The Social Stratification of Environmental and Genetic Influences on Education: New Evidence Using a Register-Based Twin Sample

Tina Baier,a Volker Langb

a) Leibniz Institute for Educational Trajectories; b) Bielefeld University

Abstract: The relative importance of genes and shared environmental influences on stratification outcomes has recently received much attention in the literature. We focus on education and the gene-environmental interplay. Specifically, we investigate whether—as proposed by the Scarr-Rowe hypothesis—genetic influences are more important in advantaged families. We argue that the social stratification of family environments affects children’s chances to actualize their genetic potential. We hypothesize that advantaged families provide more child-specific inputs, which enhance genetic expression, whereas the rearing environments of children in disadvantaged families are less adapted to children’s individual abilities, leading to a suppression of genetic potential. We test this relationship in Germany, which represents an interesting case due to its highly selective schooling system characterized by early tracking and the broad coverage of part-time schools. We use novel data from the TwinLife panel, a population-register–based sample of twins and their families. Results of ACE-variance decompositions support the Scarr-Rowe hypothesis: Shared environmental influences on education matter only in disadvantaged families, whereas genetic influences are more important in advantaged families. Our findings support the growing literature on the importance of the gene-environmental interplay and emphasize the role of the family environment as a trigger of differential genetic expression.

Keywords: educational inequality; gene-environment interplay; twin studies; Germany

The influence of family of origin on children’s education is a recurring topic in stratification research. What has largely been neglected in this literature is that not only social resources but also genetic influences lead to differences in education. Behavioral scientists, by contrast, commonly differentiate between genetic and environmental sources of interindividual variation. Recent findings based on an international meta-analysis show that shared environmental influences account for an average of about one-third of the total variation in educational attainment, whereas the relative importance of genes is only slightly higher (i.e., about 40 percent: Branigan, McCallum, and Freese 2013). Thus, both shared environmental and genetic influences are important to understanding individual variation in education. However, what is missing are studies addressing whether and how genetic influences on education depend on social conditions.

The Scarr-Rowe hypothesis claims that the relative importance of genes is higher in advantaged families, whereas shared environmental influences are more important in disadvantaged families (Rowe, Jacobson, and Van den Oord 1999; Scarr-Salapatek 1971). The underlying assumption is that environmental conditions provided by advantaged families fit better with children’s genetic dispositions,
which enhances genetic expression (Bates, Lewis, and Weiss 2013; Guo and Stearns 2002; Tucker-Drob and Bates 2016; Turkheimer et al. 2003). The Scarr-Rowe hypothesis emerged in the context of studies of intelligence (IQ). Certainly, IQ is strongly correlated with outcomes that are relevant to social stratification. Nonetheless, mechanisms that determine an individual’s social position are more complex and are not only determined by an individual’s IQ (e.g., Strenze 2007). Thus, for stratification research, it is important to extend studies on the Scarr-Rowe hypothesis beyond IQ. This article focuses on education as one of the most relevant indicators of social stratification.

To explain the relative importance of social compared to genetic influences on education, we combine sociological theories with behavioral genetic approaches. We use the conceptual differentiation between primary and secondary effects of social background to explain the substantial impact of shared environmental influences on educational attainment (Boudon 1974). To elaborate on how social background shapes the relative importance of genetic influences on educational attainment, we draw on the literature on stratified logics of parenting (e.g., Bodovski and Farkas 2008; Cheadle and Amato 2011; Kalil, Ryan, and Corey 2012; Lareau 2011; Lareau and Weininger 2003). Advantaged parents typically engage in so-called “strategic” or “planned” types of parenting; disadvantaged parents are more focused on addressing children’s basic needs, leading to so-called “natural growth” (Lareau 2011). Strategic parenting often entails a cultivation of skills and behaviors that are distinctive to higher social classes, so-called “concerted cultivation” (Lareau 2011). We argue that strategic types of parenting tend to be accompanied by the provision of rearing environments that are more adapted to children’s potential. In other words, such parenting behavior not only promotes skills that are rewarded in contemporary societies but also involves parenting strategies that are more child centered and hence individually adapted.

The specificity of rearing environments is important because the development of humans’ genetic potential is dependent on environmental conditions (e.g., Bronfenbrenner and Ceci 1994). Hence, if children from advantaged parents grow up in an environment that matches their individual abilities, this enhances the expression of genetic potentials, whereas the less-adapted environments more often provided by disadvantaged families lead to a suppression of genetic potential. This theoretical mechanism can explain why shared environmental influences on IQ—and potentially on education as well—play a stronger role in disadvantaged than in advantaged families. In addition, different opportunities for gene expression can accumulate over the life course (e.g., Dannefer 2003; DiPrete and Eirich 2006), as children are selected in distinct learning environments related to their social background.

We test the Scarr-Rowe hypothesis for educational attainment in Germany, which is an especially interesting case for two reasons: First, Germany has a highly stratified schooling system characterized by early tracking, leading to comparably strong secondary effects on educational attainment (e.g., Blossfeld and Shavit 1993; Hillmert and Jacob 2010; Neugebauer 2010). And second, schooling is often part-time, leaving more room for parents to shape the developmental environment of their children through investments in parenting and other resources. Given these
institutional features, we expect the social stratification of the relative importance of shared environmental and genetic influences to be stronger compared to countries with a comprehensive schooling system.

Our analyses are based on the TwinLife panel (Diewald et al. 2018). Since 2014, TwinLife has collected a wealth of information from monozygotic and same-sex dizygotic twins and their families residing in Germany. These data allow, for the first time, genetically sensitive analyses of educational attainment in Germany. Importantly, TwinLife overcomes one of the major weaknesses of many observational twin studies, as a probability-based register-sampling strategy was applied (Lang and Kottwitz 2017). We estimate genetic-sensitive-variance decomposition models (ACE models) of completed years of education based on twins, which enables us to distinguish between environmental and genetic influences (Guo and Wang 2002; Plomin et al. 2008; Rabe-Hesketh, Skrondal, and Gjessing 2008).

Our article makes two main contributions. First, we extend research on the Scarr-Rowe hypothesis on IQ to education and provide evidence for a social stratification of environmental and genetic influences on educational attainment in Germany. Second, we provide a theoretical account for the strong impact of shared environmental influences on education and their variation according to social background. Our explanation emphasizes the role of parental behavior in shaping children’s chances of their genetic potential being expressed.

**Theoretical Background**

**Shared Environmental Influences on Education**

The impact of shared environmental and genetic influences on education is well established (Branigan et al. 2013; Heath et al. 1985; Nielsen 2016; Silventoinen et al. 2004; Taubman 1976). According to an early study of the United States based on the Veteran Twin Registry, shared environmental influences explain about one-third of the total variation in education, and genes explain about 40 percent (Taubman 1976). In Norway (Norwegian Twin Panel) and Finland (Finnish Twin Cohort Study), genetic influences on education are about the same, but at 40 percent, the relative importance of shared environmental influences is somewhat higher (Heath et al. 1985; Silventoinen et al. 2004). The findings of a recent international meta-analysis (including Australia, Denmark, Finland, Germany, Italy, Norway, Spain, Sweden, the United Kingdom, and the United States) are in line with findings based on single countries (Branigan et al. 2013): Shared environmental influences account for an average of about 36 percent and genes account for about 40 percent of the variation in education. In almost every country, shared environmental influences account for more than 20 percent of the variation in education (Branigan et al. 2013). Most of the samples are not population based (only the samples for Finland, Italy, and parts of the United States are), and it is therefore important to replicate those findings.

Nonetheless, current findings for shared environmental influences on education diverge strongly from the pattern identified for most other adult characteristics, including those relevant to education, such as cognitive and noncognitive skills.
Here, the relative importance of genes far exceeds the relevance of shared environmental influences (e.g., Freese and Jao 2017; Nielsen 2016; Polderman et al. 2015; Turkheimer 2000).

Before we discuss current explanations for the comparatively strong impact of shared environmental influences on education, we briefly explain their meaning. Shared environmental influences represent all nongenetic influences that lead to similarities among siblings. They cannot simply be equated with indicators of shared family background because such factors can also make children from one family less alike. For example, siblings may share the experience of their parents’ divorce but can have very different reactions to it (Turkheimer and Waldron 2000). This in turn is a nonshared environmental influence, as it leads to differences among siblings (Turkheimer and Waldron 2000). Thus, only factors that affect children from one family in similar fashion are shared environmental influences. It is standard fare in the behavioral genetics literature that the definition of shared environmental influences is based on their effects (i.e., whether children are equally affected or not). Shared environmental influences on education, therefore, represent transmission mechanisms that affect siblings in a similar way—net of genetic factors (Freese and Jao 2017; Nielsen and Roos 2015).

To date, the literature provides two sets of explanations for shared environmental influences on education. The first set discusses the impact of assortative mating, sibling effects, and the measurement of education (Freese and Jao 2017; Nielsen and Roos 2015).

The standard approach to estimating environmental and genetic influences (see the section Data and Methods) assumes that spouses mate randomly with regard to the characteristic under study. This justifies the assumption that siblings share on average about 50 percent of their genes. However, previous research shows that spouses are more similar in their education than expected given random mating across Western societies (e.g., Blossfeld 2009). It is plausible to assume that spouses are also more similar in genetic endowments relevant to education, which in turn increases (on average) the similarity of their children with respect to genes. Assortative mating, therefore, leads to an underestimation of genetic influences and an overestimation of shared environmental influences because the genetic similarity of siblings is higher than assumed by standard models building on the random-mating assumption (e.g., Plomin et al. 2008).

Sibling effects may also account for comparably strong shared environmental influences on education. Such sibling effects can occur if the educational decisions of one sibling guide those of other siblings (Freese and Jao 2017; Nielsen and Roos 2015). For example, older siblings can serve as a role model for younger ones (Benin and Johnson 1984). In that case, shared environmental influences are the consequence of the mutual influences of siblings rather than being the direct consequence of parents’ actions.

Finally, the measurement of educational attainment can affect estimations of shared environmental influences. In general, fine-grained measures are better suited to detect differences between individuals. If coarse measures are used (e.g., tertiary vs. nontertiary education), the estimates for shared environments tend to be larger.
compared to more nuanced measures (e.g., years of education [Freese and Jao 2017; Nielsen and Roos 2015]).

The second set of explanations discusses family resources, parents’ educational expectations, and the broader family context (Freese and Jao 2017; Nielsen and Roos 2015). In light of the large body of sociologically inspired work that links parents’ material and nonmaterial resources with children’s educational attainment (e.g., Blau and Duncan 1967; Breen and Jonsson 2005; Haller and Portes 1973; Torche 2015), we focus on this set of explanations. We address this first set of explanations, which we label “alternative explanations,” in our analysis to rule out the possibility that our findings (with respect to shared environmental influences on education) are driven by any of these aspects.

To integrate current explanations of family resources, we draw on the theory of primary and secondary effects of social background on educational attainment (Boudon 1974). This theory is well established in the stratification literature to conceptualize how parents’ social background and related resources affect children’s education. Primary effects describe how parents influence children’s academic performance, and the most important input factors are cognitive and noncognitive skills. Advantaged parents provide more stimulating home environments, provide relevant learning materials, and often transmit cultural resources and interests, all of which further school-related skills (e.g., Cheadle and Amato 2011; Cunha and Heckman 2008; Kalil et al. 2012; Lareau 2011; Lareau and Weininger 2003). Secondary effects, by contrast, describe systematic differences in educational decisions by social background—net of academic achievement. Parents evaluate the costs, anticipated benefits, and likelihood of success differently according to their own social position and educational experience (Breen and Goldthorpe 1997; Erikson and Jonsson 1996). An important driver of stratified educational choices is the intention to avoid downward mobility. In order to maintain social status, advantaged parents tend to have higher educational aspirations for their children compared to disadvantaged parents, who can reach this goal with lower levels of education (Breen and Goldthorpe 1997). Thus, even if the academic performances of advantaged and disadvantaged children do not differ, children from higher social backgrounds are still more likely to choose higher levels of education than their counterparts from lower social backgrounds (Boudon 1974; Breen and Goldthorpe 1997; Erikson and Jonsson 1996).

Primary and secondary effects do not work independently of each other. Parents with high educational aspirations for their children are also more likely to prompt children’s academic skills accordingly (for a discussion on anticipatory effects, see Erikson et al. 2005). Such anticipatory behavior leads to an underestimation of secondary effects because such efforts affect children’s academic performance and are hence inadequately attributed to primary effects (Erikson et al. 2005). Nonetheless, this conceptual differentiation acknowledges that differences in education emerge not only because of differences in academic skills but also because of stratified cost–benefit calculations.

In light of the secondary effects, shared environmental influences on educational attainment are comparatively strong because parents choose educational routes based not only on their children’s academic performance. Parents tend to have
Baier and Lang GxE for Education in Germany

similar, social class–specific educational aspirations and educational choices for their children in order to maintain social status. As a consequence, even if their children differ in school-related skills, they still end up being more alike with regard to their education. Because similar educational choices within families affect the development of cognitive and noncognitive skills less directly than they affect educational attainment, we expect shared environmental influences to be stronger for educational attainment than for cognitive and noncognitive skills, as is reported in the literature (e.g., Polderman et al. 2015; Turkheimer 2000). Based on this framework, we formulate our first hypothesis:

Hypothesis 1 (H1): Secondary effects of social background explain a substantial part of the shared environmental influences on educational attainment.

The Social Stratification of Environmental Influences and the Gene-Environmental Interplay

The environment in which children develop their skills can also vary within socially defined groups (Bodovski and Farkas 2008; Cheadle 2008; Cheadle and Amato 2011; Kalil et al. 2012; Lareau 2011; Lareau and Weininger 2003). Specifically, how differently parents treat their children and the variety of inputs they provide can differ between advantaged and disadvantaged families. Such differences in rearing environments are important because human development is embedded in proximal processes (i.e., specific environmental influences that help individuals realize their genetic potential [e.g., conditions set by families, peers, or institutions]; Bronfenbrenner and Ceci 1994). To be effective, such environmental influences have to be encountered regularly (Bronfenbrenner and Ceci 1994). One prominent hypothesis concerning such a gene-environmental interaction is the Scarr-Rowe hypothesis, which claims that proximal processes leading to the actualization of genetic potential are more prevalent in enriched environmental settings (Rowe et al. 1999; Scarr-Salapatek 1971).

Against the backdrop of stratified family environments, we argue that this type of gene-environment interplay is important to understanding how genetic and environmental influences affect educational attainment. In her U.S. study, Lareau (2011) scrutinizes how distinct cultural habits consolidate daily parenting behavior. Different logics of parenting affect children’s development and skill-formation processes. Advantaged parents follow a parenting concept labeled concerted cultivation, which describes parenting practices that foster behaviors and skills that are distinctive to higher social classes (Lareau 2011). More importantly for this study is that parents from higher social backgrounds adopt an active role in their children’s development (Lareau 2011). Parents frequently plan activities with their children and provide educationally relevant inputs (e.g., books or other learning materials) to raise their children’s interests and motivation to learn. Parents more often structure children’s time outside of school and engage them in extracurricular activities (e.g., music lessons or sports clubs [Covay and Carbonaro 2010; Dumais 2006]). As a consequence, children from higher social backgrounds grow up in stimulating home environments. We argue that a more active role in parenting
also implies that home environments are more adapted to children’s individual potential and needs because parents focus on planned interactions and stimulating activities with their children. In addition, children learn to speak up for themselves and communicate with institutions, especially instructors and educators. Moreover, parents extend their efforts to the schooling context, as they interact frequently with teachers and intervene on behalf of their children if needed (Bodovski and Farkas 2008; Cheadle 2008; Cheadle and Amato 2011). The acquisition of such skills and parents’ efforts can lead to further individual adaptation of learning environments outside of the family context.

In contrast, disadvantaged parents tend to perceive their children’s development as a naturally evolving process (Lareau 2011) and follow a parenting concept labeled natural growth. Parenting is focused more on fixed interaction routines. Given their limited resources, disadvantaged parents are usually occupied with meeting the basic needs of children and are less, if at all, involved in children’s learning and out-of-school activities. This also implies that parents less often customize children’s leisure time or provide tailored inputs to foster children’s skills. As a consequence, a fit between environments and potential (or the specific abilities of children) is more often coincidental and less often planned than it is in the case of advantaged families. Furthermore, there are fewer discussions between parents and their children, and children follow their parent’s instructions instead of questioning them. And because disadvantaged parents more firmly believe that children’s education is the responsibility of teachers, parents question teachers’ behavior toward their children less often (Lareau 2011). Previous research shows that differences in parenting behavior are significantly associated with scholastic performance (Bodovski and Farkas 2008; Roksa and Potter 2011).

Originally, the concept of different logics of parenting was applied to understand how class-related differences between families emerge (i.e., why advantaged children tend to outperform their counterparts from disadvantaged families). However, we argue that the notion of different logics of parenting also allows for conclusions about differences between children from the same family: Advantaged parents plan and can afford investments that promote specific talents and interests of their children. Such individualized investments can be expected to be more effective. By contrast, disadvantaged parents have lower levels of resources and might not have the time to discover their children’s specific talents, or they lack the resources to further them individually.

The social stratification of parenting behavior, therefore, provides an understanding of the theoretical mechanisms underlying the Scarr-Rowe hypothesis from a sociological perspective. Stratified differences in parenting trigger different conditions for genetic expression. Disadvantaged parents provide environmental conditions that are less adapted to children’s individual abilities and hence less often match children’s genetic dispositions. This increases the relative importance of shared environmental influences on educational attainment. By contrast, advantaged parents provide inputs to actively foster children’s specific talents. Environmental conditions are more often in line with children’s genetic makeup, which enhances genetic expression.
In addition to environmental conditions provided by the family, children are also exposed to different school environments and related peer groups (Freese and Jao 2017); these are the consequence of stratified educational choices. For example, in the United States, which has a within-school tracking system, it is mostly disadvantaged students who attend lower tracks with less-qualified teachers (Heubert and Hauser 1999; Lucas 1999). Due the persistent stratification of learning environments, systematic differences in the realization of developmental potential can be exacerbated over the life course (Dannefer 2003; DiPrete and Eirich 2006). Given this literature, we formulate our second hypothesis:

Hypothesis 2 (H2): Shared environmental influences on educational attainment are more important for children from disadvantaged families, whereas genetic influences matter more for children from advantaged families.

To date, the Scarr-Rowe hypothesis has not been tested for educational attainment based on sibling or twin data. Nonetheless, studies report that the relative importance of shared environmental influences is sensitive to macrostructural differences (Branigan et al. 2013; Heath et al. 1985; Nielsen and Roos 2015). In most countries, shared environmental influences on educational attainment declined over the twentieth century, although they have increased in the United States (Branigan et al. 2013; Heath et al. 1985; Nielsen and Roos 2015).

In addition to twin studies, in which genetic influences are indirectly assessed based on their information, genetic relatedness, and common rearing (for an overview, see Plomin et al. 2008), researchers also use direct measures of genes based on DNA samples (for an overview, see Conley 2016) to study the impact of genes. Studies in this research area have used polygenic scores (PGS) for education—a measure constructed based on DNA samples—and examined whether the association between these PGS and education changed by social background (Conley et al. 2015; Domingue et al. 2015). Conley et al. (2015) report that the effect of PGS on education did not systematically vary by social background. The study by Domingue et al. (2015), by contrast, shows that the association between PGS and education decreased with social background for younger birth cohorts. Thus, previous studies based on PGS do not support the Scarr-Rowe hypothesis. PGS are conceptually much closer to the pathways of genetic transmission than twin studies. But until recently, PGS studies related to social stratification outcomes (such as educational attainment) had limited predictive power (Okbay et al. 2016). For example, the PGS used in these studies explained only about 2 to 3 percent of the total variation in educational attainment (Conley et al. 2015). However, genomic analyses are rapidly evolving, and advancements in genotyping procedures and larger sample sizes have already led to significant improvements in the predictive power of PGS for educational attainment (Okbay et al. 2016). According to a recent study, current PGS account for about one-fifth of the variation in educational attainment (Liu 2018). It is likely that these improvements also affect findings on the moderating effect of social background on genetic influences (see also Conley et al. 2015). Thus, the findings of previous studies are preliminary and need to be replicated before we can draw conclusions about the moderating effect of social background (Domingue et al. 2015).
A comparatively large body of literature has tested the Scarr-Rowe hypothesis for IQ based on twin or sibling designs (e.g., Figlio et al. 2017; Guo and Stearns 2002; Tucker-Drob and Bates 2016; Tucker-Drob, Briley, and Harden 2013; Bates et al. 2013). Here, the evidence is also mixed. Most research refers to the United States and supports the Scarr-Rowe hypothesis. However, these findings have recently been challenged by studies that include countries other than the United States (Tucker-Drob and Bates 2016), and the most recent study for the United States also finds no support for the Scarr-Rowe hypothesis regarding IQ (Figlio et al. 2017).

**The German Context**

Germany represents an especially interesting case for our study, as the link between social background and children’s education is exceptionally strong (e.g., Blossfeld and Shavit 1993; Hillmert and Jacob 2010; Neugebauer 2010). The educational system plays a pivotal role: Even though tracking is a common feature of European educational systems, children in Germany are tracked as early as the age of 10 to 12 years into one of the three hierarchically structured secondary-school types (*Hauptschule*, *Realschule*, and *Gymnasium*). Secondary-schooling tracks differ strongly in their curricula and length and are linked to postsecondary alternatives. Only a diploma from the highest secondary level (*Gymnasium*) entitles a student to study in tertiary-level institutions. Only recently have reforms been implemented to loosen the link between secondary education and postsecondary alternatives (Bethhäuser 2017).

It is possible to switch between secondary-schooling tracks. However, upward moves are rare and linked to social background (Henz 1997; Jacob and Tieben 2010). Secondary-school-type decisions are—with variations in regulations between federal states—made jointly by teachers and parents. Teacher recommendations are provided for all students, though they are not binding in every federal state. Recommendations are based on children’s academic performance and expectations about future development. However, research shows that children from lower social background are (*ceteris paribus*) less likely to receive a recommendation for the highest academic track than are children from higher social backgrounds (Baumert, Trautwein, and Artelt 2003). Furthermore, the quality of the learning environment differs between the tracks due to compositional effects (i.e., students are more similar in terms of social background and achievement) and institutional effects (i.e., curricula and teaching personnel; Maaz et al. 2008).

Another distinct feature of the German educational system is that primary and secondary schools are frequently part-time. Full-time schools, by contrast, end about late afternoon and provide food, extracurricular activities, and usually support with homework. Initiatives toward an expansion of full-time schooling took place from 2002 to 2008 (Sekretariat der Kultusministerkonferenz 2008, 2010). Despite these efforts, part-time schools are overall still widespread in Germany. However, there is large variation in part-time schooling, especially between eastern and western federal states. Previous research finds that extracurricular activities for children can lower social disparities in performance (Covay and Carbonaro 2010; Dumais 2006). Thus, part-time schools leave much more room for parents’
resources and parenting practices to influence children and can (depending on their quality) reinforce social disparities in academic performance, whereas full-time schools can serve as an equalizer (Fischer and Theis 2014; Kuhn and Fischer 2011). Taken together, the stratified schooling system and the broad coverage of part-time schools make children’s educational attainment particularly sensitive to social background. Therefore, we argue that the proposed mechanisms driving the social stratification of shared environmental influences on education are comparatively strong in Germany.

Data and Methods

Sample

We use data from the first wave of the German twin panel TwinLife (Diewald et al. 2018). TwinLife collects extended twin family information (i.e., on twins, their parents, siblings, and partners) for monozygotic (MZ) and same-sex dizygotic (DZ) twins residing in Germany. Due to the social and regional stratified probability-based sampling strategy, TwinLife provides a unique opportunity to analyze a broad range of the social spectrum with behavioral genetic methods (Lang and Kottwitz 2017). We examine twin pairs from the oldest birth cohort (1990–1993), who were between 22 and 25 years old at the time of the survey. We study twin pairs in which both twins provided valid information on their education (for 3 percent of the sample, this information is missing or incomplete). The analysis sample comprises 1,930 twins (47 percent DZ; 53 percent MZ).

Variables

Educational attainment is measured in years of education. We transformed the categorical information on respondents’ educational level (general education and professional training) into corresponding years of education using established coding schemes for Germany (Socio-Economic Panel Group 2017; see online supplement Table S1). The coding scheme is based on educational levels and not on the actual time spent in educational institutions. The continuous measure has the advantage that linear genetic-sensitive-variance decompositions can be estimated (see the section Analytical Strategy). We centered educational attainment to 13 years of education in all our analyses.

Due to the age range of twins, about 59 percent were still enrolled in professional training at the time of the interview. For these twins, we do not know whether they finish their education or not. To address this uncertainty, we ran two sets of analyses. The first set assumes that all twins currently enrolled finish their professional training and that they do not attend additional professional training leading to a higher educational degree at some later point. In this scenario, twins get assigned the associated additional years of education of the degree they are currently pursuing (so-called “upper-bound scenario” with respect to final years of education). The second set of analyses relies on the assumption that all twins currently enrolled drop out and again that they do not attend additional professional
training leading to a higher educational degree at some later point. In this case, twins do not get assigned the additional years of education of the degree they are currently pursuing (so-called “lower-bound scenario” with respect to final years of education).

To evaluate the findings, it is important to note that in Germany, the share of young adults (aged between 20 and 30 years) without vocational training or tertiary education fluctuated between 14.6 percent and 12.9 percent from 1996 to 2015 (Bundesinstitut für Berufsbildung 2017). Thus, most young adults who do not finish their current professional training will switch to another professional training and will not entirely drop out. Moreover, given the strong linkage between secondary education and postsecondary alternatives in Germany, a switch within the current type of profession (horizontal change) is more likely than a vertical switch between vocational training and tertiary education. The majority of twins currently in professional training will, therefore, receive a certificate either from vocational training or tertiary education. Thus, the final distribution of educational attainment—which is observable in a few years—is most likely closer to the upper-bound than to the lower-bound scenario.

We indicated social background with parents’ education measured in years. We chose education because of its significant role in shaping an offspring’s chances of educational attainment (e.g., Bukodi and Goldthorpe 2013; Hout and DiPrete 2006). Parents’ education covers not only transmission mechanisms that run through economic resources but also transmission mechanisms driven by socially stratified educational choices and specific parental habits and practices, all of which affect children’s skill formation (Breen and Goldthorpe 1997; Erikson and Jonsson 1996). We used the same coding scheme that we used for the twins (see Supplementary Table S1). We took the information on the parent with the highest educational level (“dominance principle”) and centered parents’ years of education on 13 years of education.

For the subgroup analyses, we distinguished between the following overlapping educational subgroups: 7 to 11, 9 to 12, 10 to 13, 11 to 14, 12 to 15, 13 to 16, 14 to 17, 15 to 18, 16 to 19, and 17 to 20 years of parental education. If we had created mutually exclusive subgroups based instead on several years of education, we would run the risk of our findings relating to a change in the variance components being driven by the cutoff points we have chosen. At the same time, however, analyses that compare variance components for every single year of parents’ education are not feasible due to small sample sizes. Overlapping groups have the advantage that our conclusions are not dependent on cutoff points and offer a more nuanced understanding of the changes in the variance components over the entire range of parents’ education.

To account for sibling effects, we included a measure of the closeness of twins in our analyses based on three items with a five-point rating scale. We used the following three items: (1) How often do you talk about important things with (name of the other twin)? (2) How often do you attempt to cheer up (name of the other twin)? (3) How close do you feel to (name of the other twin)? We used confirmatory factor analysis based on a structural equation model to construct a single indicator of these items. The coefficient of determination for this indicator was 0.90.
As information on parents’ education is sometimes missing (4.63 percent for mothers; 37.88 percent for fathers), we imputed missing information by means of multiple imputation with chained equations and created 20 imputations for each observation (van Buuren et al. 2006). All of the predictors used for the imputation were at the family level (i.e., the imputation model was based on information about the parents). Thus, the imputation model is not suitable to predict missing information for twins individually. Our results are robust with and without the imputed information on parents’ education (see online supplement Table S3 and Table S4).

Table 1 presents the distributions of the sample characteristics. The main variables are distributed fairly similarly among MZ and DZ twins. The closeness indicator is exceptional, as MZ twins are significantly closer to one another than DZ twins ($p < 0.001$). Differences in closeness between MZ and DZ have been found previously (Fortuna, Goldner, and Knafo 2010) and might be rationalized in terms of the stronger similarity of MZ twins.

Analytical Strategy

Our analyses are based on a classical twin design (CTD) (e.g., Plomin et al. 2008). The CTD is a method in behavioral genetics that is widely used to estimate the relative importance of environmental and genetic influences (Plomin et al. 2008). Twins are born and raised at the same time, and MZ twins are additionally genetically alike; DZ twins share on average 50 percent of the 1 percent of all genes in which humans tend to vary. A CTD uses this knowledge to decompose the variance of an outcome into a component associated with additive genetic influences (A), a component associated with shared environmental influences (C), and a component associated with unique environmental influences, including the error term of the decomposition (E) (Table 2). This type of analysis is called ACE-variance decomposition.

To identify the A and C components, the CTD relies on additional assumptions. First, it is assumed that genetic effects are additive, which means that the effects of different genes are independent. This implies that there is no epistasis (i.e., there are no interactions between genetic influences that affect the outcome under study). Second, the CTD assumes that the genetic and environmental components are additive (i.e., that there are no correlations or interactions between genes and their environment). This assumption is challenged by the Scarr-Rowe hypothesis and therefore is tested in the second part of our analysis focusing on H2. The third assumption is the equal environments assumption (EEA), which states that the trait under study is not affected by the fact that MZ twins are differently treated by their environment than DZ twins (Scarr and Carter-Saltzman 1979). The EEA has been critically evaluated for several, mostly psychological, traits. We are not aware of any study that investigates the EEA with regard to education, though studies report that more similar environments of MZ twins do not lead to an overestimation of genetic influences on IQ (Derks, Dolan, and Boomsma 2006). The fourth assumption states that spouses mate randomly. Given random mating, the genetic similarity of siblings is on average about 0.5. As stated above, assortative mating increases
Table 1: Descriptive statistics.

<table>
<thead>
<tr>
<th></th>
<th>DZ mean/SD</th>
<th>min</th>
<th>max</th>
<th>N</th>
<th>MZ mean/SD</th>
<th>min</th>
<th>max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual (twin) level variables:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twins’ years of education:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-bound scenario</td>
<td>12.8</td>
<td>7</td>
<td>18</td>
<td>900</td>
<td>13.0</td>
<td>7</td>
<td>18</td>
<td>1,030</td>
</tr>
<tr>
<td></td>
<td>(2.35)</td>
<td></td>
<td></td>
<td></td>
<td>(2.23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper-bound scenario</td>
<td>14.2</td>
<td>7</td>
<td>18</td>
<td>900</td>
<td>14.5</td>
<td>7</td>
<td>18</td>
<td>1,030</td>
</tr>
<tr>
<td></td>
<td>(3.04)</td>
<td></td>
<td></td>
<td></td>
<td>(2.90)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>-0.23</td>
<td>-4.06</td>
<td>1.03</td>
<td>900</td>
<td>0.21</td>
<td>-4.06</td>
<td>1.03</td>
<td>1,030</td>
</tr>
<tr>
<td></td>
<td>(1.03)</td>
<td></td>
<td></td>
<td></td>
<td>(0.93)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Family-level (twin-pair-level) variables:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twin pair’s age</td>
<td>23.03</td>
<td>21</td>
<td>25</td>
<td>450</td>
<td>23.05</td>
<td>21</td>
<td>25</td>
<td>515</td>
</tr>
<tr>
<td></td>
<td>(0.81)</td>
<td></td>
<td></td>
<td></td>
<td>(0.83)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twin pair’s sex (1 = female)</td>
<td>0.57</td>
<td>0</td>
<td>1</td>
<td>450</td>
<td>0.59</td>
<td>0</td>
<td>1</td>
<td>515</td>
</tr>
<tr>
<td></td>
<td>(0.50)</td>
<td></td>
<td></td>
<td></td>
<td>(0.49)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parents’ years of education:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest in family</td>
<td>13.37</td>
<td>7</td>
<td>20</td>
<td>449</td>
<td>13.15</td>
<td>7</td>
<td>20</td>
<td>513</td>
</tr>
<tr>
<td></td>
<td>(2.98)</td>
<td></td>
<td></td>
<td></td>
<td>(3.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest in family (imputed)</td>
<td>13.86</td>
<td>7</td>
<td>20</td>
<td>450</td>
<td>13.66</td>
<td>7</td>
<td>20</td>
<td>515</td>
</tr>
<tr>
<td></td>
<td>(2.76)</td>
<td></td>
<td></td>
<td></td>
<td>(2.80)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>12.72</td>
<td>7</td>
<td>20</td>
<td>425</td>
<td>12.46</td>
<td>7</td>
<td>20</td>
<td>482</td>
</tr>
<tr>
<td></td>
<td>(2.89)</td>
<td></td>
<td></td>
<td></td>
<td>(2.68)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother (imputed)</td>
<td>12.71</td>
<td>7</td>
<td>20</td>
<td>442</td>
<td>12.43</td>
<td>7</td>
<td>20</td>
<td>509</td>
</tr>
<tr>
<td></td>
<td>(2.84)</td>
<td></td>
<td></td>
<td></td>
<td>(2.62)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father</td>
<td>13.29</td>
<td>7</td>
<td>20</td>
<td>257</td>
<td>13.31</td>
<td>7</td>
<td>20</td>
<td>276</td>
</tr>
<tr>
<td></td>
<td>(3.09)</td>
<td></td>
<td></td>
<td></td>
<td>(3.17)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father (imputed)</td>
<td>13.20</td>
<td>7</td>
<td>20</td>
<td>406</td>
<td>13.09</td>
<td>7</td>
<td>20</td>
<td>451</td>
</tr>
<tr>
<td></td>
<td>(2.62)</td>
<td></td>
<td></td>
<td></td>
<td>(2.70)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: TwinLife wave 1; own calculations.

the genetic similarity of siblings. If information on parents’ education is available, it is possible to estimate an average genetic correlation for DZ twins that corrects for assortative mating (Loehlin, Harden, and Turkheimer 2009). The correction is given by \(0.5 + 0.5 \cdot h^2 \cdot r_p\), where \(h^2\) denotes the share of genetic influences (A) estimated without correction for assortative mating and \(r_p\) denotes the correlation of parents with respect to the trait under study (Loehlin et al. 2009). In our sample, \(r_p\) is 0.46, whereas \(h^2\) is 0.43 for the lower-bound scenario and 0.46 for the upper-bound scenario. These inputs imply an assumed average genetic correlation for DZ twins of 0.60 for both scenarios.

We test our hypotheses as follows: We first estimated ACE-variance decompositions for the lower- and upper-bound scenarios using the linear multilevel mixed-effects parameterization developed by Rabe-Hesketh et al. (2008) without controlling for parental educational background (base model). This parameterization can be extended (like the standard regression approach) by including explanatory variables on which the mean of the outcome (in our case, twins’ years of education) is regressed. We then controlled for parents’ education in model 1 to test H1. Next, we stepwise controlled for parents’ assortative mating (model 2) and
Table 2: Variance decomposition based on the CTD.

<table>
<thead>
<tr>
<th>Variance component</th>
<th>Definition</th>
<th>Differences within twin pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MZ</td>
</tr>
<tr>
<td>A</td>
<td>additive genetic influences</td>
<td>no</td>
</tr>
<tr>
<td>C</td>
<td>shared environmental influences (that make twins alike—net of genes; e.g., shared effects of parents’ education or financial resources)</td>
<td>no</td>
</tr>
<tr>
<td>E</td>
<td>nonshared environmental influences (leading to differences between twins—net of genes; e.g., selective parenting, selective peer influences, and measurement error)</td>
<td>yes</td>
</tr>
</tbody>
</table>

Sibling effects (model 3). To test H2, we estimated ACE-variance decomposition models separated by parents’ years of education. This research strategy is also known as nonparametric gene-environment interaction analysis (Guo and Wang 2002). All analyses were computed with the statistical software Stata (14.2) using acelong.ado (Lang 2017).

Results

Table 3 and Figure 1 present the results for the ACE-variance decompositions for twins’ years of education. We start with the base model. In the lower-bound scenario, the mean value for years of education was about 12.9 years (constant +13) and 14.4 years for the upper-bound scenario. The total variances were 5.3 and 8.8 years of education, respectively. Genetic influences (A) accounted for about 44 percent of the total variation in educational attainment in the lower-bound scenario and about 41 percent in the upper-bound scenario; shared environmental influences (C) accounted for 27 percent in the lower-bound scenario and about 40 percent in the upper-bound scenario. Accordingly, in the lower-bound scenario, 30 percent can be attributed to nonshared environmental influences and measurement error (E), and 20 percent can be attributed as such in the upper-bound scenario. In line with theoretical expectations and previous findings (i.e., 36 percent C and 40 percent A; see Branigan et al. 2013), we find substantial shared environmental influences on educational attainment in Germany.

In model 1, we examine the role of secondary effects and controlled for education. First of all, we found that parental education has a strong effect on twins’ education in both scenarios. In the lower-bound scenario, twins’ educational attainment increases by 0.25 years for each additional year of parental education. About 10 percent of the total variation in twins’ years of education is explained by parents’ education (see explained variance \(R^2\) in percent). In the upper-bound scenario, the effect of parents’ education is about 0.38 years, and 14 percent of the total variance is explained. In both scenarios, the relative importance of genetic influences remained stable. By contrast, the relative importance of shared environmental influences was smaller. Shared environmental influences explain about 16 percent of the
Table 3: ACE-variance decompositions for twins’ years of education.\textsuperscript{a}

<table>
<thead>
<tr>
<th></th>
<th>Lower-bound scenario</th>
<th>Upper-bound scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b/variance</td>
<td>b/variance</td>
</tr>
<tr>
<td><strong>Base model:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assumed genetic DZ correlation</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.10 (0.07)</td>
<td>1.38 (0.09)</td>
</tr>
<tr>
<td>Total variance</td>
<td>5.26 (0.23)</td>
<td>8.84 (0.23)</td>
</tr>
<tr>
<td>A in %</td>
<td>43.57 (7.97)</td>
<td>41.00 (6.52)</td>
</tr>
<tr>
<td>C in %</td>
<td>26.59 (7.14)</td>
<td>39.73 (6.21)</td>
</tr>
<tr>
<td>E in %</td>
<td>30.32 (2.88)</td>
<td>19.72 (1.78)</td>
</tr>
<tr>
<td><strong>Model 1 (parents’ education):</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assumed genetic DZ correlation</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Parents’ years of education</td>
<td>0.25\textdagger</td>
<td>0.38\textdagger</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.29 (0.06)</td>
<td>1.08 (0.09)</td>
</tr>
<tr>
<td>Total variance</td>
<td>5.26 (0.23)</td>
<td>8.84 (0.23)</td>
</tr>
<tr>
<td>A in %</td>
<td>43.36 (7.83)</td>
<td>41.50 (6.70)</td>
</tr>
<tr>
<td>C in %</td>
<td>16.33 (7.03)</td>
<td>24.46 (6.31)</td>
</tr>
<tr>
<td>E in %</td>
<td>30.34 (2.63)</td>
<td>19.67 (1.77)</td>
</tr>
<tr>
<td>Explained variance (R2) in %</td>
<td>9.97</td>
<td>14.37</td>
</tr>
<tr>
<td><strong>Model 2 (assortative mating):</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assumed genetic DZ correlation</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>Parents’ years of education</td>
<td>0.25\textdagger</td>
<td>0.38\textdagger</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.29 (0.06)</td>
<td>1.08 (0.09)</td>
</tr>
<tr>
<td>Total variance</td>
<td>5.26 (0.23)</td>
<td>8.84 (0.23)</td>
</tr>
<tr>
<td>A in %</td>
<td>54.28 (9.77)</td>
<td>51.40 (8.30)</td>
</tr>
<tr>
<td>C in %</td>
<td>5.41 (8.76)</td>
<td>14.56 (7.76)</td>
</tr>
<tr>
<td>E in %</td>
<td>30.34 (2.63)</td>
<td>19.67 (1.77)</td>
</tr>
<tr>
<td>Explained variance (R2) in %</td>
<td>9.97</td>
<td>14.37</td>
</tr>
</tbody>
</table>
Table 3 continued

<table>
<thead>
<tr>
<th>Model 3 (sibling effects):</th>
<th>Lower-bound scenario b/variance</th>
<th>Upper-bound scenario b/variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assumed genetic DZ correlation</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>Parents’ years of education</td>
<td>0.25†</td>
<td>0.38†</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>Closeness of twins</td>
<td>0.24†</td>
<td>0.28†</td>
</tr>
<tr>
<td>(0.05)</td>
<td>(0.07)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.29‡</td>
<td>1.09‡</td>
</tr>
<tr>
<td>(0.06)</td>
<td>(0.09)</td>
<td></td>
</tr>
<tr>
<td>Total variance</td>
<td>5.26†</td>
<td>8.84‡</td>
</tr>
<tr>
<td>(0.23)</td>
<td>(0.23)</td>
<td></td>
</tr>
<tr>
<td>A in %</td>
<td>55.24†</td>
<td>53.60‡</td>
</tr>
<tr>
<td>(9.71)</td>
<td>(8.38)</td>
<td></td>
</tr>
<tr>
<td>C in %</td>
<td>2.28</td>
<td>10.89</td>
</tr>
<tr>
<td>(8.58)</td>
<td>(7.84)</td>
<td></td>
</tr>
<tr>
<td>E in %</td>
<td>30.58‡</td>
<td>19.56‡</td>
</tr>
<tr>
<td>(2.64)</td>
<td>(1.78)</td>
<td></td>
</tr>
<tr>
<td>Explained variance (R2) in %</td>
<td>11.90</td>
<td>15.95</td>
</tr>
</tbody>
</table>

Source: TwinLife wave 1; own calculations. † All models are based on 965 twin pairs, and clustered standard errors are calculated at the twin-pair level. Clustered standard errors are in parentheses. ‡ P (Z > |z|) < 0.05; † P (Z > |z|) < 0.01 (two-tailed tests).

total variation in the lower-bound scenario and 25 percent in the upper-bound scenario. Hence, the results for both scenarios show that the relative importance of genes hardly changed once parents’ education was controlled for. By contrast, the relative importance of shared environmental influences was about 11 to 15 percent lower if parents’ education is controlled for. These results indicate that parents’ education mostly accounts for shared environmental influences and to a much less extent for genetic influences. As a consequence, the results offer support for H1 (i.e., that socially stratified educational choices [secondary effects of social background] explain a substantial part of shared environmental influences on education). However, a substantial impact of shared environmental influences on education remains unexplained after controlling for parents’ education.

In model 2 we analyzed the role of alternative explanations. We assess their effects in conjunction with the findings for H1. In our sample, assortative mating increased the genetic similarity to 0.6 (Table 3: assumed genetic DZ correlation). The relative importance of genetic influences was about 54 percent in the lower-bound scenario and about 51 percent in the upper-bound scenario, whereas the relative importance of shared environmental influences was only about 5 percent in the lower-bound scenario and 15 percent in the upper-bound scenario. Thus, in both scenarios, shared environmental influences are about 10 percent lower compared to the findings in which random mating is assumed. Furthermore, the impact of the shared environmental component is no longer significant. Thus, both parents’
education and assortative mating account for shared environmental influences on education in Germany.

In model 3, we accounted for sibling effects. The closeness of twins had a positive significant effect on their educational attainment. An increase of one standard deviation in twins’ closeness is associated with an increase of about 0.25 years of education and an about 2 percent increase in explained variance compared to model 2. Hence, sibling effects account for shared environmental influences to some extent, but they were not as strong of an explanatory factor as parental education and parental assortative mating.

In sum, when all theoretical explanations are accounted for, the relative importance of shared environmental influences on educational attainment is no longer significant. In fact, shared environmental influences were only about 2 percent in the lower-bound scenario and about 11 percent in the upper-bound scenario.

Now we turn to the results for the social stratification of shared environmental and genetic influences on education. The findings are visualized in Figure 2 (see online supplement Table S2 for additional information on the estimation results). The following subgroup-specific ACE decompositions are conditional on parental sociological science | www.sociologicalscience.com 159 February 2019 | Volume 6
education. As a consequence, the variance related to the effect of parental education on children’s educational attainment is excluded in these variance decompositions. Our results in models 1 to 3 have shown that this variance is mainly associated with shared environment influences.

In line with theoretical expectations and our previous analyses (Table 3, model 1), twins’ mean level of educational attainment increases continuously with parents’ education in both scenarios. In the lower-bound scenario, we found that twins whose parents have the lowest level of education (7 to 11 years of education) have about 11.7 years of education. Twins whose parents have the highest level of education (17 to 20 years) have about 13.9 years of education. For the upper-bound scenario, mean values range from 12.7 to about 16.1 years. Overall, the total variances of educational attainment were quite stable across the subgroups.

In the lower-bound scenario, shared environmental influences accounted for about one-third of the variation among twins whose parents have the lowest level of education. Shared environmental influences mattered most (about 44 percent) among twins whose parents have 9 to 12 years of education and were absent among twins whose parents have 16 years of education or more. By contrast, the relative importance of genetic influences was smallest (about 24 percent) among
twins whose parents have 9 to 12 years of education and highest (about 70 percent) among twins whose parents have the highest level of education. Thus, the relative importance of shared environmental influences tended to be smaller among twins whose parents have more education, whereas the relative importance of genetic influences tended to be larger in these educational subgroups.

This pattern was even more pronounced in the upper-bound scenario: The relative importance of shared environmental influences was most relevant (about 44 percent) among twins whose parents have the lowest level of education, whereas shared environmental influences are absent for twins whose parents have the highest level of education. The relative importance of genetic influences was about 40 percent among twins whose parents have the lowest level of education and about 80 percent among twins whose parents have the highest level of education.

These findings support our second hypothesis: In both scenarios, shared environmental influences were most pronounced in families with lower levels of education, whereas genes were more important in families with higher levels of education. In neither the lower-bound nor the upper-bound scenario did we find shared environmental influences among twins whose parents are highly educated over and above shared environmental influences induced by parents’ education. Interpreting these findings in conjunction with the mean levels of education, our results show that shared environmental influences in less-educated families are rather detrimental for educational attainment because, on average, twins attained lower levels of education than twins with highly educated parents where shared environmental influences were absent.

Conclusion and Discussion

In this article, we investigated how shared environmental and genetic influences affect educational attainment and tested the Scarr-Rowe hypothesis for educational attainment (i.e., we analyzed whether shared environmental and genetic influences are socially stratified).

Our theoretical explanations combine sociological perspectives with behavioral genetic approaches. Firstly, to explain the comparatively strong shared environmental influences on education compared to IQ, we use the framework of primary and secondary effects of social background on educational attainment (Boudon 1974). Because schooling decisions are socially stratified and operate over and above academic performance, we proposed that secondary effects explain a substantial part of the shared environmental influences on education. Secondly, to motivate the social stratification of gene expression proposed by the Scarr-Rowe hypothesis, we extend the sociological literature on stratified parenting (e.g., Lareau 2011): We argue that advantaged parents provide learning environments that are more child centered and adapted to children’s potential and needs. This individual adaptation of children’s learning environment matters because it leads to better conditions for gene expression. In consequence, we expected the relative importance of genetic influences to be stronger in advantaged families and shared environmental influences to be more important in disadvantaged families. Furthermore, it is plausible that if learning environments are persistently socially stratified, then the different
likelihoods of genetic potentials being realized are amplified as children get older (Dannefer 2003; DiPrete and Eirich 2006). Mechanisms of cumulative advantage might therefore explain why differences in parenting can have a long-lasting impact on children’s educational biographies.

Our results based on a socially stratified random sample of young-adult twins show that shared environmental influences account for an average of about one-third of the total variation in education in Germany. These findings are in line with previous findings (Branigan et al. 2013). The impact of shared environmental influences was driven partly by parents’ education, which supports our first hypothesis on the importance of socially stratified educational choices. However, assortative mating also accounts for shared environmental influences on education in Germany. Thus, both parents’ education and assortative mating are main explanatory mechanisms of shared environmental influences of education. Future research should, therefore, systematically account for assortative mating to avoid an overestimation of shared environmental influences on education.

Our findings on the social stratification of shared environmental and genetic influences provide evidence for the Scarr-Rowe hypothesis applied to education. For twins whose parents have low levels of education (up to 12 years), shared environmental influences accounted for about 40 percent of the total variation in educational attainment. By contrast, for twins whose parents are highly educated (more than 17 years), shared environmental influences were absent, and genetic influences accounted for about 75 percent of the total variation in educational attainment. These findings provide support for the impact of socially stratified family environments and different logics of parenting on gene expression with respect to educational attainment: More-educated parents provide rearing and learning environments that are more often in line with children’s individual abilities and genetic makeup, helping them to realize their developmental potential. In disadvantaged families, environmental influences are less adapted to children’s potential and needs and hence less specific to children’s genetic disposition. This can explain why genetic influences are much weaker, and thus why relative shared environmental influences are much stronger, in less-educated families. In addition, lower mean levels of education for young adults whose parents are less educated indicate that those shared environmental influences tend to be detrimental for children’s education and constrain the realization of children’s innate talents.

Our results relate to Germany, which has a highly stratified schooling system and a broad coverage of part-time schools. Both of these institutional facets imply a higher sensitivity to social background influences compared to educational systems that lack these features. Given these institutional features, the social stratification we found can be expected to be strong in comparison to other national contexts. Future research is needed to systematically assess the effects of cross-country differences in educational institutions on gene expression with regard to education and to examine to what extent the social stratification of environmental and genetic influences depends on factors such as the timing of tracking or the coverage of full-time schools.

It is important to keep in mind that the individuals we studied are predominantly still in education. We addressed the uncertainty by providing lower-
Baier and Lang GxE for Education in Germany

upper-bound estimates of environmental and genetic influences on education. These estimates are based on assumptions about the future educational career of the twins we analyzed, which can also influence our estimates of variance components. It is therefore important to replicate our results once twins finished their education. Nonetheless, results for both scenarios provided a clear pattern that supports the Scarr-Rowe hypothesis.

Furthermore, we did not explicitly analyze factors leading to primary and secondary effects (e.g., educational choices or aspirations, extracurricular activities, and parenting behaviors). Thus, more genetically informed research on these mechanisms is needed (e.g., to assess how parenting behaviors suppress or enhance children’s potential to realize their genetic disposition).

Finally, two limitations that come along with the CTD need to be discussed: First, we cannot rule out that our findings are driven by systematic differences in genetic variation across the social strata. Such differences could stem from stratified assortative-mating patterns: If assortative mating is less pronounced among more-educated compared to less-educated parents, our findings on the relative importance of genetic influences on education would be upwardly biased (Conley et al. 2015). Although this argument is reasonable, current evidence based on PGS on education within sibling samples does not support this assumption (Conley et al. 2015). Sibling analysis provides a powerful tool in this context because siblings are exposed to the same family background influences, whereas differences in their genes are random. The results show that the effect of PGS on education is smaller in models that examine children from different families (“between-family analysis”) compared to models that examine different children from the same family (“within-family analysis”). Thus, controlling for unobserved influences within the family, the effect of genetic endowment measured through PGS on educational attainment is stronger than without these controls. These results contradict the expectation that greater genetic variance among more-educated parents upwardly biases the findings on the relative importance of genes for educational attainment.

Second, we cannot rule out that gene-by-gene instead of gene-environmental interactions are at work. In other words, it could be that genetic variants of children interact with genetic variants of parents that are associated with parents’ education. Such mechanisms can only be detected if genotyped data of parents and children are available. Such data are not available in Germany yet. However, additional analysis from the study of Conley et al. (2015) shows that parental genotype has not had an independent effect on children’s education if children’s PGS and parents’ schooling are controlled for.

In light of recent developments in molecular genetics, findings on the Scarr-Rowe hypothesis on education need to be replicated for Germany based on genotyped data. Nonetheless, we acknowledge the added value of the CTD. Molecular genetic studies are interested in scrutinizing the role of specific genetic variants, whereas behavioral genetic studies in general are also interested in the relative importance of overall genetic influences on individual variation. Complex traits, such as educational outcomes, are influenced by many different genetic and environmental factors, with each single factor having a rather small effect. Twin designs, however, facilitate the investigation of whole-genome effects—rather than specific
genetic variants—across different social conditions. Importantly, these approaches are not mutually exclusive, and we believe that they fruitfully complement each other, which will significantly enhance our current understanding on how social and genetic influences shape individuals’ life chances.

Overall, our results provide evidence for socially stratified environmental and genetic influences on educational attainment in Germany based on a quantitative genetic design. We provided a theoretical account for the underlying mechanisms that are rooted in differences in the quality of the learning environment and related parenting. The social stratification of learning environments shapes the realization of genetic predispositions and thus contributes to social disparities in educational attainment.

Notes

1 Neither of the logics of parenting is better or worse, per se. Different practices provoke different types of skills (e.g., disadvantaged children can become more autonomous as they decide by themselves what they want to do in their leisure time; they are also better in learning by experience). The concept of concerted cultivation, however, fosters skills that are rewarded more by contemporary educational institutions.

2 During 2002 and 2008, the share of full-time schooling increased substantially, although different school tracks were expanded unevenly: Full-time schooling in primary schools and intermediate secondary-schools (Realschule) tripled and quadrupled in the lowest secondary school track (Hauptschule), whereas full-time schools for the highest schooling track (Gymnasium) doubled (Sekretariat der Kultusministerkonferenz 2008, 2010).

3 Results remain robust when we used parents’ occupational status (as indicated by the International Socio-Economic Index of Occupational Status [ISEI]) instead of parents’ education (results are available upon request).

4 In addition to the codes used for the young-adult twins, we included a doctoral category for the parents (see online supplement Table S1: coding scheme for years of education).

5 Because MZ twins are closer to each other than DZ twins, we run an additional analysis in which we accounted for differences in closeness by zygosity. In addition, we examined whether the closeness of MZ and DZ twins varies by parental education. The results remained stable for both the upper- and lower-bound scenarios (results are available upon request).

6 The correction for assortative mating changes only the estimations for the relative importance of shared environmental and genetic influences.

7 Furthermore, we base our analysis for the subgroups by parental education on the assumption of random mating because the majority of assortative mating between parents based on education is already accounted for by conditioning on parental education.

References


Baier and Lang GxE for Education in Germany


Polderman, Tinca J. C., Beben Benyamin, Christiaan A. de Leeuw, Patrick F. Sullivan, Arjen van Bochoven, Peter M. Visscher, and Danielle Posthuma. 2015. “Meta-analysis of the Heritability of Human Traits Based on Fifty Years of Twin Studies.” *Nature Genetics* 47:702–9. [https://doi.org/10.1038/ng.3285](https://doi.org/10.1038/ng.3285).


Acknowledgments: This article was supported by a grant from the German Research Foundation (awarded to Martin Diewald [DI 759/11-1], Rainer Riemann [RI 595/8-1], and Frank M. Spinath [SP 610/6-1]) and the European Consortium for Sociological Research (ECSR) internship grant. We received excellent comments from participants in the ECSR conference in Milan in August 2017, the “Reading Group” held at the University of Oxford in November 2017, and the “Social Inequality and Social Demography” colloquium held at Humboldt University of Berlin in January 2018. We also would like to thank especially Martin Diewald, Anette Fasang, and the editors for their valuable feedback on an earlier version of the article.

Tina Baier: Department Educational Decisions and Processes, Migration, Returns to Education, Leibniz Institute for Educational Trajectories. E-mail: tina.baier@lifbi.de.

Volker Lang: Department of Sociology, Bielefeld University. E-mail: volker.lang@uni-bielefeld.de.