naar de overlevingskansen van Joden in de Nederlandse gemeenten, 1940-1945*. Aksant Academic Publishers, Amsterdam.
100. Ineke Nagel (2004), *"Cultuurdeelname in de levensloop."* ICS dissertation, Utrecht
101. Marieke van der Wal (2004), *"Competencies to participate in life: Measurement and the impact of school."* ICS dissertation, Groningen
102. Vivian Meertens (2004), *"Depressive symptoms in the general population: A multifactorial social approach."* ICS dissertation, Nijmegen
103. Hanneke Schuurmans (2004), *"Promoting well-being in frail elderly people: Theory and intervention."* ICS dissertation, Groningen
104. Javier Arregui Moreno (2004), *"Negotiation in legislative decision-making in the European Union."* ICS dissertation, Groningen
105. Tamar Fischer (2004), *"Parental divorce, conflict and resources: The effects on children's behaviour problems, socioeconomic attainment, and transitions in the demographic career."* ICS dissertation, Nijmegen
106. René Bekkers (2004), *"Giving and volunteering in the Netherlands: Sociological and psychological perspectives."* ICS dissertation, Utrecht

107. Renée van der Hulst (2004), *"Gender differences in workplace authority: An empirical study on social networks."* ICS dissertation, Groningen
108. Rita Smaniotto (2004), *"'You scratch my back and I scratch yours' versus 'love thy neighbour': Two proximate mechanisms of reciprocal altruism."* ICS dissertation, Groningen
109. Maurice Gesthuizen (2004), *"The life course of the low-educated in the Netherlands: Social and economic risks."* ICS dissertation, Nijmegen
110. Carlijne Philips (2005), *"Vakantiegemeenschappen: Kwalitatief en kwantitatief onderzoek naar gelegenheid en refreshergemeenschap tijdens de vakantie."* ICS dissertation, Groningen
111. Esther de Ruijter (2005), *"Household outsourcing."* ICS dissertation, Utrecht
112. Frank van Tubergen (2005), *"The integration of immigrants in cross-national perspective: Origin, destination, and community effects."* ICS dissertation, Utrecht
113. Ferry Koster (2005), *"For the time being: Accounting for inconclusive findings concerning the effects of temporary employment relationships on solidary behavior of employees."* ICS dissertation, Groningen
114. Carolien Klein Haarhuis (2005), *"Promoting anti-corruption reforms: Evaluating the implementation of a World Bank anti-corruption program in seven African countries (1999-2001)."* ICS dissertation, Utrecht
115. Martin van der Gaag (2005), *"Measurement of individual social capital."* ICS dissertation, Groningen
116. Johan Hansen (2005), *"Shaping careers of men and women in organizational contexts."* ICS dissertation, Utrecht
117. Davide Barrera (2005), *"Trust in embedded settings."* ICS dissertation, Utrecht
118. Mattijs Lambooij (2005), *"Promoting cooperation: Studies into the effects of long-term and short-term rewards on cooperation of employees."* ICS dissertation, Utrecht
119. Lotte Vermeij (2006), *"What's cooking? Cultural boundaries among Dutch teenagers of different ethnic origins in the context of school."* ICS dissertation, Utrecht
120. Mathilde Strating (2006), *"Facing the challenge of rheumatoid arthritis: A 13-year prospective study among patients and cross-sectional study among their partners."* ICS dissertation, Groningen
121. Jannes de Vries (2006), *"Measurement error in family background variables: The bias in the intergenerational transmission of status, cultural consumption, party preference, and religiosity."* ICS dissertation, Nijmegen
122. Stefan Thau (2006), *"Workplace deviance: Four studies on employee motives and self-regulation."* ICS dissertation, Groningen
123. Mirjam Plantinga (2006), *"Employee motivation and employee performance in child care: The effects of the introduction of market forces on employees in the Dutch child-care sector."* ICS dissertation, Groningen
124. Helga de Valk (2006), *"Pathways into adulthood: A comparative study on family life transitions among migrant and Dutch youth."* ICS dissertation, Utrecht
125. Henrike Elzen (2006), *"Self-Management for chronically ill older people."* ICS dissertation, Groningen
126. Ayse Güveli (2007), *"New social classes within the service class in the Netherlands and Britain: Adjusting the EGP class schema for the technocrats and the social and cultural specialists."* ICS dissertation, Nijmegen
127. Willem-Jan Verhoeven (2007), *"Income attainment in post-communist societies."* ICS dissertation, Utrecht
128. Marieke Voorpostel (2007), *"Sibling support: The exchange of help among brothers and sisters in the Netherlands."* ICS dissertation, Utrecht
129. Jacob Dijkstra (2007), *"The effects of externalities on partner choice and payoffs in exchange networks."* ICS dissertation, Groningen
130. Patricia van Echtelt (2007), *"Time-greedy employment relationships: Four studies on the time claims of post-Fordist work."* ICS dissertation, Groningen
131. Sonja Vogt (2007), *"Heterogeneity in social dilemmas: The case of social support."* ICS dissertation, Utrecht
132. Michael Schweinberger (2007), *"Statistical methods for studying the evolution of networks and behavior."* ICS dissertation, Groningen
133. István Henrik Back (2007), *"Commitment and evolution: Connecting emotion and reason in long-term relationships."* ICS dissertation, Groningen
134. Ruben van Gaalen (2007), *"Solidarity and ambivalence in parent-child relationships."* ICS dissertation, Utrecht
135. Jan Reitsma (2007), *"Religiosity and solidarity: Dsimensions and relationships disentangled and tested."* ICS dissertation, Nijmegen

136. Jan Kornelis Dijkstra (2007), *"Status and affection among (pre)adolescents and their relation with antisocial and prosocial behavior."* ICS dissertation, Groningen
137. Wouter van Gils (2007), *"Full-time working couples in the Netherlands: Causes and consequences."* ICS dissertation, Nijmegen
138. Djamila Schans (2007), *"Ethnic diversity in intergenerational solidarity."* ICS dissertation, Utrecht
139. Ruud van der Meulen (2007), *"Brug over woelig water: Lidmaatschap van sportverenigingen, vriendschappen, kennissenkringen en veralgemeend vertrouwen."* ICS dissertation, Nijmegen
140. Andrea Knecht (2008), *"Friendship selection and friends' influence: Dynamics of networks and actor attributes in early adolescence."* ICS dissertation, Utrecht
141. Ingrid Doorten (2008), *"The division of unpaid work in the household: A stubborn pattern?"* ICS dissertation, Utrecht
142. Stijn Ruiter (2008), *"Association in context and association as context: Causes and consequences of voluntary association involvement."* ICS dissertation, Nijmegen
143. Janneke Joly (2008), *"People on our minds: When humanized contexts activate social norms."* ICS dissertation, Groningen
144. Margreet Frieling (2008), *"'Joint production' als motor voor actief burgerschap in de buurt."* ICS dissertation, Groningen
145. Ellen Verbakel (2008), *"The partner as resource or restriction? Labour market careers of husbands and wives and the consequences for inequality between couples."* ICS dissertation, Nijmegen
146. Gijs van Houten (2008), *"Beleidsuitvoering in gelaagde stelsels: De doorwerking van aanbevelingen van de Stichting van de Arbeid in het CAO-overleg."* ICS dissertation, Utrecht
147. Eva Jaspers (2008), *"Intolerance over time: Macro- and micro-level questions on attitudes towards euthanasia, homosexuality and ethnic minorities."* ICS dissertation, Nijmegen
148. Gijs Weijters (2008), *"Youth delinquency in Dutch cities and schools: A multilevel approach."* ICS dissertation, Nijmegen
149. Jessica Nooij (2009), *"The self in social rejection."* ICS dissertation, Groningen
150. Gerald Mollenhorst (2009), *"Networks in contexts: How meeting opportunities affect personal relationships."* ICS dissertation, Utrecht
151. Tom van der Meer (2009), *"States of freely associating citizens: Comparative studies into the impact of state institutions on social, civic and political participation."* ICS dissertation, Nijmegen
152. Manuela Vieth (2009), *"Commitments and reciprocity in trust situations: Experimental studies on obligation, indignation, and self-consistency."* ICS dissertation, Utrecht
153. Rense Corten (2009), *"Co-evolution of social networks and behavior in social dilemmas: Theoretical and empirical perspectives."* ICS dissertation, Utrecht
154. Arieke Rijken (2009), *"Happy families, high fertility? Childbearing choices in the context of family and partner relationships."* ICS dissertation, Utrecht
155. Jochem Tolsma (2009), *"Ethnic hostility among ethnic majority and minority groups in the Netherlands: An investigation into the impact of social mobility experiences, the local living environment and educational attainment on ethnic hostility."* ICS dissertation, Nijmegen
156. Freek Bucx (2009), *"Linked lives: Young adults' life course and relations with parents."* ICS dissertation, Utrecht
157. Philip Wotschack (2009), *"Household governance and time allocation: Four studies on the combination of work and care."* ICS dissertation, Groningen
158. Nienke Moor (2009), *"Explaining worldwide religious diversity: The relationship between subsistence technologies and ideas about the unknown in pre-industrial and (post-)industrial societies."* ICS dissertation, Nijmegen
159. Lieke ten Brummelhuis (2009), *"Family matters at work: Depleting and enriching effects of employees' family lives on work outcomes."* ICS dissertation, Utrecht
160. Renske Keizer (2010), *"Remaining childless: Causes and consequences from a life-course perspective."* ICS dissertation, Utrecht
161. Miranda Sentse (2010), *"Bridging contexts: The interplay between family, child, and peers in explaining problem behavior in early adolescence."* ICS dissertation, Groningen

162. Nicole Tieben (2010), *"Transitions, tracks and transformations: Social inequality in transitions into, through and out of secondary education in the Netherlands for cohorts born between 1914 and 1985."* ICS dissertation, Nijmegen
163. Birgit Pauksztat (2010), *"Speaking up in organizations: Four studies on employee voice."* ICS dissertation, Groningen
164. Richard Zijdeman (2010), *"Status attainment in the Netherlands, 1811-1941: Spatial and temporal variation before and during industrialization."* ICS dissertation, Utrecht
165. Rianne Kloosterman (2010), *"Social background and children's educational careers: The primary and secondary effects of social background over transitions and over time in the Netherlands."* ICS dissertation, Nijmegen
166. Olav Aarts (2010), *"Religious diversity and religious involvement: A study of religious markets in Western societies at the end of the twentieth century."* ICS dissertation, Nijmegen
167. Stephanie Wiesmann (2010), *"24/7 Negotiation in couples' transition to parenthood."* ICS dissertation, Utrecht
168. Borja Martinovic (2010), *"Interethnic contacts: A dynamic analysis of interaction between immigrants and natives in Western countries."* ICS dissertation, Utrecht
169. Anne Roeters (2010), *"Family life under pressure? Parents' paid work and the quantity and quality of parent-child and family time."* ICS dissertation, Utrecht
170. Jelle Sijtsema (2010), *"Adolescent aggressive behavior: Status and stimulation goals in relation to the peer context."* ICS dissertation, Groningen
171. Kees Keizer (2010), *"The spreading of disorder."* ICS dissertation, Groningen
172. Michael Mäs (2010), *"The diversity puzzle. explaining clustering and polarization of opinions."* ICS dissertation, Groningen
173. Marie-Louise Damen (2010), *"Cultuurdeelname en CKV: Studies naar effecten van kunsteducatie op de cultuurdeelname van leerlingen tijdens en na het voortgezet onderwijs."* ICS dissertation, Utrecht
174. Marieke van de Rakt (2011), *"Two generations of crime: The intergenerational transmission of convictions over the life course."* ICS dissertation, Nijmegen
175. Willem Huijnk (2011), *"Family life and ethnic attitudes: The role of the family for attitudes towards intermarriage and acculturation among minority and majority groups."* ICS dissertation, Utrecht
176. Tim Huijts (2011), *"Social ties and health in Europe: Individual associations, cross-national variations, and contextual explanations."* ICS dissertation, Nijmegen
177. Wouter Steenbeek (2011), *"Social and physical disorder: How community, business presence and entrepreneurs influence disorder in Dutch neighborhoods."* ICS dissertation, Utrecht
178. Miranda Vervoort (2011), *"Living together apart? Ethnic concentration in the neighborhood and ethnic minorities' social contacts and language practices."* ICS dissertation, Utrecht
179. Agnieszka Kanas (2011), *"The economic performance of immigrants: The role of human and social capital."* ICS dissertation, Utrecht
180. Lea Ellwardt (2011), *"Gossip in organizations: A social network study."* ICS dissertation, Groningen
181. Annemarije Oosterwaal (2011), *"The gap between decision and implementation: Decision making, delegation and compliance in governmental and organizational settings."* ICS dissertation, Utrecht
182. Natascha Notten (2011), *"Parents and the media: Causes and consequences of parental media socialization."* ICS dissertation, Nijmegen
183. Tobias Stark (2011), *"Integration in schools: A process perspective on students' interethnic attitudes and interpersonal relationships."* ICS dissertation, Groningen
184. Giedo Jansen (2011), *"Social cleavages and political choices: Large-scale comparisons of social class, religion and voting behavior in Western democracies."* ICS dissertation, Nijmegen
185. Ruud van der Horst (2011), *"Network effects on treatment results in a closed forensic psychiatric setting."* ICS dissertation, Groningen
186. Mark Levels (2011), *"Abortion laws in European Countries between 1960 and 2010: Legislative developments and their consequences for women's reproductive decision making."* ICS dissertation, Nijmegen
187. Marieke van Londen (2012), *"Exclusion of ethnic minorities in the Netherlands: The effects of individual and situational characteristics on opposition to ethnic policy and ethnically mixed neighbourhoods."* ICS dissertation, Nijmegen

188. Sigrid Mohnen (2012), *"Neighborhood context and health: How neighborhood social capital affects individual health."* ICS dissertation, Utrecht
189. Asya Zhelyazkova (2012), *"Compliance under controversy: Analysis of the transposition of European directives and their provisions."* ICS dissertation, Utrecht
190. Valeska Korff (2012), *"Between cause and control: Management in a humanitarian organization."* ICS dissertation, Groningen
191. Maike Gieling (2012), *"Dealing with diversity: Adolescents' support for civil liberties and immigrant rights."* ICS dissertation, Utrecht
192. Katya Ivanova (2012), *"From parents to partners: The impact of family on romantic relationships in adolescence and emerging adulthood."* ICS dissertation, Groningen
193. Jelmer Schalk (2012), *"The performance of public corporate actors: Essays on effects of institutional and network embeddedness in supranational, national, and local collaborative contexts."* ICS dissertation, Utrecht
194. Alona Labun (2012), *"Social networks and informal power in organizations."* ICS dissertation, Groningen
195. Michal Bojanowski (2012), *"Essays on social network formation in heterogeneous populations: Models, methods, and empirical analyses."* ICS dissertation, Utrecht
196. Anca Minescu (2012), *"Relative group position and intergroup attitudes in Russia."* ICS dissertation, Utrecht
197. Marieke van Schellen (2012), *"Marriage and crime over the life course: The criminal careers of convicts and their spouses."* ICS dissertation, Utrecht
198. Mieke Maliepaard (2012), *"Religious trends and social integration: Muslim minorities in the Netherlands."* ICS dissertation, Utrecht
199. Fransje Smits (2012), *"Turks and Moroccans in the Low Countries around the year 2000: Determinants of religiosity, trend in religiosity and determinants of the trend."* ICS dissertation, Nijmegen
200. Roderick Sluiter (2012), *"The diffusion of morality policies among Western European countries between 1960 and 2010: A comparison of temporal and spatial diffusion patterns of six morality and eleven non-morality policies."* ICS dissertation, Nijmegen
201. Nicoletta Balbo (2012), *"Family, friends and fertility."* ICS dissertation, Groningen
202. Anke Munniksma (2013), *"Crossing ethnic boundaries: Parental resistance to and consequences of adolescents' cross-ethnic peer relations."* ICS dissertation, Groningen
203. Anja-Kristin Abendroth (2013), *"Working women in Europe: How the country, workplace, and family context matter."* ICS dissertation, Utrecht
204. Katia Begall (2013), *"Occupational hazard? The relationship between working conditions and fertility."* ICS dissertation, Groningen
205. Hidde Bekhuis (2013), *"The popularity of domestic cultural products: Cross-national differences and the relation to globalization."* ICS dissertation, Utrecht
206. Lieselotte Blommaert (2013), *"Are Joris and Renske more employable than Rashid and Samira? A study on the prevalence and sources of ethnic discrimination in recruitment in the Netherlands using experimental and survey data."* ICS dissertation, Utrecht
207. Wiebke Schulz (2013), *"Careers of men and women in the 19th and 20th centuries."* ICS dissertation, Utrecht
208. Ozan Aksoy (2013), *"Essays on social preferences and beliefs in non-embedded social dilemmas."* ICS dissertation, Utrecht
209. Dominik Morbitzer (2013), *"Limited farsightedness in network formation."* ICS dissertation, Utrecht
210. Thomas de Vroome (2013), *"Earning your place: The relation between immigrants' economic and psychological integration in the Netherlands."* ICS dissertation, Utrecht
211. Marloes de Lange (2013), *"Causes and consequences of employment flexibility among young people: Recent developments in the Netherlands and Europe."* ICS dissertation, Nijmegen
212. Roza Meuleman (2014), *"Consuming the Nation: Domestic cultural consumption: Its stratification and relation with nationalist attitudes."* ICS dissertation, Utrecht
213. Esther Havekes (2014), *"Putting interethnic attitudes in context: The relationship between neighbourhood characteristics, interethnic attitudes and residential behaviour."* ICS dissertation, Utrecht
214. Zoltán Lippényi (2014), *"Transitions toward an open society? Intergenerational occupational mobility in Hungary in the 19th and 20th centuries."* ICS dissertation, Utrecht

215. Anouk Smeekes (2014), *"The presence of the past: Historical rooting of national identity and current group dynamics."* ICS dissertation, Utrecht
216. Michael Savelkoul (2014), *"Ethnic diversity and social capital: Testing underlying explanations derived from conflict and contact theories in Europe and the United States."* ICS dissertation, Nijmegen
217. Martijn Hogerbrugge (2014), *"Misfortune and family: How negative events, family ties, and lives are linked."* ICS dissertation, Utrecht
218. Gina-Felicia Potarca (2014), *"Modern love: Comparative insights in online dating preferences and assortative mating."* ICS dissertation, Groningen
219. Mariska van der Horst (2014), *"Gender, aspirations, and achievements: Relating work and family aspirations to occupational outcomes."* ICS dissertation, Utrecht
220. Gijs Huitsing (2014), *"A social network perspective on bullying."* ICS dissertation, Groningen
221. Thomas Kowalewski (2015), *"Personal growth in organizational contexts."* ICS dissertation, Groningen
222. Manuel Muñoz-Herrera (2015), *"The impact of individual differences on network relations: Social exclusion and inequality in productive exchange and coordination games."* ICS dissertation, Groningen
223. Tim Immerzeel (2015), *"Voting for a change: The democratic lure of populist radical right parties in voting behavior."* ICS dissertation, Utrecht
224. Fernando Nieto Morales (2015), *"The control imperative: Studies on reorganization in the public and private sectors."* ICS dissertation, Groningen
225. Jellie Sierksma (2015), *"Bounded helping: How morality and intergroup relations shape children's reasoning about helping."* ICS dissertation, Utrecht
226. Tinka Veldhuis (2015), *"Captivated by fear: An evaluation of terrorism detention policy."* ICS dissertation, Groningen
227. Miranda Visser (2015), *"Loyality in humanity: Turnover among expatriate humanitarian aid workers."* ICS dissertation, Groningen
228. Sarah Westphal (2015), *"Are the kids alright? Essays on postdivorce residence arrangements and children's well-being."* ICS dissertation, Utrecht
229. Britta Rüschoff (2015), *"Peers in careers: Peer relationships in the transition from school to work."* ICS dissertation, Groningen
230. Nynke van Miltenburg (2015), *"Cooperation under peer sanctioning institutions: Collective decisions, noise, and endogenous implementation."* ICS dissertation, Utrecht
231. Antonie Knigge (2015), *"Sources of sibling similarity: Status attainment in the Netherlands during modernization."* ICS dissertation, Utrecht
232. Sanne Smith (2015), *"Ethnic segregation in friendship networks: Studies of its determinants in English, German, Dutch, and Swedish school classes."* ICS dissertation, Utrecht
233. Patrick Präg (2015), *"Social stratification and health: Four essays on social determinants of health and wellbeing."* ICS dissertation, Groningen
234. Wike Been (2015), *"European top managers' support for work-life arrangements."* ICS dissertation, Utrecht
235. André Grow (2016), *"Status differentiation: New insights from agent-based modeling and social network analysis."* ICS dissertation, Groningen
236. Jesper Rözer (2016), *"Family and personal networks: How a partner and children affect social relationships."* ICS dissertation, Utrecht
237. Kim Pattiselanno (2016), *"At your own risk: The importance of group dynamics and peer processes in adolescent peer groups for adolescents' involvement in risk behaviors."* ICS- dissertation, Groningen
238. Vincenz Frey (2016), *"Network formation and trust."* ICS dissertation, Utrecht
239. Rozemarijn van der Ploeg (2016), *"Be a buddy, not a bully? Four studies on social and emotional processes related to bullying, defending, and victimization."* ICS dissertation, Groningen
240. Tali Spiegel (2016), *"Identity, career trajectories and wellbeing: A closer look at individuals with degenerative eye conditions."* ICS- dissertation, Groningen
241. Felix Christian Tropf (2016), *"Social science genetics and fertility."* ICS dissertation, Groningen
242. Sara Geven (2016), *"Adolescent problem behavior in school: The role of peer networks."* ICS dissertation, Utrecht
243. Josja Rokven (2016), *"The victimization-offending relationship from a longitudinal perspective."* ICS dissertation, Nijmegen

244. Maja Djundeva (2016), *"Healthy ageing in context: Family welfare state and the life course."* ICS dissertation, Groningen
245. Mark Visser (2017), *"Inequality between older workers and older couples in the Netherlands: A dynamic life course perspective on educational and social class differences in the late career."* ICS dissertation, Nijmegen
246. Beau Oldenburg (2017), *"Bullying in schools: The role of teachers and classmates."* ICS dissertation, Groningen
247. Tatang Muttaqin (2017), *"The education divide in Indonesia: Four essays on determinants of unequal access to and quality of education."* ICS dissertation, Groningen
248. Margriet van Hek (2017), *"Gender inequality in educational attainment and reading performance: A contextual approach."* ICS dissertation, Nijmegen
249. Melissa Verhoef-van Dorp (2017), *"Work schedules, childcare and well-being: Essays on the associations between modern-day job characteristics, childcare arrangements and the well-being of parents and children."* ICS dissertation, Utrecht
250. Timo Septer (2017), *"Goal priorities, cognition and conflict: Analyses of cognitive maps concerning organizational change."* ICS dissertation, Groningen
251. Bas Hofstra (2017), *"Online social networks: Essays on membership, privacy, and structure."* ICS dissertation, Utrecht
252. Yassine Khoudja (2018), *"Women's labor market participation across ethnic groups: The role of household conditions, gender role attitudes, and religiosity in different national contexts."* ICS dissertation, Utrecht
253. Joran Laméris (2018), *"Living together in diversity: Whether, why and where ethnic diversity affects social cohesion."* ICS dissertation, Nijmegen
254. Maaike van der Vleuten (2018), *"Gendered Choices: Fields of study of adolescents in the Netherlands."* ICS dissertation, Utrecht
255. Mala Sondang Silitonga (2018), *"Corruption in Indonesia: The impact of institutional change, norms, and networks."* ICS dissertation, Groningen
256. Manja Coopmans (2018), *"Rituals of the past in the context of the present: The role of Remembrance Day and Liberation Day in Dutch society."* ICS dissertation, Utrecht
257. Paul Hindriks (2018), *"The struggle for power: Attitudes towards the political participation of ethnic minorities."* ICS dissertation, Utrecht
258. Nynke Niezink (2018), *"Modeling the dynamics of networks and continuous behavior."* ICS dissertation, Groningen
259. Simon de Bruijn (2018), *"Reaching agreement after divorce and separation: Essays on the effectiveness of parenting plans and divorce mediation."* ICS dissertation, Utrecht
260. Susanne van 't Hoff-de Goede (2018), *"While you were locked up: An empirical study on the characteristics, social surroundings and wellbeing of partners of prisoners in The Netherlands."* ICS dissertation, Utrecht
261. Loes van Rijsewijk (2018), *"Antecedents and consequences of helping among adolescents."* ICS dissertation, Groningen
262. Mariola Gremmen (2018), *"Social network processes and academic functioning: The role of peers in students' school well-being, academic engagement, and academic achievement."* ICS dissertation, Groningen
263. Jeanette Renema (2018), *"Immigrants' support for welfare spending: The causes and consequences of welfare usage and welfare knowledgeability."* ICS dissertation, Nijmegen
264. Suwatin Miharti (2018), *"Community health centers in Indonesia in the era of decentralization: The impact of structure, staff composition and management on health outcomes."* ICS dissertation, Groningen
265. Chaïm la Roi (2019), *"Stigma and stress: Studies on attitudes towards sexual minority orientations and the association between sexual orientation and mental health."* ICS dissertation, Groningen
266. Jelle Lössbroek (2019), *"Turning grey into gold: Employer-employee interplay in an ageing workforce."* ICS dissertation, Utrecht
267. Nikki van Gerwen (2019), *"Employee cooperation through training: A multi-method approach."* ICS dissertation, Utrecht
268. Paula Thijs (2019), *"Trends in cultural conservatism: The role of educational expansion, secularisation, and changing national contexts."* ICS dissertation, Nijmegen
269. Renske Verweij (2019), *"Understanding childlessness: Unravelling the link with genes and the socio-environment."* ICS dissertation, Groningen

270. Niels Blom (2019), *"Partner relationship quality under pressing work conditions: Longitudinal and cross-national investigation."* ICS dissertation, Nijmegen
271. Müge Simsek (2019), *"The dynamics of religion among native and immigrant youth in Western Europe."* ICS dissertation, Utrecht
272. Leonie van Breeschoten (2019), *"Combining a career and childcare: The use and usefulness of work-family policies in European organizations."* ICS dissertation, Utrecht
273. Roos van der Zwan (2019), "*The political representation of ethnic minorities and their vote choice."* ICS dissertation, Nijmegen
274. Ashwin Rambaran (2019), "*The classroom as context for bullying: A social network approach*." ICS dissertation, Groningen
275. Dieko Bakker (2019), *"Cooperation and social control: Effects of preferences, institutions, and social structure."* ICS dissertation, Groningen
276. Femke van der Werf (2019), "*Shadow of a rainbow? National and ethnic belonging in Mauritius*." ICS dissertation, Utrecht
277. Robert Krause (2019), *"Multiple imputation for missing network data."* ICS dissertation, Groningen
278. Take Sipma (2020), *"Economic insecurity and populist radical right voting."* ICS dissertation, Nijmegen
279. Mathijs Kros (2020), "*The nature of negative contact: Studies on interethnic relations in Western societies."* ICS dissertation, Utrecht
280. Lonneke van den Berg (2020), *"Time to leave: Individual and contextual explanations for the timing of leaving home."* ICS dissertation, Amsterdam
281. Marianne Hooijsma (2020), "*Clashrooms: Interethnic peer relationships in schools."* ICS dissertation, Groningen
282. Marina Tulin (2020*), "Blind spots in social resource theory: Essays on the creation, maintenance and returns of social capital."* ICS dissertation, Amsterdam
283. Tessa Kaufman (2020), "*Toward tailored interventions: Explaining, assessing, and preventing persistent victimization of bullying."* ICS dissertation, Groningen
284. Lex Thijssen (2020), *"Racial and ethnic discrimination in western labor markets: Empirical evidence from field experiments."* ICS dissertation, Utrecht
285. Lukas Norbutas (2020), *"Trust on the dark web: An analysis of illegal online drug markets."* ICS dissertation, Utrecht
286. Tomáš Diviák (2020), *"Criminal networks: Actors, mechanisms, and structures."* ICS dissertation, Groningen
287. Tery Setiawan (2020), "*Support for interreligious conflict in Indonesia*." ICS dissertation, Nijmegen
288. Vera de Bel (2020), "*The ripple effect in family networks: Relational structures and well-being in divorced and non-divorced families*." ICS dissertation, Groningen
289. Diego Palacios (2020), *"How context and the perception of peers' behaviors shape relationships in adolescence: A multiplex social network perspective."* ICS dissertation, Groningen
290. Saskia Glas (2020), *"Where are the Muslim Feminists? Religiosity and support for gender equality in the Arab region*." ICS dissertation, Nijmegen
291. Tomas Turner-Zwinkels (2020), *"A new macro-micro approach to the study of political careers: Theoretical, Methodological and empirical challenges and solutions."* ICS dissertation, Groningen
292. Lotte Scheeren (2020), "*Not on the same track? Tracking age and gender inequality in education."* ICS dissertation, Amsterdam
293. Joris Broere (2020), "*Essays on how social network structure affects asymmetric coordination and trust*." ICS dissertation, Utrecht
294. Marcus Kristiansen (2021*), "Contact with benefits: How social networks affect benefit receipt dynamics in the Netherlands."* ICS dissertation, Utrecht
295. Judith Kas (2021), "*Trust and reputation in the peer-to-peer platform economy*." ICS dissertation, Utrecht
296. Andrea Forster (2021), "*Navigating educational institutions: Mechanisms of educational inequality and social mobility in different educational systems."* ICS dissertation, Amsterdam
297. Jannes ten Berge (2021), "*Technological change and work: The relation between technology implementation within organizations and changes in workers' employment."* ICS dissertation, Utrecht
298. Jolien Geerlings (2021), "*Teaching in culturally diverse classrooms: The importance of dyadic relations between teachers and children*." ICS dissertation, Utrecht

299. Kirsten van Houdt (2021), "*Stepfamilies in adulthood: Solidarity between parents and adult children*." ICS dissertation, Amsterdam
300. Suzanne de Leeuw (2021), "*The intergenerational transmission of educational attainment after divorce and remarriage."* ICS dissertation, Amsterdam
301. Fleur Goedkoop (2021), "*Involvement in bottom-up energy transitions: The role of local and contextual embeddedness*." ICS dissertation, Groningen
302. Eva Vriens (2021), "*Mutualism in the 21st century: The why, when, and how behind successful risk-sharing institutions*." ICS dissertation, Utrecht
303. Ardita Muja (2021), "*From school to work: The role of the vocational specificity of education in young people's labor market integration*." ICS dissertation, Nijmegen
304. Siyang Kong (2021), "*Women and work in contemporary China: The effect of market transition on women's employment, earnings, and status attainment*." ICS dissertation, Utrecht
305. Marijn Keijzer (2022), *"Opinion dynamics in online social media."* ICS dissertation, Groningen
306. Sander Kunst (2022), *"The educational divide in openness towards globalisation in Western Europe."* ICS dissertation, Amsterdam
307. Nella Geurts (2022), "*Puzzling pathways: The integration paradox among migrants in Western Europe."* ICS dissertation, Nijmegen
308. Dragana Stojmenovska (2022), *"Men's place: The incomplete integration of women in workplace authority."* ICS dissertation, Amsterdam
309. Bram Hogendoorn (2022), *"Divorce and inequality: Stratification in the risk and consequences of union dissolution."* ICS dissertation, Amsterdam
310. Tom Nijs (2022), *"This place is ours: Collective psychological ownership and its social consequences."* ICS dissertation, Utrecht
311. Nora Storz (2022), *"'This land is ours—but is it also theirs?' Collective ownership beliefs and reconciliation in territorial conflict regions."* ICS dissertation, Utrecht
312. Tara Koster (2022), *"Parenting and fairness in diverse families."* ICS dissertation, Utrecht
313. Danelien van Aalst (2022), "*Elements Contributing to Teachers' Role in Bullying."* ICS dissertation, Groningen
314. Wybren Nooitgedagt (2022), *"Who owns the country? Collective psychological ownership and intergroup relations in settler societies."* ICS dissertation, Utrecht
315. Marija Dangubić (2022), *"Rejecting Muslim minority practices: Principles and prejudices."* ICS dissertation, Utrecht
316. Sara Cvetkovska (2022), *"Lines in the shifting sand: The implications of being tolerated."* ICS dissertation, Utrecht
317. Maaike Hornstra (2022), "*Going beyond the dyad: Adult intergenerational closeness after divorce and remarriage."* ICS dissertation, Amsterdam
318. Wouter Kiekens (2022), *"Sexual and gender minority youth's mental health and substance use: Disparities, mechanisms, and protective factors."* ICS dissertation, Groningen
319. Carlijn Bussemakers (2022), *"Adversity and educational inequality: The interplay between adverse experiences and parental resources for children's educational attainment."* ICS dissertation, Nijmegen
320. Evi Velthuis (2022), "*'To tolerate or not to tolerate?' Reasons for tolerance of minority group practices among majority members in the Netherlands and Germany."* ICS dissertation, Utrecht
321. Hendrik Nunner (2023), *"Co-evolution of social networks and infectious diseases."* ICS dissertation, Utrecht
322. Carly van Mensvoort (2023), *"Inspiring leaders: An empirical study of female supervisors at work."* ICS dissertation, Nijmegen
323. Anne van der Put (2023), *"Healthy at work: The role of the work environment in worksite health promotion."* ICS dissertation, Utrecht
324. Rowan ten Kate (2023), "*Understanding loneliness among older migrants*." ICS dissertation, Groningen
325. Sanne Kellij (2023), "*I see, I see what you don't see: Neural and behavioral social-cognitive processes underlying (persistent) victimization. "* ICS dissertation, Groningen
326. Inge Hendriks (2023), "*Changes and contrasts in attitudes towards ethnic minorities*." ICS dissertation, Nijmegen
327. Eleonora Marucci (2023), "*Antecedents and consequences of teacher attunement in primary and secondary school classrooms."* ICS dissertation, Groningen

328. Kasper Otten (2023), "*Cooperation in changing groups: How newcomers and norms shape public good provision in the lab, online games, and the field.*" ICS dissertation, Utrecht
329. Carlos de Matos Fernandes (2023), *"In or out? The paradox of exclusionary mechanisms in keeping cooperation going."* ICS dissertation, Groningen
330. Ruohuang Jiao (2023), *"Reputation effects in peer-to-peer online markets: meta-analyses and laboratory experiments."* ICS dissertation, Utrecht
331. Marlou Ramaekers (2024), "*Informal helping. Insights from a dyadic, family and societal perspective.*" ICS dissertation, Nijmegen.
332. Kim Stienstra (2024),*"Educational quality and inequality: The interplay between schools, families, and genes."* ICS dissertation, Utrecht.

In many countries, the quality and equality of education are under pressure. Investigating the relationship between three key elements that shape pupils' performance – schools, families, and genes – is essential to understand and reduce educational inequality. This dissertation studies these three elements simultaneously because they may be mistaken for each other otherwise and may interact in ways that increase or decrease educational inequality. Theory and methods from sociology and behavioral genetics are used to analyze educational performance data of twins enriched with information on the school environment. Differences in educational performance related to family socioeconomic status and genes are smaller in more advantageous school environments. This shows that schools have the potential to compensate for educational inequality. While the importance of school quality has been highly debated, the classroom environment and socioeconomic composition are more relevant for reducing educational inequality.

Kim Stienstra (1995) obtained a bachelor's degree in Sociology and a research master's degree in Sociology and Social Research (cum laude) at Utrecht University. She conducted the present study at the department of Sociology at Utrecht University, as part of the Interuniversity Centre for Social Science Theory and Methodology (ICS).