The Social Stratification of Skills from Infancy to Adolescence – Evidence from an Accelerated Longitudinal Design

Skopek, J. & Passaretta, G.

Department of Sociology, Trinity College Dublin.

Both authors contributed equally to this work.

(skopekj@tcd.ie – passareg@tcd.ie)

Discussion Paper

Abstract

When do socially-determined gaps in cognitive achievement emerge, how large are they before children enter school, and how do they develop over schooling? Exceeding previous research, we study the evolution over time of social gaps in achievement from birth to adolescence in Germany. Our theoretical framework juxtaposes two sets of compensation and polarization mechanisms which shape the overall extent of social inequality in learning as children navigate through early years and school. For Germany, with the most stratified education system in the Western world, our theoretical considerations predict that achievement gaps will be particularly magnified during tracked secondary schooling. We exploit data from 57 competence tests taken from age of 7 months to 16 years by the National Educational Panel Study. Composite skill measures and domain-specific skills were analyzed. Our findings point to striking socio-economic status (SES) gaps in skills which emerge and expand long before children enter school but then remain astoundingly stable throughout their school careers. We tentatively conclude that schooling decreases social inequality in learning. Our study adds to an emerging body of longitudinal research concerned with the evolution of inequality of educational opportunity in the early life course.

Keywords: inequality of educational opportunity; achievement gaps; role of schooling; longitudinal research

Acknowledgements: The authors acknowledge funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 727069 (ISOTIS). We are grateful to Henrik Daae Zachrisson and Thomas Van Huizen who provided valuable comments on earlier versions of the paper.

INTRODUCTION

The association between social background and academic achievement is an essential measure of inequality of educational opportunity (Coleman 1966; Schütz, Ursprung, and Woessmann 2008). Gaps in cognitive and academic achievement by familial socioeconomic status (hereafter: SES) not only explain a large part of the association between SES and educational pathways and destinations (Blossfeld et al. 2016; Jackson 2013), but also predict SES inequality in later career outcomes over and above educational attainment (Barone and Werfhorst 2011; Hanushek and Woessmann 2008; Lleras 2008). Consequently, social inequality in cognitive achievement arising early in children's lives may underpin a whole range of stratification processes which occur later in the life course (Farkas 2003). Meanwhile, understanding how and why SES shapes children's skill formation so profoundly has become a shared scientific concern at the emerging interdisciplinary crossroads of research in human and child development, the neurosciences, and the educational and social sciences (Leopoldina 2014; Shonkoff and Phillips 2000).

More than half a century of research has established profound associations between children's cognitive and academic achievement and the SES of their family. Still, most of this knowledge is cross-sectional and leaves many questions unanswered. When and how do SES gaps in child achievement arise and evolve over the early life course? How large are these SES gaps before children enter school? Do the gaps change while children navigate through primary and secondary schooling? Longitudinal studies that trace socially-determined inequality in achievement from early childhood to adolescence are rare. The few available studies examine exclusively Anglophone countries with a high degree of institutional similarity, and often reach inconclusive findings. Our study addresses this lacuna by providing comprehensive and recent evidence on the evolution of SES gaps in cognitive and academic achievement from infancy to

adolescence for the case of Germany – the prototypical institutional contrast to the Anglophone countries analyzed elsewhere. Tying in with recent US debate about the role of schooling for social inequality in educational achievement, our study reconstructs from an institutional viewpoint how SES gaps in achievement develop as children transition from infancy and preschool to school life and, during school, from primary to secondary schooling.

Research on Achievement Gaps

A wealth of cross-sectional research demonstrates an association between family SES and student achievement. Meta-analyses report average correlations of about .3 but also reveal considerable variation depending on the skill domain, measures of SES, grade level, and study designs (Sirin 2005; White 1982). Trend studies reveal that SES gaps in achievement have been on the rise over the past 50 years in the US and in many other countries (Chmielewski 2017; Reardon 2011). Moreover, academic and public reports based on large-scale assessment studies such as PISA (age 15), PIRLS (Grade 4) or TIMSS (Grades 4 and 8) have helped cross-national research on SES gaps in math, reading, and science skills to proliferate (e.g., Bos 2007; Marks 2005; Mullis, Martin, and Loveless 2016; OECD 2016).

However, when in the life course do achievement gaps arise and how do they develop as children grow? As purely cross-sectional data provide fragmented answers at best, recent research has attempted to integrate cross-sectional test score data (such as PIRLS and PISA) into pseudo-longitudinal or pseudo-cohort designs (Dämmrich and Triventi 2018; Rözer and Werfhorst 2017; Schubert and Becker 2010). For instance, using a pseudo-cohort design involving PIRLS, TIMSS, PISA, and PIACC data, Dämmrich and Triventi (2018) found persistent SES gaps in reading and gaps in mathematics which increase slightly over children's life courses.

Longitudinal studies relying on child cohort data have facilitated our understanding of the temporal dynamics of SES gaps in children's cognitive development and achievement. In his influential analysis of the 1970 British Cohort Survey, Feinstein (2003) finds profound SES gaps in cognitive achievement for ages as early as 22 months, which continue to grow up to age 10. From an analysis of 42 US families, Hart and Risley (2003) report that by age three, upper-class children used twice as many words as lower-class children. Analyzing a sample of infant children in the US, Fernald and colleagues (2013) detected significant SES gaps in vocabulary and language processing when children were 18 months old; those gaps grew to a 6-month SES gap in processing skills critical to language development by 24 months of age.

Based on representative US data from the Kindergarten Cohort of the Early Childhood Longitudinal Study (ECLS-K), Lee and Burkam (2002) find substantial SES gaps in early reading and math ability at the transition to kindergarten. For the same data, it was later found that SES gaps in math and reading skills at kindergarten age grow further over years of elementary schooling (Cheadle 2008; Morgan, Farkas, and Hibel 2008; Potter and Roksa 2013). Farkas and Beron (2004) study age trajectories in oral vocabulary from age 3 to 13 using pooled data from the NLSY79 (US); SES gaps in vocabulary seem to increase in preschool age but remain stable when children navigate through school. Similarly, SES gaps in math are invariant during primary schooling, as Caro and colleagues (2009) establish in their analysis of the Canadian National Longitudinal Study of Children and Youth; however, these gaps widen sharply in the transition from elementary school to secondary schooling (age 11 to 15). Most notably, Bradbury and colleagues (2015) analyze the evolution of SES achievement gaps in a comparative perspective by drawing on cohort datasets from the US, Canada, the UK, and Australia. SES gaps in language and reading as observed from age 5 to 11 remained mostly stable, albeit with a tendency to grow rather than to shrink.

All in all, the available evidence suggests that SES gaps in cognitive and academic achievement are already profound before children enter school, and in some cases increase further during schooling. There are, however, several significant limitations. First, the reported studies are not all strictly comparable in terms of populations and sampling; sometimes evidence comes from non-representative or time-worn datasets. Moreover, studies differ with regard to age ranges, measures for SES, or the nature of gap measures (e.g., absolute score measures versus relative score measures). Most studies start at age five, thus neglecting SES gaps in early abilities. Furthermore, most studies focus on one or two cognitive domains only (such as reading and math). Consequently, the extent to which the evolution of SES gaps depends on the intellectual domain under investigation remains unclear; certainly, the findings from the meta-analyses mentioned above do suggest that SES gaps vary by domain. Most importantly, the longitudinal literature on achievement gaps is restricted to Anglophone countries with commonalities both in their education systems and in the overall societal context of liberal welfare states. Thus, in absence of further evidence from other national contexts, the generalizability of these findings remains unclear.

Research Objectives and Questions

This study examines the evolution of SES gaps in cognitive and academic achievement over an extended period ranging from infancy to adolescence. Our empirical investigation exploits 57 very recent child and student assessments which have been conducted by the German National Educational Panel Study (NEPS) for multiple cohort samples of children. The NEPS is the largest ever (and still active) panel study in Europe to research competence development and educational trajectories over the life course. We adopt an accelerated longitudinal cohort design by pooling multiple cohort data to reconstruct the development of cognitive and aca-

demic achievement gaps over three important institutional stages in the early educational career: infancy and preschool, primary schooling, and secondary schooling. We address four research questions: (1) When do SES gaps in cognitive achievement emerge? (2) How large are the gaps before children enter school? (3) How do these gaps develop during primary schooling? (4) How do gaps develop after children make the transition to secondary school? By incorporating composite skill measures as well as measures from multiple intellectual domains, our study can also examine the degree to which answers to these questions depend explicitly on the cognitive domain under study. Germany provides an interesting context to study the evolution of SES gaps in achievement because its tripartite educational system is known to be heavily stratifying, as it places students in different educational tracks very early.

THEORY

Socio-economic Status and Child Achievement

Family SES is a multifaceted construct shaping children's development in profound ways. In general, SES refers to a family's social position in a stratification system which includes access to financial and material resources, parents' skills and knowledge, and parents' social capital and prestige (Bradley and Corwyn 2002; Conger, Conger, and Martin 2010; Duncan and Magnuson 2003; Oakes and Rossi 2003). Widespread agreement exists as to the tripartite nature of family SES. The most important dimensions are parental education, income, wealth, and occupation, and, while all of them are interrelated, each of them is theoretically governed by distinct SES inequality mechanisms (Duncan and Magnuson 2003; Duncan, Magnuson, and Votruba-Drzal 2015).

In the literature on child development there are two prominent models which aim to explain how family SES may influence child achievement: family investment and stress reaction (Bradley and Corwyn 2002; Conger and Donnellan 2007). Models of family investment

originate in educational economics and stress the role of the resources available to families for investing in their young children's development. Accordingly, inequality in children's development results from the differences in opportunities and constraints with respect to the material, cultural, and social resources that are available to support children's skill production (Conger and Donnellan 2007). For example, higher SES families are more affluent, which enables them to provide better material resources at home, higher quality out-of-home child care, access to better schools, or private tutoring. Moreover, better educated parents are likely to have more knowledge and skills that can be transmitted to children directly, through everyday social interaction, or indirectly, through beneficial parenting practices (Ermisch 2008) or stronger involvement in their children's school life (Domina 2005; Fan and Chen 2001).

The second model, family stress, argues that the experience of economic hardship (like low income, or a high ratio of debts to assets) creates pressure and emotional distress for parents, with negative consequences on their parenting and on family functioning in general (Conger and Donnellan 2007). Unsuitable or even harmful parenting practices, in turn, have detrimental effects on children's cognitive development by contributing to their psychological maladjustment (McLoyd 1998). For example, one of the parents losing their job may destabilize household relationships and worsen the quality of the time parents devote to their children, thus undermining the latter's cognitive development. More recent research suggests that both models, educational investment and psychosocial adjustment to stress, may interact in explaining social class differences in the educational achievement of children (Layte 2017).

There are good reasons to expect SES inequality in cognitive achievement, based both on the foregoing theoretical arguments and on the abundance of studies that unequivocally establish positive associations between familial SES and a wide array of children's cognitive and non-cognitive skills. Therefore, the baseline expectation for our study is that we will find SES gaps in cognitive achievement at all ages, even before children enter school.

However, formulating unambiguous hypotheses on the evolution of achievement gaps from infancy to adolescence is more challenging, as the aforementioned theories are less concrete regarding biographical and institutional contexts. Do SES inequalities in children's cognitive and academic achievement grow as children navigate through the education system? Do patterns of SES inequalities remain persistent? Or do we even find evidence of convergence, i.e., decreasing SES inequality? To tackle these questions, in the next section we discuss two competing theoretical ideas on gap development: *polarization*, on the one hand, whereby social inequalities in achievement are intensified by the educational system, and *compensation* on the other, whereby social inequalities are mitigated through schooling.

Mechanisms of Gap Polarization

Learning may resemble a path-dependent process of cumulative advantage, if achievement gains depend on previous cognitive and academic performance (DiPrete and Eirich 2006). On the individual level, these so-called 'Matthew effects' in learning may occur as a result of the inherent self-productive logic of learning operating as a hierarchical process in which 'skill begets skill' (Heckman and Cunha 2007). For instance, the development of reading capabilities may embody a cumulative advantage process because "vocabulary knowledge substantially facilitates reading development, and reading itself is a major mechanism of vocabulary growth – which in turn will enable more efficient reading" (Stanovich 1986:380). Hence, even small initial ability differences in children from different SES groups may evolve to sizeable SES gaps in achievement as children mature and progress in school.

Ability sorting in school may further amplify pre-existing achievement gaps. Heterogeneous school and classroom environments may not only create differences in curriculum, as well as instructional quality and pace; sorting students by their ability may also enact reinforcing peer effects (van Ewijk and Sleegers 2010) or self-fulfilling prophecy effects through

teacher expectations, labeling, or teacher-student interactions (Eder 1981). For instance, skill grouping in elementary school may promote unequal achievement gains (Condrón 2008). In secondary schooling, all educational systems at some point – more or less explicitly – track or stream students to different school types, curricula, or performance classes based on their learning attainment (Blossfeld et al. 2016). The grouping and tracking of students during successive stages in the educational career may foster ‘institutional’ Matthew effects by exerting additional cumulative effects on learning, leading to increasing dispersion in achievement (Hanushek and Woessmann 2006; Kerckhoff and Glennie 1999). As a consequence, to the degree that prior achievement is associated with SES, ability-based sorting may contribute to the widening SES gaps in achievement, as higher SES children will over-proportionally reap the benefits from sorting.

In addition, family SES may determine school environments beyond children’s previous achievement. Research on the primary and secondary effects of social background (Boudon 1974) demonstrates that SES gaps in prior achievement only partially explain socially stratified transition rates and exposure to more demanding academic school tracks. Even if pre-tracking performance differences in children are taken into account, children from more advantaged families are more likely to enjoy more favorable and prestigious academic curricula (Blossfeld et al. 2016; Jackson 2013). From a rational choice perspective, such secondary effects of SES are rooted in SES differences in the parameters of the decision-making process (Breen and Goldthorpe 1997; Jackson 2013). These arguments are consistent with the idea of ‘effectively maintained inequality’, which postulates that higher SES parents will strategically exploit opportunities of differentiation – be they quantitative (e.g., length of schooling) or qualitative (e.g., quality of instruction) in nature – in education and school systems to secure advantages for their offspring (Lucas 2001).

For tracking to have consequences for SES inequality in achievement, educational segregation within and between schools must have an effect on students' learning. While proponents of tracking stress gains arising from the efficient sorting of students to tailored instruction, critics of tracking are often concerned by its potential to intensify SES gaps in achievement, either by impeding low performing children or over-proportionally benefitting high performing children (Gamoran and Mare 1989). Cross-national evidence on educational systems has disappointed optimistic accounts of tracking by showing that tracking seems to increase achievement inequality without increasing achievement efficiency (Van de Werfhorst and Mijs 2010).

The bottom line of the reproductivist argument here is that social inequality in educational opportunities tends to be reinforced by the institutional organization of education. Taken together, arguments of polarization emphasize the effects of cumulative advantage in learning and the role of schooling as an institution of social reproduction. Hence, polarization contributes to an enlargement of SES gaps in achievement throughout children's schooling careers. Moreover, polarization forces should be particularly strong once children transition to secondary school and are navigating through more differentiated and stratified school environments.

Mechanisms of Gap Compensation

However, there are contradicting views on the role of schooling in social inequality: instead of reproducing inequality, there are a number of arguments which propose that schooling may have an equalizing effect on inequality of educational opportunities. Counterfactual models of schooling suggest that schooling may act as an equalizer rather than a de-equalizer (Downey and Condron 2016; Raudenbush and Eschmann 2015); schooling is conceived of here as a dose treatment whose effects depend on the contrast in instructional quality between school

and the home. Because schools are more standardized than family home environments, inequality in instructional regimes in school are smaller than out of school (the home environment). If learning in schools is more equal than at home, low SES children's gains from schooling tend to be larger than high SES children's gains. Research on the summer/school gap supports this view by demonstrating that SES gaps in learning rise during the out-of-school season but not over the school year (Alexander, Entwisle, and Olson 2001; Downey, von Hippel, and Broh 2004). The equalizing effect of schooling may thus operate towards a compensation of SES gaps in achievement to the degree that lower SES students gain more from schooling than their higher SES counterparts. Importantly, this argument also holds for tracked schooling. Although tracking may increase heterogeneity in instruction, and lower SES children tend to be sorted to the less demanding tracks, they may still benefit relative to the counterfactual out-of-school state.

A second argument is that achievement gaps could be further compensated by ceiling effects in learning; these ceiling effects could arise psychologically from diminishing returns to learning as students reach the highest expected levels in skill attainment (Protopapas et al. 2011). In contrast to the cumulative concept of Matthew effects, 'learning plateaus' may give rise to convergence by damping the learning growth of high achievers and allowing low achievers to catch up (Baumert, Nagy, and Lehmann 2012). Hence, diminishing returns and learning plateaus may counteract a further divergence of SES gaps in achievement.

Furthermore, children's skill attainment is restricted by the limits of instruction. Hence, ceiling effects may also arise from the centralization and standardization of schooling, which equalizes learning environments. This argument underlines the importance of the national context of educational systems. For instance, a comparative study by Stevenson and Baker (1991) found that, in education systems where control over the curriculum was located at the national

level, teachers were more likely to teach uniformly, with the curricular level being less determined by students' characteristics than in education systems with more local autonomy. Thus, the institutional context of education, in particular the degree to which an educational system is standardized and centralized, seems to be relevant for the operation of ceiling effects in learning. Greater levels of centralization and standardization in school systems may limit the consequences of tracking on SES gaps in achievement (Ayalon and Gamoran 2000; Bol et al. 2014; Horn 2009; Van de Werfhorst and Mijs 2010).

Over time, mechanisms of compensation contribute to a convergence of SES gaps in achievement. Note, however, that polarization and compensation mechanisms may interact in complex ways. For instance, the equalizing effect of schooling may be offset to the degree that SES shapes students' learning gains. This would be the case if children's capacity to benefit from a certain form of instruction is a function of their attained skill level (Heckman and Cunha 2007; Raudenbush and Eschmann 2015). In an extreme scenario, higher SES children may enjoy better instruction than lower SES children both at home and at school. Such advantages may enable high SES children to benefit more from further instruction at later stages in their school careers, with the result that high SES students learn at higher rates and achievement gaps will eventually diverge over the time spent in schooling.

Summary

Our theoretical discussion suggests that the evolution of SES gaps in achievement is the result of a complex web of interacting and counteracting mechanisms. Although our study cannot test the manifold mechanisms shaping educational inequalities, the analytical framework we have developed informs our investigation, which aims to empirically determine both the magnitude and the evolution of SES gaps in cognitive and academic achievement at differ-

ent stages of children's educational careers. Furthermore, our framework underlines the importance of considering the social and institutional contexts in which mechanisms of polarization and compensation may operate. In the next step, we will address the national context of Germany's education system, since it differs in various ways to the Anglophone contexts that have been studied in the past.

THE GERMAN CONTEXT OF EDUCATION

In contrast to the 'liberal' welfare state contexts studied so far, the national setup of Germany prototypically embodies the model of 'conservative' welfare statism, characterized by a system of comparably generous welfare benefits (social insurance model), a preservation of status hierarchies, and a marked principle of subsidiarity in social policy (Arts and Gelissen 2002; Esping-Andersen 1990). Such contextual differences in social policy may be consequential for the production of educational inequality at multiple levels.

First, compared to liberal regimes such as the United States or the United Kingdom, income inequality and social disparity in living conditions are considerably lower in Germany (Smeeding and Rainwater 2002). Based on the arguments of family investment and stress reaction, less heterogeneous living conditions in families of different SES should translate into smaller SES gaps in the cognitive development of children. While lower inequality of conditions is likely to compress achievement gaps in children, it is also likely to reduce the equalizing effect of schooling, as the relative gains from schooling at the lower end decrease. Thus, in Germany, comparatively small SES gaps in achievement may coexist with comparatively weak equalizing effects of schooling.

Second, there are pronounced differences in the organization of education. Early formal childcare and preschool is semi-standardized in Germany. Subsidiarity shapes the social organization of childcare and, although rates of enrolment to formal childcare (such as crèches

or child care centers) are rising, there is no universal provision of early education and care services below age three (BMFSFJ 2015). Kindergarten (ages three to five and six) is the most significant preschool institution, with a long historical tradition and with a nearly universal attendance rate of more than 90 percent (OECD 2015). By age six children enter primary school which is – despite some regional heterogeneity from state to state – highly standardized in terms of organizational structures, curricula, and teachers’ professional training (Allmendinger 1989). This standardization may amplify mechanisms of gap compensation in primary school.

Germany’s education system is known as one of the world’s most stratifying sorting machines, adopting as it does an early tracking model in secondary education (Allmendinger 1989; Blossfeld et al. 2016; Bol and van de Werfhorst 2011). Nearly all of the 16 German states track students after Grade 4 at ages 10 and 11. School tracks – typically academic track (Gymnasium), poly-technical track (Realschule), and lower vocational track (Hauptschule) – embody different curricula, and by and large different school types. Despite complex regional heterogeneity in the organization and structure of tracking (Helbig and Nikolai 2015), the tripartite system is deeply rooted in the socio-historical context of Germany (Benavot and Resnik 2004). A common element across German states is the primary school recommendation (Gresch, Baumert, and Maaz 2010; Neugebauer 2010): at the end of primary school, teachers provide a formal recommendation for students stating the secondary school track most appropriate to their abilities (usually marks in math and German), behavior, and talent. Specifically, the recommendation indicates children’s eligibility for the academic track. The regulating power of the recommendation and its consequences, however, varies by federal state, from being just a suggestion to the parents to having legally binding character (Helbig and Nikolai 2015).

Previous research on Germany seems to be quite conclusive about the polarizing effects of the tracking system. Access to the academic track is highly socially selective, which is only

partly explained by SES differences in students' prior achievement (Buchholz et al. 2016; Neugebauer et al. 2013; Stocké 2007). Track mobility during secondary education years is rare, mostly downward directed, and socially selective even after accounting for students' achievements (Buchholz et al. 2016). Thus, track mobility does not reduce but rather amplifies SES disparities in track placement (Ditton 2013).

When applying our theoretical arguments in the German context, we conclude that the reproduction mechanisms of gap polarization should be particularly pronounced, and may in fact be the dominant force shaping social inequality in educational achievement. Therefore, as a baseline hypothesis for our study, we expect to find rising SES gaps in achievement over children's school careers, and particularly after they transfer to the highly stratified stage of secondary education in Germany.

DATA AND METHODS

Data

Our empirical analysis draws on data from the German National Education Panel Study (NEPS, see Blossfeld, Roßbach, and von Maurice 2011). NEPS data is extraordinarily well-suited to our purposes, and represents a unique and unprecedented source of information on competence development from infancy to adulthood in a European country.

Implementing a multi-cohort sequence design, NEPS collects longitudinal data on children and adults from a total of six cohort samples. We used data from three starting cohorts (SC): the *Newborns Cohort (SC1)*, the *Kindergarten Cohort (SC2)*, and the *Grade 5 cohort (SC3)*. SC1 covers a nationally representative sample of children (N=3,481) who were born in the first half of 2012 (Würbach, Zinn, and Aßmann 2016). Children are assessed for the first time at an age of about 7 months. Both SC2 and SC3 adopted a two-stage sampling approach with schools/institutions being sampled in the first stage. SC2 covers a representative sample

(N=2,996) of children attending kindergarten in 2010/2011 and entering primary school in 2012/2013, and SC3 a representative sample (N=6,112) of children attending schools in Grade 5 (for most German states the first level of secondary schooling) in the school year 2010/11 (Steinhauer et al. 2015). Both SC2 and SC3 cohort panels were supplemented by refreshment samples in Wave 3 (N=6,341 for SC2 and N=2,205 for SC3).

An Accelerated Longitudinal Design

In order to investigate achievement gaps from infancy to adolescence we apply an accelerated longitudinal cohort design (see Figure B1 in Appendix B), which involves linking adjacent data from the three NEPS cohorts to reconstruct a synthetic early life span that ranges from 7 months of age to the end of compulsory schooling (age 15 and a half). More precisely, through weighting (see below), we created a synthetic pseudo-cohort of children for which we studied SES gaps in cognitive achievement throughout the stages of *infancy and preschool age* (age 7 months to 6 years), *primary schooling* (age 7 to 10), and *lower secondary schooling* (age 11 to 15 and a half).

Table 1 provides an overview to the study material (cohorts, waves, sample size, and test information) that entered our design. Cognitive tests are partly hierarchically classified into *domain-general cognitive functioning*, a set of *domain-specific skills* such as language (listening comprehension and reading), science, or math, and, finally, *stage-specific skills* (such as orthography at the end of primary and during secondary schooling). A more detailed account of the competence assessments and the correlation matrices among all tests available in each life stage are given in Tables A1, A2, A3, and A4 in Appendix A. For all tests, we used global scores as provided by NEPS (Pohl and Carstensen 2012). Note that we did not include Wave 2 of SC1 and Wave 4 of SC3 because no appropriate test data was available. In total, we drew upon 15 waves (4 for SC1, 6 for SC2, and 5 for SC3) and 57 child competence assessments.

Our study included overall 54,526 observations obtained from 16,512 children. The specific sample size for the analyses of the composite measures and the single competence domains are reported in Table A5 in Appendix A.

– TABLE 1 HERE –

Measurement of Skill Stratification

Studying the evolution of achievement gaps over an extended and heterogeneous early life span creates substantial challenges in measurement. One of the most obvious aspects is rooted in the hierarchical nature of development itself: children master different skills at different times. For instance, attaining math skills in school requires the ability to read (e.g., a textbook), and learning to read requires the ability to comprehend words and sentences, an ability to be mastered before school. Developmental changes in the early life span are particularly rapid and complex, and as children’s minds evolve the quality of their skills changes (Zeanah, Boris, and Larrieu 1997). Consequently, qualitative change in psychological development renders a consistent measurement of skills on absolute scales infeasible (Feinstein 2003). A pragmatic challenge results from the fact that NEPS tests are comparable over time only for selected domains (math and reading), only at school age, and only within cohorts (Fischer et al. 2016).

Our primary measurement strategy therefore relies on the construction of *relative test score differences* rather than absolute test score differences. Relative measures express test score gaps in relation to the variation in overall test scores, and are useful for measuring inequality in the distribution of achievement across groups (Reardon 2008). For the purpose of our study, a relative approach is particularly appropriate, as we are interested in analyzing the

evolution over time of the SES stratification of skills. In a relative approach, test score distributions are standardized to have a mean of 0 and a uniform standard deviation (SD) of 1 (subtract from each score the mean score and divide the result by the standard deviation of the scores). Hence, standardized scores (z-scores) are easy to interpret and, in relative terms, comparable across different domains. For example, a standardized math score of 0.5 represents a math test score that is 50% of a SD larger than the average math score. For approximately normal distributed test scores, a child with a z-score of 0.5 would be ‘smarter’ in math than 69% of all children. If, say, children from high SES families averaged 0.5 SD and children from low SES -0.5 SD, the achievement gap would amount to 1 SD, implying that an average high SES child scores better than 84% of all low SES children. Relative measures have crucial advantages in the context of our study but also come with some significant limitations, which we will discuss in more detail later. To compensate for some of the limitations, we also inspected absolute score gaps for selected competence domains with a longitudinally coherent scale (math, science, and reading).

We constructed *two types* of relative score outcomes. First, we generated a *composite* measure by combining all test score data available in a particular life stage. As it exploits all 57 assessments, the composite measure summarizes a child’s position in the multivariate distribution of skills at each life stage, ranging from 7 months until age 15 and a half. For our purposes, this represents a major benefit as it provides a comprehensive portrait of the gradual evolution of skill stratification in childhood. Moreover, by focusing on children’s relative positions in a series of age appropriate tests, the composite measure addresses the problem of qualitative change in skills.

Second, we also analyzed z-scores for single domains, as our composite measure may mask important heterogeneities in gap size and the evolution of the gaps for specific competence domains. We did so for six types of abilities: vocabulary¹, reading, orthography, math,

science, and basic cognitive skills. The six domains form an approximation of the skill set necessary for educational success, labor market prospects, and participation in society. Moreover, these are the only domains for which we had at least four measurements in the observation window. Appendix C1 provides more details on the six ability domains.

Note that our analyses of single domains are, depending on the domain, restricted in terms of measurement points and overall observation window. Altogether, the combination of a composite measure with single domain measures offers a complementary portrait of the evolution of those disparities in children's positions within stage-specific skill distributions which can be attributed to differences in SES.

Construction of Scores

We construct the composite score in four steps. First, we standardize all test scores within each wave. Second, we take the average across z-scores from different tests for each wave. Third, we purge the average z-scores from age variation by calculating z-score residuals unrelated to age (for a similar procedure see Bradbury et al. 2015). Fourth, we standardize the residuals. Steps one, three (using the only z-score available), and four were equivalently applied to the single domain measures. Appendix C2 provides a detailed account of the procedures involved.

Limitations

Using z-scores as measures of inequality in achievement has many advantages for our purposes. Nevertheless, certain caveats and limitations should be kept in mind and, although these are discussed extensively in the achievement gap literature (e.g., Bradbury et al. 2015; Reardon 2008), we want to highlight some issues that are particularly germane to our case.

First and most importantly, relative and absolute skill levels should not be confused. Based on z-scores alone we cannot say anything about the evolution of absolute skill levels,

nor about the difference in absolute skills between SES groups. For example, rising SES inequality in math as expressed by widening absolute scores may occur even though the SES differences in math z-scores remain constant. Moreover, z-scores are a measure of inequality; they are a function of the overall score heterogeneity (expressed by the standard deviation). Therefore, it is also possible, for example, that SES gaps in z-scores will increase over time as a result of increasing variance in absolute skill levels, although SES gaps in absolute skill levels (e.g., SES specific proficiency levels in math) remain constant. Based on these considerations, we conduct some additional analyses using linked scores provided by NEPS (math, science, and reading), which turned out to be very consistent with our main results (see Figures B2 in Appendix B).

Second, like absolute scores, z-scores hinge on the assumption of perfectly interval scaled test score variables. This assumption is relaxed by metric-free measures that have purely ordinal features such as percentile ranks or relative distribution measures (Reardon 2008) that, nonetheless, make less efficient use of the data if it is actually interval scaled. In our case, sensitivity analyses on the composite measure using percentile ranks yielded very similar results (see Figure B3 in Appendix B).

Third, z-score levels for SES groups are sensitive to the marginal distribution of those SES groups, although of no consequence for between-group differences in the average z-score (the real quantity of interest in our context). Intuitively, this occurs because the average z-score within groups is influenced by the relative groups' size, an issue not explicitly discussed in the existing literature (further information is available from the authors upon request). In order to overcome this issue, we standardized marginal distributions across the five different sampling points in our data as part of a weighting strategy (see below).

Measurement of Family SES

For the purpose of our study we take parental education as an indicator for family SES for various theoretical and practical reasons. First, parental education is more stable over the life course, less subject to fluctuations, and less prone to measurement error than parental occupation, earnings, and family income. Second, education is an antecedent variable and reliable predictor of occupational outcomes and income. Third, parental education turned out to be a very important predictor of early cognitive development and, therefore, closely tied to the process under study, i.e. educational inequality. Fourth, by using parental education, our findings can be linked to previous stratification research on educational achievement and attainment, which routinely relies on parental education as an SES indicator (e.g., Blossfeld et al. 2016; Bradbury et al. 2015; Buchmann 2002; Jackson 2013).

Using data from interviews with parents, we construct parental education by averaging the number of years spent in education by the children's parents. Averaging ensures that both parents' educational levels are considered for measuring SES. In case of single parents, we only took the educational level of the parent respondent (mother or father). Note that, in our case, years of education did not reflect the total years of schooling and training parents had experienced, but the years of education that are necessary to attain a certain educational certificate (the information is provided in a variable generated by NEPS, see Skopek and Munz 2016). Years of education was coded 8 (basic schooling without formal qualifications), 9 (lower secondary degree), 10 (intermediate secondary degree), 12 (lower secondary and vocational training), 13 (intermediate secondary degree and vocational training, or, upper secondary degree), 15 (upper secondary and vocational training), 16 (lower tertiary education), and 18 (higher tertiary, university level education). Henceforth, we use parental education and SES synonymously. Sensitivity analyses using a threefold classification of parental education based

on the most-educated parent led to very similar results (available from the authors upon request).

All our models controlled for migration background. We constructed a dummy variable that distinguishes children with no migration background (children born in Germany whose parents were also born in Germany) from children with migration background (first- or second-generation migrants). Thus, our findings pertain to SES gaps net of migration-related differences.

Weighting and Adjustments

In our analyses we apply weighting to address three issues related to the accelerated longitudinal design. We give a non-technical description here but describe all related procedures in detail in Appendix C3.

First, we must consider the specific sampling design of the NEPS cohorts. Therefore, we used design weights as provided by NEPS that adjust for non-randomness in the selection of children and students (Steinhauer et al. 2015; Würbach et al. 2016). These initial weights ensure that the data in the respective panel samples is representative.

Second, marginal distributions of parental education, migration background, and sex varied across the five initial samples (3 cohorts plus the 2 refreshment samples). Marginal distributions (e.g. the proportions of students with high SES parents), however, affect the z-score levels (e.g., the larger the proportion of high SES students, the closer their z-score will get to 0, the average). To avoid variation in marginal distributions affecting our estimates, we link the initial samples by additionally constructing weights that standardize the marginal distributions of SES, migration background, and sex across all the five sampling waves. That way, we construct a pseudo-cohort that has constant marginal distributions over time and for which z-score levels of SES groups are perfectly comparable.

Third, not all children were successfully followed up on over all waves, thus leading to unbalanced panel data. Hence, we need to consider selective panel attrition, as it could induce bias into our estimation of the evolution of achievement gaps (for instance, if poorly performing children tend to drop out). We use inverse probability weighting to account for panel attrition; to put in concisely, we predict for each wave the probability of participating in the next wave based on all our analytical variables, including the outcome (achievement scores) interacted with SES. Design weights, cohort standardization weights, and the inversed conditional probabilities are finally combined to give longitudinal weights.

Estimation Methods

We use linear regression models to estimate the association of SES and achievement net of migration background. We run regressions separately for the seven outcomes (composite score and the six single domain scores) and for each age point t in our data design. Regression models were specified as:

$$y_{i|t} = \alpha_t + b_{1t}P_i + b_{2t}M_i + e_{i|t} \quad (1)$$

where the score outcome at age t of student i is a function of parental education (P), migration background (M), and an error term e . Parameter b_{1t} measures the association between parental education and standardized achievement scores at age t net of migration. Due to the large number of estimates for each age and score, we present estimation results in visual plots in addition to tabular form. These plots illustrate achievement gaps between children with ‘low’, ‘medium’ and ‘high’ SES backgrounds, defined by 12, 14, and 16 years of average parental education. We deliberately do not present graphs showing extreme group comparisons,

i.e. 18 versus 8 years of parental education, although these can be easily obtained by multiplying b_{1t} by 10. Rather, the chosen cut-offs represent typical values. About 20 per cent of children in our sample have parents with 12 years of education or less, another 20 per cent have parents with 16 or more years of education, and roughly 60 per cent have parents with more than 12 but less than 16 years of education.

In a second step, we estimate a time-trend parameter for the evolution of achievement gaps within each of the three institutional periods considered: preschool, primary, and secondary education. Trend-parameters could be estimated only for the composite measure, since it is the only outcome for which we have a minimum of three observation points within each institutional period. These models reflect in statistical ways the central ideas of our theoretical discussion, as they allow us to assess more formally whether SES gaps in achievement evolve differently in the preschool, primary, and lower secondary education periods. To this end, we pool data on the composite score and create additional time (age) variables (age equals 0 in the beginning of each institutional period). The three models, one for each institutional period, are specified as:

$$y_{it|p} = \alpha_p + b_{1p}P_i + b_{2p}A_{it} + b_{3p}M_i + b_{4p}P_i \cdot A_{it} + b_{5p}M_i \cdot A_{it} + e_{it|p} \quad (2)$$

where the score outcome y in the institutional period p for the individual i is regressed on parental education (P), the mean age at test day at time t (A), migration background (M), and the two-term interactions between parental education and migration background with age, respectively. In this setting, the three coefficients b_{4p} inform us whether a) linear rates of change of SES disparities are statistically significant within the institutional periods and b) whether these rates of change differ in preschool, primary, and secondary education. We apply cluster-

robust estimation to adjust standard errors for similarities in repeated observations within children.

RESULTS

Composite Measure

We begin by reporting findings on the evolution of achievement gaps using the composite score. Parameter estimates for Equation 1 are presented in Table 2 (column ‘Composite’). By plotting predictions for three SES groups (low, medium, and high), Figure 1 illustrates the evolution of SES gaps in cognitive and academic achievement from infancy to adolescence. Dashed lines were added to emphasize the temporal dynamics of achievement gaps, from preschool and throughout the school period. When interpreting those dynamics, one should recall that there is only a single assessment entering the composite index at 7 months of age (sensorimotor skills) and age 4 (mathematical competence), while all other measurement occasions involve multiple measures.

The plot reveals that SES gaps in achievement emerge very early in childhood. SES gaps in skills are already visible 7 months after birth, albeit they are relatively small (.16 SD between ‘high’ and ‘low’ or an estimated .4 SD for the maximum SES gap). During preschool years this gap grows to about 50 per cent of a SD (or 1.2-1.3 SD for the maximum SES gap) as children from lower SES families increasingly lose ground relative to children from higher SES families. For primary school years, we notice larger gaps but also more fluctuating patterns: although the gap appears most extreme in the first year of primary schooling, we interpret this outlying value with some caution, as it might be driven by the specific combination of tests (test correlations are high for this sweep). Towards end of primary schooling and shortly before children transition to secondary schooling, the gap in achievement between children of the most and the least educated parents grew to about 65 per cent of a SD (or 1.6 SD between the

extreme groups respectively). For secondary school years, the gap first shrinks slightly, to 56-60 per cent of a SD, only to increase further to .8 SD at age 15 and about .7 SD at age 15.4 in Grade 9.

Despite some fluctuating patterns, the analysis based on the composite score suggests increasing SES inequality: at 7 months of age the high-low gap amounts to .16 SD; from age 2 to 6 to an average of .47 SD; during primary school years (age 7 to 10) to an average of .64 SD; and during lower secondary school years (age 11 to 15-15.4) to an average of .68 SD. It is worth noting that the achievement gap is substantial even before children enter school (age 5-6) since its magnitude corresponds to more than two thirds of the achievement gap found at the end of lower secondary (age 15-15.4).

Table 3 reports selected results from the pooled models (Equation 2) estimating linear trends in the evolution of SES gaps within each of the three institutional periods: preschool, primary, and lower secondary education. Full estimates can be found in Appendix A (Table A6). Table 3 confirms the overall increase of SES inequalities moving from preschool to the school years (the main effect of parental education increases moving from preschool to schooling), and offers a valuable synthesis of the evolution of SES gaps within the three periods. Coefficients of the interaction terms between parental education and age indicate clearly that SES gaps increase from 7 months of age until the end of kindergarten (age 5-6), remain remarkably stable during primary education, and tend to increase only slightly over secondary schooling. While being both statistically significant, the linear rate of growth is almost doubled in preschool compared to lower secondary education (0.015 vs 0.08, respectively, $p = 0.042$, one-sided test). Moreover, the linear rate of growth in secondary education is relatively small in magnitude: the high-low gap (4 years of parental education) grows only by about 0.15 SD over the 4.5 years of secondary schooling. Therefore, the findings do not lend overwhelming

support to the hypothesis that achievement gaps should widen sharply after children enter the highly stratified and tracked system of secondary education in Germany.

– FIGURE 1 HERE –

– TABLE 2 HERE –

– TABLE 3 HERE –

Selected Domains

Table 2 also includes estimates of the SES coefficient for the single domains available. In addition, Figure 2 illustrates the evolution over time of SES-determined inequality in the selected domains for the three SES levels. The general findings obtained from the composite score analysis seem to hold also for qualitatively different domains. For mathematical competence, we observe a rise in the average high-low gap from preschool (.44 SD) to primary school (.63 SD) and secondary school (.64 SD), which parallels science (.40 SD, .60 SD, and .60 SD) and vocabulary (.41 SD, .58 SD, .58 SD). The same pattern of increasing inequality from preschool to lower secondary education can be found looking at basic cognitive skills (.16 SD, .28 SD, .34 SD), although the magnitude of the SES gaps here is significantly lower compared to the other competence domains (approximately halved). For reading and orthography, average gaps seem to increase moving from primary to secondary schooling (from .52 SD to .61 SD and from .44 SD to .57 SD respectively), but, unfortunately, we cannot evaluate the evolution of gaps from preschool to primary education due to lack of measures in the early years. Overall, in Grade 9, SES inequality in ability is similar for all domains related to learning – ranging from .16 SD to .18 SD differences per year of parental education – while being significantly lower for basic cognitive skills (0.11 SD per year of education).

We also conduct additional analyses that examine the evolution of achievement gaps in absolute terms (Figures B2 in Appendix B). While allowing us to assess conclusions which differ completely depending on whether we are looking at relative or absolute SES gaps, the analysis of absolute scores was only feasible for selected competence domains and starting cohorts (math and science in SC2, reading in SC3). Overall, findings in absolute terms parallel findings in relative terms. SES gaps in math and science seem to increase slightly from pre-school to the end of primary education, while, in the case of reading, they are quite constant throughout lower secondary education (the increase in SES-determined gaps in reading only pertains to primary education, see figure 2).

Figure B2 (see Appendix B) may erroneously suggest a limited increase in absolute skill gaps in math and science. To illustrate the relevance in the absolute size of gaps, we can express them as time-lags in learning. Estimating the linear rate of growth within the period, we find that the gap between high and low SES children (4 years of parental education) increases, from preschool towards the end of primary education, from a 5 to an 8.5-month time-lag in learning for math, and for science from a 7 to a 9.5-month time-lag. These striking increases in learning gaps, from roughly 55 to 94 percent and from 77 to 105 percent of a school year (a school year in Germany has about 9 months of instruction), underline the substantive significance of the initial gaps and their increase over children's school careers.

– FIGURE 2 HERE –

DISCUSSION

Taken together, our results indicate that the inequality attributable to socio-economic status in children's cognitive and academic achievement immediately begins to grow in the first year of life, and continues to do so, albeit more slowly, right until the end of compulsory

schooling. Consistent with theories of family investment, which suggest stratified developmental conditions for children, we found that cognitive gaps arise rapidly in the early years before children enter school. Once children enter primary school, however, we found SES inequality in overall achievement to remain stable, rather than increasing substantially.

But what does this empirical evidence say about the overall role of schooling for the reproduction of SES inequality? Does schooling increase or reduce social inequality? Unfortunately, our design cannot provide direct and conclusive evidence on the matter, as we lack a suitable counterfactual scenario in which to observe how inequality might evolve beyond the preschool years in a world without – or more realistically with less – instructional time spent in school. Still – based on key assumptions – our results may offer some insights in this recent debate.

Assuming that SES inequality were to keep increasing for children of school age at the same rate as during the preschool period if they were educated at home rather than in school (and in absence of cohort effects), the dual findings that gaps grow in preschool and remain constant in primary school lend support to the idea that school acts as an equalizer, effectively impeding additional gap increases that might have happened in absence of schooling. Figure 3 presents a data-fed illustration of the argument by comparing observed trends in high-low SES gaps (linear interpolation based on pooled models, Table 3) with trends extrapolated from gap growth in the preschool phase (linear extrapolation based on the preschool model, Table 3). Although SES gaps at the end of primary school are substantial – both in relative and absolute terms –, it appears likely that schooling contributes to a compression of inequality.

Furthermore, theoretical arguments regarding sorting and tracking predict a marked increase in achievement gaps once children transition from the more comprehensive primary to the tracked secondary school system. Our findings shed more light on this hypothesis: on the one hand, we find that gaps do indeed increase during secondary schooling. On the other hand,

this increase was markedly less dramatic than one might expect based on the social reproduction argument prevailing in stratification literature; in fact, it turned out that gap increases in secondary schooling occur, if at all, at modest rates. In light of the indisputable truth that Germany's education system resembles a huge early-stratifying sorting machine, this finding is astounding. As the conceptual Figure 3 suggests, schooling (even tracked schooling) acts towards compensating overall levels of inequality. Of course, that argument does not imply that tracking does not foster inequality in educational achievement. Comparative research on educational systems has shown that more school tracking compared to less school tracking is associated with greater inequality in educational achievement (Hanushek and Woessmann 2006; Van de Werfhorst and Mijs 2010). Based on our findings, we argue, however, that even though tracking might be less effective in compensating educational inequality, tracked schooling nevertheless has a compensatory effect.

While there is no need to discuss the multiple theoretical and practical caveats that may invalidate the assumptions underlying these interpretations, it is noteworthy that our conclusions are consistent with findings from the US, both on the summer/school gap in learning and on the consequences of preschool expansion (for a review see Raudenbush and Eschmann, 2015).²

– FIGURE 3 HERE –

CONCLUSIONS

When do SES gaps in cognitive achievement emerge, how large are the gaps before children enter school, and how do these gaps develop during the course of primary and secondary schooling? Investigating these questions for the case of Germany by employing the most recent, representative, and extraordinarily rich competence data, our paper contributes to

the hitherto fragmented and Anglo-centric empirical evidence on the development of educational inequality from infancy to adolescence. For Germany – with one of the most infamously stratifying education systems in the Western world – we can answer those questions as follows: socially-determined gaps in early cognitive achievement arise as early as 7 months after birth, translate into remarkable achievement gaps even before children enter school, and remain – notwithstanding slight increases – fairly stable as children navigate through school.

Although our research design is genuinely limited to assessing the overall role of schooling for social inequality, our longitudinal findings on achievement gaps carry valuable insights to this fundamental debate. In line with older and more recent arguments in the US debate (Coleman 1966; Downey and Condron 2016), our findings invite more critical reflection on the dominant view of schooling as a locus of reproduction, that is, an institution that exacerbates inequality in educational outcomes by social background. Our study inspecting actual educational outcomes in terms of educational competencies and skills does not yield overwhelming support for such reproductivist ideas, even in what can be considered one of most stratifying education systems in the Western world.

Our study also contributes an analytical framework that may inform future research on achievement gaps. The longitudinal study of inequality *outcomes* we have proposed represents an important addition to research that studies the potentially counteracting *mechanisms* of polarization and compensation which produce those outcomes. Our systematic review of the mechanisms related to the family environment, the school system, and the natural process of learning may help future studies to frame their contributions and to reflect on the relevance of the specific mechanisms involved for the overall outcomes of inequality. Moreover, our analytical framework stresses the role of institutional characteristics that may affect specific mechanisms of inequality and, in turn, the overall extent and shape over time of inequality at the group level. While the analysis of achievement gaps in Germany may be seen as a complement

to studies in the US and other liberal countries, sociological science would benefit from extending the scope to a variety of education systems.

NOTES

¹ Vocabulary scores assessing ‘listening comprehension at the word level’ are only provided by the NEPS in the first year of kindergarten (age 5), grade 1 (age 6.9), grade 3 (age 8.7) and grade 6 (age 11.9). We complemented the analyses of these four observation points using scores assessing ‘listening comprehension at the text level’ (listening) in the end of grade 9 (age 15.4).

² Note that this recent literature is not free from key assumptions (for a discussion see Downey and Condon 2016).

TABLES AND FIGURES

Figure 1. Composite scores for children of parents with low, medium, and high educational levels.

Notes: Predictions based on life stage specific regression models (see Equation 1). 95% confidence intervals for predictions shown. Stage: K = Kindergarten, G = Grade level in school. Long-dashed black lines connect data within the same NEPS cohort.

Figure 2. Domain specific z-scores for the children of parents with low, medium, and high educational levels.

Notes: Predictions based on life stage specific regression models (see Equation 1). 95% confidence intervals for predictions shown.

Table 1. Overview of NEPS cohorts, waves, samples and assessment data

	Wave	Sample	Test date	Life-course		Assessment data ^c								Stage-specific tests
						LC			R					
Starting cohort	#	N ^a	(mean)	Stage	Age ^b	CB	G	V	RC	RS	S	M		
SC1: <i>Newborns</i>	1	2,944	10/2012	Infancy	0.6									Sensorimotor development
	2	—							—					
	3	2,085	6/2014		2.2		X	X						Morphology
	4	1,813	6/2015		3.2			X						Executive control, Phonological working memory
	5	2,024	6/2016		4.2							X		
SC2: <i>Kindergarten</i>	1	2,515	3/2011	Preschool	1	5.0		X	X			X		
	2	2,341	2/2012		2	5.9	X					X		Knowledge of letters, Phonological working memory and awareness
+ refreshment sample	3	6,027	3/2013	Primary	1	7.0		X	X			X	X	
	4	5,717	11/2013	school	2	7.7	X			X	X		X	
	5	5,145	11/2014	(Grade)	3	8.7			X			X		
	6	6,191	11/2015		4	9.7				X			X	Orthography
SC3: <i>Grade 5</i>	1	3,732	11/2010	Secondary	5	10.9	X			X	X		X	Orthography
	2	3,514	11/2011	school	6	11.9			X			X		
+ refreshment sample	3	4,019	11/2012	(Grade)	7	12.9				X			X	Orthography
	4	—			8					—				
	5	3,317	11/2014	(Autumn)	9	14.9				X		X	X	Orthography
	6	3,142	4/2015	(Spring)	9	15.4	X	X		X				

Notes: ^a Number of children in the overall analytical sample (see table A5 in Appendix A for the sample size related to each outcome). ^b Mean age at (first) test day. ^c Five types of assessments. (1) Domain general cognitive functioning: CB = cognitive (non-verbal) basic skills (two tests: perceptual speed and reasoning), (2) Language: LC = listening comprehension comprising G = grammar (LC at sentence/text level), V = vocabulary (LC at word level), R = reading comprising RC = reading competence, RS = reading speed, (3) Science: S = scientific literacy, (4) Math: M = mathematical competence, (5) stage-specific tests. For more details on tests see Table A1 in appendix A.

Table 2. Regression results by outcome and time point

Time		Outcomes (z-scores)														
Age	Stage	Composite		Cog. Basic		Math		Ortho		Read		Science		Vocab		
0.6	<i>Infancy and preschool</i>	0.04	(0.01)	—		—		—		—		—		—		
2.2		0.11	(0.02)	—		—		—		—		—		0.10	(0.01)	
3.2		0.10	(0.02)	—		—		—		—		—		0.10	(0.02)	
4.2		0.13	(0.02)	—		0.12	(0.02)	—		—		—		—		
5.0	K1	0.13	(0.01)	—		—		—		—		0.10	(0.01)	0.11	(0.01)	
5.9	K2	0.12	(0.01)	0.04	(0.01)	0.10	(0.01)	—		—		—		—		
7.0	<i>Primary school</i>	G1	0.19	(0.01)	—		0.15	(0.01)	—		—		0.15	(0.02)	0.15	(0.01)
7.7		G2	0.13	(0.02)	0.07	(0.01)	0.16	(0.01)	—		0.10	(0.02)	—		—	
8.7		G3	0.16	(0.01)	—		—		—		—		0.15	(0.01)	0.14	(0.01)
9.7		G4	0.16	(0.02)	—		0.16	(0.01)	0.11	(0.02)	0.16	(0.01)	—		—	
10.9	<i>Secondary school</i>	G5	0.14	(0.01)	0.06	(0.01)	0.14	(0.01)	0.11	(0.01)	0.14	(0.01)	—		—	
11.9		G6	0.15	(0.01)	—		—		—		—		0.14	(0.01)	0.11	(0.01)
12.9		G7	0.19	(0.01)	—		0.17	(0.01)	0.16	(0.01)	0.16	(0.01)	—		—	
14.9		G9	0.20	(0.01)	—		0.17	(0.01)	0.16	(0.01)	—		0.16	(0.01)	—	
15.4			0.17	(0.01)	0.11	(0.01)	—		—		0.16	(0.01)	—		0.18	(0.02)

Notes: Regression coefficients for parental education (see Equation 1). Standard errors in parentheses. All estimations adjusted for migration background.

Outcomes: Composite score, cognitive basic skills, mathematical competence, orthography, reading competence, scientific literacy, vocabulary.

Table 3. Parametric models for linear time trends in SES gaps in preschool, primary, and lower secondary education.

	Preschool	Primary	Secondary
<i>Parental education</i>	0.061*** (0.014)	0.163*** (0.012)	0.153*** (0.008)
<i>Age</i>	-0.196*** (0.053)	0.008 (0.076)	-0.118*** (0.035)
<i>Parental education*Age</i>	0.015*** (0.003)	-0.002 (0.005)	0.008** (0.002)
Constant	-0.863***	-2.207***	-2.112***
N	13092	20082	17279

Notes: Three regression models pooling all observations on composite scores within the three institutional periods: preschool, primary and lower secondary education. Time axis (age) as linear variable (age=0 in the beginning of each institutional period). All models adjusted for migration background and interaction with time axis. Standard errors clustered by individual. Weighted data. Standard errors in parentheses.

Significance (two-sided test): * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Figure 3. Conceptual Graph

APPENDIX A. TABLES

Table A1. List of the competence assessments used in our study with details on cohort, institutional period, wave, grade, assessment method, and score type.

	Start- ing cohort	Institu- tional period	Wave	Grade	Assessment methods	Score type
<i>7 months</i>						
Sensorimotor development	1	Infancy	1	–	Observer rating	WLE
<i>2.2 years</i>						
Active vocabulary	1	Infancy	3	–	Parents report*	–
Syntax	1	Infancy	3	–	Parents report*	–
Morphology	1	Infancy	3	–	Parents report*	–
<i>3.2 months</i>						
Vocabulary	1	Infancy	4	–	Computer-based	SUM
Executive control	1	Infancy	4	–	Computer-based	SUM
Phonological working memory	1	Infancy	4	–	Computer-based	SUM
<i>4.2 months</i>						
Mathematics	1	Infancy	5	–	Computer-based	WLE
<i>5 years</i>						
Vocabulary	2	Preschool	1	–	Paper-based	SUM
Grammar	2	Preschool	1	–	Paper-based	SUM
Scientific literacy	2	Preschool	1	–	Paper-based	WLE
<i>5.9 years</i>						
Cognitive basics skills 1	2	Preschool	2	–	Paper-based	SUM
Cognitive basics skills 2	2	Preschool	2	–	Paper-based	SUM
Mathematics	2	Preschool	2	–	Paper-based	WLE
Early knowledge of letters	2	Preschool	2	–	Paper-based	SUM
Phonolog. work. memory 1	2	Preschool	2	–	Paper-based	SUM
Phonolog. work. memory 2	2	Preschool	2	–	Paper-based	SUM
Phonolog. awareness 1	2	Preschool	2	–	Paper-based	SUM
Phonolog. awareness 2	2	Preschool	2	–	Paper-based	SUM
Phonolog. awareness 3	2	Preschool	2	–	Paper-based	SUM
<i>7 years</i>						
Vocabulary	2	Primary	3	1	Paper-based	SUM
Grammar	2	Primary	3	1	Paper-based	WLE
Scientific competence	2	Primary	3	1	Paper-based	WLE
Mathematics	2	Primary	3	1	Paper-based	WLE
<i>7.7 years</i>						
Cognitive basics skills 1	2	Primary	4	2	Paper-based	SUM
Cognitive basics skills 2	2	Primary	4	2	Paper-based	SUM
Mathematics	2	Primary	4	2	Paper-based	WLE
Early reading	2	Primary	4	2	Paper-based	SUM
Reading speed	2	Primary	4	2	Paper-based	SUM
<i>8.7 years</i>						
Vocabulary	2	Primary	5	3	Paper-based	SUM
Scientific competence	2	Primary	5	3	Paper-based	WLE
<i>9.7 years</i>						
Early reading	2	Primary	6	4	Paper-based	SUM
Mathematics	2	Primary	6	4	Paper-based	WLE
Orthography 1	2	Primary	6	4	Paper-based	WLE
Orthography 2	2	Primary	6	4	Paper-based	WLE
<i>10.9 years</i>						
Cognitive basics skills 1	3	Secondary	1	5	Paper-based	SUM
Cognitive basics skills 2	3	Secondary	1	5	Paper-based	SUM
Reading speed	3	Secondary	1	5	Paper-based	SUM
Reading competence	3	Secondary	1	5	Paper-based	WLE

Mathematics	3	Secondary	1	5	Paper-based	WLE
Orthography 1	3	Secondary	1	5	Paper-based	WLE
Orthography 2	3	Secondary	1	5	Paper-based	WLE
<i>11.9 years</i>						
Vocabulary	3	Secondary	2	6	Paper-based	SUM
Scientific competence	3	Secondary	2	6	Paper-based	WLE
<i>12.9 years</i>						
Reading competence	3	Secondary	3	7	Paper-based	WLE
Mathematics	3	Secondary	3	7	Paper-based	WLE
Orthography 1	3	Secondary	3	7	Paper-based	WLE
Orthography 2	3	Secondary	3	7	Paper-based	WLE
<i>14.9 years</i>						
Reading speed	3	Secondary	5	9–fall	Paper-based	SUM
Mathematics	3	Secondary	5	9–fall	Paper-based	WLE
Scientific competence	3	Secondary	5	9–fall	Paper-based	WLE
Orthography 1	3	Secondary	5	9–fall	Paper-based	WLE
Orthography 2	3	Secondary	5	9–fall	Paper-based	WLE
<i>15.4 years</i>						
Cognitive basics skills 1	3	Secondary	6	9–spring	Paper-based	SUM
Cognitive basics skills 2	3	Secondary	6	9–spring	Paper-based	SUM
Reading competence	3	Secondary	6	9–spring	Paper-based	WLE
Listening	3	Secondary	6	9–spring	Paper-based	WLE

Notes: Score types: WLE= weighted maximum likelihood estimates, SUM=sum scores. Starting Cohorts: 1 = Newborns, 2 = Kindergarten, 3 = Grade 5. *No formal testing.

Table A2. Test score correlations for Starting Cohort *Newborns* (SC1).

<i>7 months</i>				
Sensorimotor development	1			
<i>2.2 years</i>				
Active vocabulary	1			
Syntax	0.81	1		
Morphology	0.78	0.82	1	
<i>3.2 months</i>				
Executive control	1			
Vocabulary	0.12	1		
Phonological working memory	0.18	0.28	1	
<i>4.2 months</i>				
Mathematics	1			

Notes: scores at age 2.2 are reported by parents.

Table A3. Test score correlations for Starting Cohort *Kindergarten* (SC2).

<i>5 years</i>									
Vocabulary	1								
Grammar	0.69	1							
Scientific literacy	0.62	0.60	1						
<i>5.9 years</i>									
Cognitive basics skills 1	1								
Cognitive basics skills 2	0.24	1							
Mathematical competence	0.35	0.35	1						
Early knowledge of letters	0.24	0.19	0.38	1					
Phonological working memory 1	0.24	0.17	0.42	0.3	1				
Phonological working memory 2	0.26	0.23	0.49	0.29	0.41	1			
Phonological awareness 1	0.15	0.19	0.40	0.48	0.32	0.32	1		
Phonological awareness 2	0.14	0.18	0.43	0.21	0.32	0.35	0.3	1	
Phonological awareness 3	0.26	0.23	0.47	0.38	0.41	0.44	0.37	0.46	1
<i>7 years</i>									
Vocabulary	1								
Grammar	0.62	1							
Scientific competence	0.56	0.56	1						
Mathematical competence	0.47	0.48	0.65	1					
<i>7.7 years</i>									
Cognitive basics skills 1	1								
Cognitive basics skills 2	0.13	1							
Mathematical competence	0.16	0.39	1						
Early reading	0.12	0.29	0.40	1					
Reading speed	0.21	0.29	0.35	0.78	1				
<i>8.7 years</i>									
Vocabulary	1								
Scientific competence	0.52	1							
<i>9.6 years</i>									
Early reading	1								
Mathematical competence	0.63	1							
Orthography 1	0.50	0.55	1						
Orthography 2	0.50	0.55	0.96	1					

Table A4. Test-score correlations for Grade 5 cohort (SC3).

<i>10.9 years</i>							
Cognitive basics skills 1	1						
Cognitive basics skills 2	0.15	1					
Reading speed	0.30	0.20	1				
Reading competence	0.08	0.42	0.34	1			
Mathematical competence	0.11	0.54	0.32	0.61	1		
Orthography 1	0.11	0.33	0.45	0.54	0.51	1	
Orthography 2	0.11	0.34	0.43	0.54	0.52	0.96	1
<i>11.9 years</i>							
Vocabulary	1						
Scientific competence	0.61	1					
<i>12.9 years</i>							
Reading competence	1						
Mathematical competence	0.63	1					
Orthography 1	0.58	0.49	1				
Orthography 2	0.60	0.51	0.96	1			
<i>14.9 years</i>							
Reading speed	1						
Mathematical competence	0.31	1					
Scientific competence	0.31	0.69	1				
Orthography 1	0.51	0.52	0.45	1			
Orthography 2	0.50	0.51	0.45	0.95	1		
<i>15.4 years</i>							
Cognitive basics skills 1	1						
Cognitive basics skills 2	0.18	1					
Reading competence	0.40	0.13	1				
Listening	0.40	0.10	0.61	1			

Table A5. Analytical samples by outcome.

Starting cohort	Wave	Age	Composite	Cog. Basic	Math	Ortho	Read	Science	Vocab
SC1: <i>Newborns</i>	1	0.6	2944	—	—	—	—	—	—
	3	2.2	2085	—	—	—	—	—	2085
	4	3.2	1386	—	—	—	—	—	1730
	5	4.2	2024	—	2024	—	—	—	—
SC2: <i>Kindergarten</i>	1	5.0	2456	—	—	—	—	2503	2468
	2	5.9	2197	2285	2341	—	—	—	—
	3	7.0	5572	—	5815	—	—	5825	5806
	4	7.7	5141	5523	5459	—	5398	—	—
	5	8.7	4830	—	—	—	—	4996	4979
	6	9.7	4539	—	5966	4750	5000	—	—
SC3: <i>Grade 5</i>	1	10.9	3703	3718	3721	3732	3723	—	—
	2	11.9	3371	—	—	—	—	3511	3514
	3	12.9	4012	—	4017	4016	4015	—	—
	5	14.9	3193	—	3208	3205	—	3250	—
	6	15.4	3000	3032	—	—	3024	—	3073
N observations	—	—	50453	14558	32551	15703	21160	20085	23655
N children	—	—	16141	11896	14798	9697	10622	12396	14621

Table A6. Parametric models for linear time trend in SES gaps in preschool, primary, and lower secondary education (composite only).

	Preschool	Primary	Secondary
Parental education	0.061*** (0.014)	0.163*** (0.012)	0.153*** (0.008)
Age	-0.196*** (0.053)	0.008 (0.076)	-0.118*** (0.035)
Parental education*Age	0.015*** (0.003)	-0.002 (0.005)	0.008** (0.002)
Migration background (yes)	0.066 (0.074)	-0.232** (0.080)	-0.125* (0.049)
Migration background*Age	-0.060** (0.019)	0.102*** (0.030)	0.053** (0.016)
Constant	-0.863***	-2.207***	-2.112***
N	13092	20082	17279

Notes: Three regression models pooling all observations on composite scores within the three institutional periods: preschool, primary, and lower secondary education. Time axis (age) as linear variable (age=0 in the beginning of each institutional period). All models adjusted for migration background and interaction with time axis. Standard errors clustered by individual. Weighted data. Standard errors in parentheses. Significance (two-sided test): * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

APPENDIX B. FIGURES

Figure B1. Structure of the accelerated longitudinal design.

Figure B2. Evolution of SES gaps in absolute terms using linked scores provided by the NEPS: Math, Science, and Reading.

Figure B3. Analysis of the composite score using metric-free measures of achievement: Ranks (instead of standardized scores).

APPENDIX C. TECHNICAL DETAILS

C1. Details on Language, Math, Science, and Cognitive Basic Skills

A first set of key skills relate to language competencies which involve the production and reception of oral and written language (Berendes et al. 2013). *Vocabulary* captures lexical knowledge of words and, as a subconcept of listening comprehension, is a regular test domain in the NEPS. Undoubtedly, *reading literacy* is not only a crucial competence for educational success, but a fundamental skill for a wide variety of social and economic outcomes. Therefore, assessment of reading competence as a receptive language component is a central concern in the measurement strategy of NEPS, which is based on international standards and involves longitudinal modelling of reading competencies over the life span (Gehrer et al. 2013). *Orthography* is measured primarily at the transition to and in secondary education (as a stage specific measure) and captures spelling achievement in German language – a skill that is crucial for academic success in the school system, much like math and reading (Schneider et al. 2004). As another component of language (writing literacy), the concept of orthography comprises mainly the ability to understand the syllabic structure of words, their inflected and derived forms, principles of word formation (e.g., compounding in German), spelling exceptions, and proper capitalization (Blatt and Prosch 2014; Jarinski et al. 2016).

Mathematical literacy is another domain subject to coherent assessment in the NEPS. Math tests were developed based on a lifespan approach and test frameworks involve content-related (major mathematical ideas) and cognitive dimensions (key abilities to solve math problems) (Neumann et al. 2013).

The NEPS testing of *scientific literacy* encompasses both knowledge of scientific concepts and facts as well as an understanding of scientific processes (Hahn et al. 2013). NEPS assessments of reading and of scientific and mathematical literacy have all been constructed in

line with national (German Education Standards) and international benchmarks and standards (such as PISA).

Finally, we inspected the evolution over time of SES inequalities in *cognitive basic skills*, which are universal rather than domain-specific skills, as they are related to the concept of ‘fluid’ intelligence or ‘cognitive mechanics’ (Brunner, Lang, and Lüdtke 2014). In contrast to the more environmentally-determined ‘crystallized’ intelligence reflected in learning attainment (such as reading or math), fluid intelligence appears to be more biologically determined (Horn and Cattell 1966). Furthermore, general cognitive skills seem to accurately predict learning, educational achievement, occupational status, and job performance (Bertua, Anderson, and Salgado 2005; Deary et al. 2007). In the context of our study, inspecting SES gaps in basic cognitive skills is therefore useful to gauge those facets of socially-determined inequality in cognitive achievement that are less dependent on environmental input. Importantly, cognitive basic skills are tested non-verbally in two picture-based tests relating to *perceptual speed* and *reasoning*, with the first employing the Picture Symbol Test (matching figures or numbers with graphical symbols), and the latter using the matrices test (subjects are asked to complete fields in a matrix of geometrical symbols) (Haberkorn and Pohl 2013).

C2. Composite Measure and Single Domains: Calculating z-scores

The construction of the composite measure of achievement followed four steps. In the first step, we standardize all test scores available within each wave in order to reflect the relative position of children in the age-specific distribution of each competence domain.

In the second step, we take the average across z-scores from different tests for each wave in order to create a synthetic measure of achievement. We opted for simple averaging instead of a more complex factor or component-based scoring solution (as applied by Feinstein 2003) to avoid sampling error entering the construction of our dependent variable. However,

using weights obtained from principal component analysis yielded very similar results (available upon request).

In the third step, we purge the averaged z-scores from the confounding effect of non-SES related age variations when children performed the tests, separately in each wave. We had to consider age variations among children in the same life stage; NEPS has a ‘stage’ design and thus age variations among children in the same life stage may occur. While previous studies based on similar data rule out the confounding effect of age-variations *tout court* (for example, Bradbury et al. 2015), we believe some of these variations must be preserved, as they substantially reflect the overall influence of SES. For example, part of the age variations in primary or secondary education may reflect the role of grade retention, which may occur more frequently among low SES children. This is probably the reason why, in our sample, low SES children are slightly older than high SES children in the same grade in secondary education. At the same time, among children in the same grade in secondary education, age is negatively correlated with test scores, net of SES (details available upon request). In this scenario, controlling for age without further detail would severely underestimate SES gaps in secondary education and, therefore, strongly bias the interpretation of the evolution of SES inequalities over the entire observation window. Insofar as we aim to explore the evolution of achievement gaps over the educational career through an institutional lens, potential SES-related age variations among children in the same educational stage must be preserved, in our view. Step 3 comprises two stages in order to achieve this goal: first, we residualize the age of children at test day as a function of SES (P), separately at each time point t:

$$Age_{i|t} = \alpha_t + b_{1t}P_i + e_{i|t} \quad (C1)$$

$$\hat{u}_{i|t} = Age_{i|t} - \widehat{Age}_{i|t} \quad (C2)$$

where the residuals $\hat{u}_{i|t}$ by construction reflect non-SES related age variations among children in the same life stage. Second, we residualize the averaged z-scores at each time point t as a cubic function of the non-SES related variations in children's age ($\hat{u}_{it}$):

$$AVzscore_{i|t} = \alpha_t + b_{1t}\hat{u}_{i|t} + b_{2t}\hat{u}_{i|t}^2 + b_{3t}\hat{u}_{i|t}^3 + e_{i|t} \quad (C3)$$

$$\hat{\delta}_{i|t} = AVzscore_{i|t} - \widehat{AVzscore}_{i|t} \quad (C4)$$

where the residuals $\hat{\delta}_{i|t}$ by construction reflect variations in the standardized scores that are independent from non-SES related age variations among children in the same life stage.

In the fourth and final step, we standardize within each time point (wave) the residuals $\hat{\delta}_{i|t}$ obtained in C4 and use them as our final independent variable. Steps one, three (with the only z-score available for each domain), and four were applied equally to the single domain measures. Basic cognitive skills and orthography are an exception since they are always captured by two tests. Therefore, the procedure for these two domains is the same as for the composite measure (in step two we average the available z-scores).

C3. Weighting Approach

A complex weighting approach is applied to adjust for (a) the sampling design of the NEPS cohorts; (b) the varying marginal distribution of parental education, migration background, and sex across the five initial samples (3 starting cohorts plus 2 refreshment samples); and (c) selective panel attrition. The weighting procedure below is applied separately to each of the seven outcomes (the composite measure and the six competence domains).

The overall weight in wave j is calculated as the product of the wave-specific inverse probabilities π_t :

$$w_j = \prod_{t=1}^j \pi_t \quad (C5)$$

where π_t represents different quantities in the five sampling waves (SC1–wave 1, SC2–wave 1, SC2–wave 3, SC3–wave 1, SC3–wave 3) and the following waves.

In the five sampling waves k , π_t is the product of the sampling design weight A_k and the cohort standardisation weight B_k . A_k is provided by the NEPS and is defined as the inverse of the probability of being included in the initial sample ($A_k = \frac{1}{p_0}$). B_k is the cohort standardization weight, which we computed as the inverse of the conditional probability of being part of the initial sample in sampling wave k ($B_k = \frac{1}{\Pr(k|X)}$), estimated from a pooled multinomial logistic model including all observations from children that are non-missing on the analytical variables in the five sampling waves:

$$\text{logit}(p_i^k) = \alpha^k + b_1^k P_i + b_2^k M_i + b_3^k G_i + b_4^k P_i \cdot M_i + e_i^k \quad (C6)$$

where P is parental education, M migration background, and G the sex of the child. Children missing on the outcome variable in their sampling wave k and being non-missing at later time points are assigned the cohort standardization weight B_k according to the predictors included in C6. A_k adjusts for the non-randomness of inclusion in the five initial samples, while B_k adjusts for disproportions in the distribution of some key observable characteristics across the five sampling waves.

In the waves following the sampling wave, we compute π_t as the inverse of the response rates of the groups defined by quintiles of the distribution of the predicted conditional probabilities of participating in wave t given participation at $t-1$ ($\pi_t = \frac{1}{\Pr(t|t-1, X)}$). We use the actual response rate of the groups in order to avoid large adjustment factors due to very low predicted probabilities. This procedure is known as *response propensity score stratification* and is quite robust to misspecification of the weighting model (see Chen et al. 2012). The probability of participating in wave t given participation at $t-1$ is predicted using the following logistic model:

$$\begin{aligned} \text{logit}(p_{it|t-1}) = & \alpha_t + b_{1t}P_i + b_{2t}M_i + b_{3t}G_i + b_{4t} \cdot \text{zscore}_{i,t-1} \\ & + b_{5t}P_i \cdot \text{zscore}_{i,t-1} + e_{it} \end{aligned} \quad (\text{C7})$$

where P stands for parental education, M for migration background, G for sex, and zscore for the outcome under investigation. In order to preserve a good sample size, in case of subjects re-entering the sample at time t after non-participation at time $t-1$ (they are by construction missing on $\Pr(t|t-1)$ and, therefore, on π_t), we assign the overall response rate π_t in that wave. For simplicity, we do not model missingness on the analytical variables in the 5 sampling waves and consider non-participation and list-wise missingness on the analytical variables in the following waves to be equivalent.

Weights accounting for sampling design, disproportion in the marginal distributions of key variables, and selective attrition are used throughout the construction of the outcome variables as well as the empirical analyses.

REFERENCES

- Alexander, K. L., D. R. Entwisle, and L. S. Olson. 2001. "Schools, Achievement, and Inequality: A Seasonal Perspective." *Educational Evaluation and Policy Analysis* 23(2):171–91.
- Allmendinger, Jutta. 1989. "Educational Systems and Labor Market Outcomes." *European Sociological Review* 5(3):231–50.
- Arts, Wilhelmus Antonius and John Gelissen. 2002. "Three Worlds of Welfare Capitalism or More? A State-of-the-Art Report." *Journal of European Social Policy* 12(2):137–58.
- Ayalon, Hanna and Adam Gamoran. 2000. "Stratification in Academic Secondary Programs and Educational Inequality in Israel and the United States." *Comparative Education Review* 54–80.
- Barone, Carlo and Herman G. Van De Werfhorst. 2011. "Education, Cognitive Skills and Earnings in Comparative Perspective." *International Sociology* 26(4):483–502.
- Baumert, Jürgen, Gabriel Nagy, and Rainer Lehmann. 2012. "Cumulative Advantages and the Emergence of Social and Ethnic Inequality: Matthew Effects in Reading and Mathematics Development Within Elementary Schools?" *Child Development* 83(August):1347–67.
- Benavot, Aaron and Julia Resnik. 2004. "Lessons from the Past: A Comparative Socio-Historical Analysis of Primary and Secondary Education." in *Global Educational Expansion. Historical Legacies and Political Obstacles*, edited by A. Benavot, J. Resnik, and J. Corrales. Cambridge: American Academic of Arts and Sciences.
- Berendes, Karin, Sabine Weinert, Stefan Zimmermann, and Cordula Artelt. 2013. "Assessing Language Indicators across the Lifespan within the German National Educational Panel Study (NEPS)." *Journal for Educational Research Online* 5(2):15–49.
- Bertua, Cristina, Neil Anderson, and Jesús F. Salgado. 2005. "The Predictive Validity of Cognitive Ability Tests: A UK Meta-Analysis." *Journal of Occupational and Organizational Psychology* 78(3):387–409.
- Blatt, Inge and Anna Prosch. 2014. *The Assessment of Orthographic Competence*. 50. Bamberg.
- Blossfeld, Hans-Peter, Sandra Buchholz, Jan Skopek, and Moris Triventi, eds. 2016. *Models of Secondary Education and Social Inequality: An International Comparison*. Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
- Blossfeld, Hans-Peter, Hans-Günther Roßbach, and Jutta von Maurice, eds. 2011. *Education as a Lifelong Process - The German National Educational Panel Study (NEPS)*.

Heidelberg: Springer VS.

BMFSFJ. 2015. *Fünfter Bericht Zur Evaluation Des Kinderförderungsgesetzes*.

Bol, Thijs and Herman G. van de Werfhorst. 2011. "Measuring Educational Institutional Diversity: External Differentiation , Vocational Orientation and Standardization Thijs Bol." *Amsterdam Centre for Inequality Studies Working Paper Series* 1(November).

Bol, Thijs, Jacqueline Witschge, Herman H. Van de Werfhorst, and Jaap Dronkers. 2014. "Curricular Tracking and Central Examinations: Counterbalancing the Impact of Social Background on Student Achievement in 36 Countries." *Social Forces* (1):1–48.

Bos, Wilfried. 2007. *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland Im Internationalen Vergleich*. Münster: Waxmann Verlag.

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

Bradbury, Bruce, Miles Corak, Jane Waldfogel, and Elizabeth Washbrook. 2015. *Too Many Children Left Behind. The U.S. Achievement Gap in Comparative Perspective*. New York: Russell Sage Foundation.

Bradley, Robert H. and Robert F. Corwyn. 2002. "Socioeconomic Status and Child Development." *Annual Review of Psychology* 53(1):371–99.

Breen, Richard and J. H. Goldthorpe. 1997. "Explaining Educational Differentials: Towards a Formal Rational Action Theory." *Rationality and Society* 9(3):275–305.

Brunner, Martin, Frieder R. Lang, and Oliver Lüdtke. 2014. *Erfassung Der Fluiden Kognitiven Leistungsfähigkeit Über Die Lebensspanne Im Rahmen Der National Educational Panel Study: Expertise*. 42. Bamberg.

Buchholz, Sandra et al. 2016. "Secondary School Differentiation and Inequality of Educational Opportunity in Germany." Pp. 79–92 in *Models of Secondary Education and Social Inequality. An International Comparison*., edited by H.-P. Blossfeld, S. Buchholz, J. Skopek, and M. Triventi. Celtenham, UK: Edward Elgar.

Buchmann, Claudia. 2002. "Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges." Pp. 150–97 in *Methodological Advances in Cross-National Surveys of Educational Achievement*, edited by A. C. Porter and A. Gamoran. Washington, D.C.: National Academy Press.

Caro, Daniel H., James Ted McDonald, and J. Douglas Willms. 2009. "Socio-Economic Status and Academic Achievement Trajectories from Childhood to Adolescence." *Canadian Journal of Education* 32(3):558–90.

Cheadle, Jacob E. 2008. "Educational Investment, Family Context, and Children's Math and

- Reading Growth from Kindergarten Through the Third Grade.” *Sociology of Education* 81(1):1–31.
- Chen, Qixuan, Andrew Gelman, Melissa Tracy, Fran H. Norris, and Sandro Galea. 2012. *Weighting Adjustments for Panel Nonresponse*.
- Chmielewski, Anna K. 2017. *The Global Increase in the Socioeconomic Achievement Gap, 1964-2015*. 17-04.
- Coleman, James S. 1966. *Equality of Educational Opportunity*. Washington.
- Condrón, Dennis J. 2008. “An Early Start: Skill Grouping and Unequal Reading Gains in the Elementary Years.” *The Sociological Quarterly* 49(2):363–94.
- Conger, Rand D., Katherine J. Conger, and Monica J. Martin. 2010. “Socioeconomic Status, Family Processes, and Individual Development.” *Journal of Marriage and Family* 72(3):685–704.
- Conger, Rand D. and M. Brent Donnellan. 2007. “An Interactionist Perspective on the Socioeconomic Context of Human Development.” *Annual Review of Psychology* 58(1):175–99.
- Dämmrich, Johanna and Moris Triventi. 2018. “The Dynamics of Social Inequalities in Cognitive-Related Competencies along the Early Life Course – A Comparative Study.” *International Journal of Educational Research* 88(January):73–84.
- Deary, Ian J., Steve Strand, Pauline Smith, and Cres Fernandes. 2007. “Intelligence and Educational Achievement.” *Intelligence* 35(1):13–21.
- DiPrete, Thomas a. and Gregory M. Eirich. 2006. “Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments.” *Annual Review of Sociology* 32:271–97.
- Ditton, Hartmut. 2013. “Bildungsverläufe in Der Sekundarstufe. Ergebnisse Einer Längsschnittstudie Zu Wechseln Der Schulform Und Des Bildungsgangs.” *Zeitschrift Für Pädagogik* (6):887–911.
- Domina, Thurston. 2005. “Leveling the Home Advantage : Assessing the Effectiveness of Parental Involvement in Elementary School.” *Sociology of Education* 78:233–49.
- Downey, Douglas B. and Dennis J. Condrón. 2016. “Fifty Years since the Coleman Report: Rethinking the Relationship between Schools and Inequality.” *Sociology of Education* 89(3):207–20.
- Downey, Douglas B., Paul T. von Hippel, and Beckett A. Broh. 2004. “Are Schools the Great Equalizer? Cognitive Inequality during the Summer Months and the School Year.” *American Sociological Review* 69(5):613–35.

- Duncan, Greg J. and Katherine A. Magnuson. 2003. "Off with Hollingshead: Socioeconomic Resources, Parenting, and Child Development." Pp. 83–106 in *Socioeconomic status, parenting, and child development*, edited by M. H. Bornstein and R. H. Bradley. Mahwah, New Jersey: Lawrence Erlbaum.
- Duncan, Greg J., Katherine Magnuson, and Elizabeth Votruba-Drzal. 2015. "Children and Socioeconomic Status." Pp. 534–73 in *Handbook of Child Psychology and Developmental Science: Ecological Settings and Processes*.
- Eder, Donna. 1981. "Ability Grouping as a Self-Fulfilling Prophecy: A Micro-Analysis of Teacher-Student." *Sociology of Education* 54(3):151–62.
- Ermisch, John. 2008. "Origins of Social Immobility and Inequality: Parenting and Early Child Development." (205):62–71.
- Esping-Andersen, G. 1990. *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.
- van Ewijk, Reyn and Peter Sleegers. 2010. "The Effect of Peer Socioeconomic Status on Student Achievement: A Meta-Analysis." *Educational Research Review* 5(2):134–50.
- Fan, Xitao and Michael Chen. 2001. "Parental Involvement and Student's Academic Achievement: A Meta-Analysis." *Educational Psychology Review* 13(1):1–22.
- Farkas, George. 2003. "Cognitive Skills and Noncognitive Traits and Behaviours in Stratification Processes." *Annual Review of Sociology* 29:541–62.
- Farkas, George and Kurt Beron. 2004. "The Detailed Age Trajectory of Oral Vocabulary Knowledge: Differences by Class and Race." *Social Science Research* 33(3):464–97.
- Feinstein, Leon. 2003. "Inequality in the Early Cognitive Development of British Children in the 1970 Cohort." *Economica* 70(277):73–97.
- Fernald, Anne, Virginia A. Marchman, and Adriana Weisleder. 2013. "SES Differences in Language Processing Skill and Vocabulary Are Evident at 18 Months." *Developmental Science* 16(2):234–48.
- Fischer, Luise, Theresa Rohm, Timo Gnamb, and Claus H. Carstensen. 2016. *Linking the Data of the Competence Tests*. 1. Bamberg.
- Gamoran, Adam and Robert D. Mare. 1989. "Secondary School Tracking and Educational Inequality: Compensation, Reinforcement or Neutrality?" *American Journal of Sociology* 84(5):1146–83.
- Gehrer, Karin, Stefan Zimmermann, Cordula Artelt, and Sabine Weinert. 2013. "NEPS Framework for Assessing Reading Competence and Results from an Adult Pilot Study." *Journal for Educational Research Online* 5(2):50–79.
- Gresch, Cornelia, Jürgen Baumert, and Kai Maaz. 2010. "Empfehlungsstatus,

Übergangsempfehlung Und Der Wechsel in Die Sekundarstufe I: Bildungsentscheidungen Und Soziale Ungleichheit.” Pp. 230–56 in *Bildungsentscheidungen*.

Haberkorn, Kerstin and Steffi Pohl. 2013. *Cognitive Basic Skills - Data in the Scientific Use File*. Bamberg.

Hahn, Inga et al. 2013. “Assessing Scientific Literacy over the Lifespan – A Description of the NEPS Science Framework and the Test Development.” *Journal for Educational Research Online* 5(2):110.

Hanushek, Eric A. and Ludger Woessmann. 2006. “Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries.” *The Economic Journal* 116(1984):63–76.

Hanushek, Eric A. and Ludger Woessmann. 2008. “The Role of Cognitive Skills in Economic Development.” *Journal of Economic Literature* 46(3):607–68.

Hart, Betty and Todd R. Risley. 2003. “The Early Catastrophe. The 30 Million Word Gap.” *American Educator* 27(1):4–9.

Heckman, James J. and Flavio Cunha. 2007. “The Technology of Skill Formation.” *The American Economic Review* 97(2):31–47.

Helbig, Marcel and Rita Nikolai. 2015. *Die Unvergleichbaren. Der Wandel Der Schulsysteme in Den Deutschen Bundesländern Seit 1949*. Bad Heilbrunn: Verlag Julius Klinkhardt.

Horn, Daniel. 2009. “Age of Selection Counts: A Cross- Country Analysis of Educational Institutions.” *Educational Research and Evaluation* 343–66.

Horn, John L. and Raymond B. Cattell. 1966. “Refinement and Test of the Theory of Fluid and Crystallized General Intelligences.” *Journal of Educational Psychology* 57(5):253–70.

Jackson, Michelle. 2013. *Determined to Succeed? Performance versus Choice in Educational Attainment*. Stanford: Stanford University Press.

Jarinski, Stephan, Sarah Frahm, Inge Blatt, Wilfried Bos, and Michael Kanders. 2016. “Assessing Spelling Competence Development in the National Educational Panel Study.” Pp. 427–39 in *Methodological Issues of Longitudinal Surveys: The Example of the National Educational Panel Study*, edited by H.-P. Blossfeld, J. von Maurice, M. Bayer, and J. Skopek. Wiesbaden: Springer VS.

Kerckhoff, Alan C. and Elizabeth Glennie. 1999. “The Matthew Effect in American Education.” *Research in Sociology of Education and Socialization* 12:35–66.

Layte, Richard. 2017. “Why Do Working-Class Kids Do Worse in School? An Empirical Test of Two Theories of Educational Disadvantage.” *European Sociological Review*

33(4):489–503.

- Lee, Valerie E. and David T. Burkam. 2002. *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School*. Washington, D.C.: Economic Policy Institute.
- Leopoldina. 2014. *Socialisation in Early Childhood. Biological, Psychological, Linguistic, Sociological and Economic Perspectives*. Berlin: German National Academy of Sciences Leopoldina.
- Lleras, Christy. 2008. “Do Skills and Behaviors in High School Matter? The Contribution of Noncognitive Factors in Explaining Differences in Educational Attainment and Earnings.” *Social Science Research* 37:888–902.
- Lucas, Samuel R. 2001. “Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects.” *American Journal of Sociology* 106(6):1642–90.
- Marks, G. N. 2005. “Cross-National Differences and Accounting for Social Class Inequalities in Education.” *International Sociology* 20(4):483–505.
- McLoyd, Vonnice C. 1998. “Socioeconomic Disadvantage and Child Development.” *American Psychologist* 53(2):185–204.
- Morgan, Paul L., George Farkas, and Jacob Hibbel. 2008. “Matthew Effects for Whom?” *Learning Disability Quarterly* 31(4):187–98.
- Mullis, Ina V. S., Michael O. Martin, and Tom Loveless. 2016. *20 Years of TIMSS 20 Years of TIMSS International Trends in Mathematics*. Chestnut Hill: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Neugebauer, Martin. 2010. “Bildungsungleichheit Und Grundschulempfehlung Beim Übergang Auf Das Gymnasium: Eine Dekomposition Primärer Und Sekundärer Herkunftseffekte.” *Zeitschrift Für Soziologie* 39(3):202–14.
- Neugebauer, Martin, David Reimer, Steffen Schindler, and Volker Stocké. 2013. “Inequality in Transitions to Secondary School and Tertiary Education in Germany.” Pp. 56–88 in *Determined to Succeed?*, edited by M. Jackson. Stanford University Press.
- Neumann, Irene et al. 2013. “Modeling and Assessing Mathematical Competence over the Lifespan.” *Journal for Educational Research Online* 5(2):80–109.
- Oakes, J. Michael and Peter H. Rossi. 2003. “The Measurement of SES in Health Research: Current Practice and Steps toward a New Approach.” 56:769–84.
- OECD. 2016. *PISA 2015 Results in Focus (Volume I)*. Vol. I.
- OECD. 2015. *Starting Strong IV*. OECD Publishing.

- Pohl, Steffi and Claus H. Carstensen. 2012. "NEPS Working Papers." (14).
- Potter, Daniel and Josipa Roksa. 2013. "Accumulating Advantages over Time: Family Experiences and Social Class Inequality in Academic Achievement." *Social Science Research* 42(4):1018–32.
- Protopapas, A., G. D. Sideridis, A. Mouzaki, and P. G. Simos. 2011. "Matthew Effects in Reading Comprehension: Myth or Reality?" *Journal of Learning Disabilities* 44(5):402–20.
- Raudenbush, Stephen W. and Robert D. Eschmann. 2015. "Does Schooling Increase or Reduce Social Inequality?" *Annual Review of Sociology* 41(1):443–70.
- Reardon, Sean F. 2011. "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations." Pp. 91–116 in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by G. J. Duncan and R. J. Murnane. New York: Russell Sage Foundation.
- Reardon, Sean F. 2008. *Thirteen Ways of Looking at the Black-White Test Score Gap*.
- Rözer, Jesper and Herman Van De Werfhorst. 2017. *Inequalities in Educational Opportunities by Socioeconomic and Migration Background: A Comparative Assessment Across European Societies*.
- Schneider, Wolfgang, Dieter Wolke, Matthias Schlagmüller, and Renate Meyer. 2004. "Pathways to School Achievement in Very Preterm and Full Term Children." *European Journal of Psychology of Education* 19(4):385–406.
- Schubert, Frank and Rolf Becker. 2010. "Social Inequality of Reading Literacy. A Longitudinal Analysis with Cross-Sectional Data of PIRLS 2001 and PISA 2000 Utilizing the Pair Wise Matching Procedure." *Research in Social Stratification and Mobility* 28(1):109–33.
- Schütz, Gabriela, Heinrich Ursprung, and Ludger Woessmann. 2008. "Education Policy and Equality of Opportunity." *Kyklos* 61(2):279–308.
- Shonkoff, Jack P. and Deborah A. Phillips. 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Vol. 23. Washington, D.C.: National Academy Press.
- Sirin, Selcuk R. 2005. "Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research." *Review of Educational Research* 75(3):417–53.
- Skopek, J. and M. Munz. 2016. "Life-Course Data and the Longitudinal Classification of Education." Pp. 669–90 in *Methodological Issues of Longitudinal Surveys: The Example of the National Educational Panel Study*, edited by H.-P. Blossfeld, J. von Maurice, M. Bayer, and J. Skopek. Wiesbaden: Springer VS.

- Smeeding, Timothy M. and Lee Rainwater. 2002. *Comparing Living Standards Across Nations: Real Incomes at the Top, the Bottom and the Middle*. 266. Luxembourg.
- Stanovich, Keith E. 1986. "Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy." *Reading Research Quarterly* 21(4):360–407.
- Steinhauer, Hans Walter, Christian Aßmann, Sabine Zinn, Solange Goßmann, and Susanne Rässler. 2015. "Sampling and Weighting Cohort Samples in Institutional Contexts: The National Educational Panel Study Cohort Samples of Kindergarten Children, Students in Grade 5 and in Grade 9." *AStA Wirtschafts- Und Sozialstatistisches Archiv* 9(2):131–57.
- Stevenson, David Lee and David P. Baker. 1991. "State Control of the Curriculum and Classroom Instruction." *Sociology of Education* 64(1):1–10.
- Stocké, Volker. 2007. "Explaining Educational Decision and Effects of Families' Social Class Position: An Empirical Test of the Breen-Goldthorpe Model of Educational Attainment." *European Sociological Review* 23(1974):505–19.
- Van de Werfhorst, Herman G. and Jonathan J. B. Mijs. 2010. "Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective." *Annual Review of Sociology* 36(1):407–28.
- White, Karl R. 1982. "The Relation between Socioeconomic Status and Academic Achievement." *Psychological Bulletin* 91(3):461–81.
- Würbach, Ariane, Sabine Zinn, and Christian Aßmann. 2016. *Sample Weights and Non-Response: The Early Childhood Cohort of the National Educational Panel Study (Wave 1 to 3)*. 8. Bamberg.
- Zeanah, Charles H., Neil Boris, and Julie A. Larrieu. 1997. "Infant Development and Developmental Risk : A Review of the Past 10 Years." *Journal of the American Academy of Child & Adolescent Psychiatry* 36(2):165–78.