Steinmann, Isa; Strietholt, Rolf

Student achievement and educational inequality in half- and all-day schools. Evidence from Germany

International journal for research on extended education : IJREE 6 (2018) 2, S. 175-197

IJREE

Issue 2 2018

International Journal for **Research** on **Extended Education**

ISSN 2196-3673

Verlag **Barbara Budrich**

International Journal for Research on
Extended Education, Volume 6/2018

Content

GENERAL CONTRIBUTIONS

DEVELOPMENTS in the Field of Extended Education

Student achievement and educational inequality in half- and all-day schools: Evidence from Germany

Isa Steinmann & Rolf Strietholt[*]

Abstract: Several countries have expanded extended education in recent years. In Germany, the most substantial educational reform is the ongoing transformation of the traditional half-day school system into an all-day school system. Among politicians, expectations are high that all-day schools will promote student achievement and reduce social achievement inequalities. To test these assumptions, we used representative data from the National Educational Panel Study (NEPS) to estimate two-level latent growth models for achievement in grades 5, 7, and 9. The analyses revealed initial achievement differences but no differences in achievement growth or changes in inequality throughout secondary school. This suggests that selection mechanisms are at work but that half- and all-day schools are not differentially effective. We discuss these findings in light of the international debate on the quality of extended education.

Keywords: extended education, all-day school, mathematics achievement, reading achievement, educational inequality

Introduction and Research Question

Learning takes place in various contexts, which can be depicted on a continuum from informal to formal settings. Informal learning is not organized and takes place unintentionally and continuously in everyday life. By contrast, formal learning takes place in organized, highly structured contexts that are designated for learning (e.g., regular school lessons). Between these two poles lies non-formal learning (Werquin, 2010). Extended education—like private tutoring or extra-curricular activities at schools—is a group of non-formal contexts that is intended to promote learning and is pedagogically structured but less formalized than regular classes (see Stecher & Maschke, 2013). Especially school-based, non-formal extended education is increasingly politically relevant in many countries because it is (a) expected to improve student learning outcomes and (b) more open to external influence than,

[*] Acknowledgements: We would like to thank Roisin Cronin for copy editing the manuscript.

for instance, private tutoring (see Kuger, 2016; Plantenga & Remery, 2013; Vest, Mahoney, & Simpkins, 2013).

In Germany, the largest education policy reform of the past decades concerns school-based extended education. In the 20th century, the school day in most schools consisted of morning classes but no afternoon program.[1] Between 2003 and 2009, a federal investment program of more than four billion euros (BMBF, 2003) prompted a large increase in the proportion of all-day schools from 16% in 2002 to 68% in 2016 (KMK, 2008, 2018). The program supported the founding of new all-day schools and the infrastructural development of existing ones. One aim behind the massive expansion of all-day schools was that the extended supervision of children would facilitate maternal employment (e.g., Fischer, Theis, & Züchner, 2014; Plantenga & Remery, 2013). However, politicians also expected the schools to increase opportunities to support learning and reduce educational gaps between social groups (e.g., BMBF, 2003; Fischer et al., 2014). The definition of all-day schools adopted by the investment program was: (a) All-day schools provide lunch and at least a seven-hour program on at least three days per week. (b) This program is supervised by (and conducted in close cooperation with) the school administration, which is also accountable for the program. (c) It is conceptually connected to the regular classes (KMK, 2008). Expectations were high that all-day schooling improves learning outcomes because all-day schools provide additional time and personnel resources that can be used to enhance schooling, for example, by providing more individual support and diverse learning activities (BMBF, 2003). Learning benefits were especially expected for socially disadvantaged students who have less stimulating environments in the afternoon at home and for children whose families do not speak German at home. Therefore, all-day schooling was expected to contribute to a reduction of social achievement gaps (e.g., BMBF, 2003; Steiner, 2009; Züchner & Fischer, 2014).

Based on the aforementioned all-day school definition, it is reasonable to expect that all-day schooling would extend formal learning opportunities. In fact, however, at all-day schools the regular morning classes are typically supplemented by non-formal components in the afternoon, not by additional formal instruction. Although the definition requires that the all-day program and regular classes are conceptually related, research indicates that this does not hold at about half of all-day schools (StEG, 2015). Generally, the definition gives schools considerable room for interpretation: It does not define qualitative characteristics such as the pedagogical content or the qualifications of the supervising staff. The schools decide whether participation in afternoon activities is nonmandatory, partially mandatory (e.g., for students in certain grades), or fully mandatory for all students. As a result, the organization of all-day schools varies within and between states, primary and secondary schools, and school types/academic tracks (KMK, 2015; StEG, 2015). In addition, all-day programs are often not supervised by teachers but by differently and often less qualified staff. Programs designed to promote student learning are much less common than those promoting leisure and sports activities (StEG, 2013, 2016). Moreover, the student enroll-

1 In the 1970s, some initiatives started to establish "Gesamtschulen" in Germany. "Gesamtschulen" were a new school type that intended to overcome the traditional ability grouping in different school types and combine ability tracks within schools. Furthermore, they were organized as all-day schools. However, the share of those schools was small and they were limited to few federal states.

ment rates are often low, even at schools with a broad all-day program, which sets limitations to the potential benefit of all-day over half-day schools. However, students with working parents, low socioeconomic status, or an immigrant background are especially likely to make use of all-day schooling, even though many schools charge fees for lunch and/or attending programs (StEG, 2016; Steiner, 2011). The observation that all-day programs seem to be able to reach disadvantaged groups reinforces the idea that all all-day schools may decrease social inequalities in achievement (cf. also Steiner, 2009; Steinmann, 2018; Züchner & Fischer, 2014).

The heterogeneous characteristics and implications of all-day schooling are especially relevant in light of models on the effectiveness of extended education, which highlight the importance of *how* extended time resources are used. Such models suggest that both program characteristics and student participation mediate effects on student outcomes. Among other things, the assumed quality prerequisites for effective extended education include the types of activities and their level of structure (e.g., curriculum, alignment with learning) as well as personnel resources (e.g., small staff-to-child ratios, high staff qualifications) and students' frequent and intensive participation (cf. Fischer & Klieme, 2013; Miller & Truong, 2009). The models suggest that extended education programs that are closer to the formal pole of the informal-formal learning continuum are more successful in promoting student learning. Therefore, one empirical question is whether non-formal learning provided in addition to regular schooling at all-day schools actually improves the learning outcomes of (groups of) students. In the present study, we investigated whether all-day schools are more successful in promoting student achievement and reducing educational inequality between social groups than traditional half-day schools.

Review of Literature

Despite the substantial recent investment in all-day schooling in Germany, few studies have evaluated its effects on achievement and inequality. Studies with robust research designs are particularly rare. One group of longitudinal studies compared students who participated in non-formal all-day programs with students who did not attend all-day programs. None of these studies found effects on student achievement after controlling for prior achievement and further background characteristics (Bellin & Tamke, 2010; Fischer, Sauerwein, Theis, & Wolgast, 2016; Linberg, Struck, & Bäumer, 2018; Lossen, Tillmann, Holtappels, Rollett, & Hannemann, 2016; Steinmann, Strietholt, & Caro, 2018). Bellin and Tamke (2010) further investigated if students with a migration background profit more from the all-day participation than native peers but found no support for this assumption. However, there are general issues with studies at the student level, resulting from possible spillover effects within schools that are related to remedial education measures (e.g., nonparticipating students may receive more attention during regular classes).

Studies that investigated all-day schooling as a school-level measure circumvent some issues related to selection mechanisms and spillover effects that operate at the individual level within schools. Only three studies used proxies or test measures of prior achievement as controls to compare performance at half- and all-day schools or at schools with and

without afternoon programs offering homework supervision and remedial courses. Again, none of these studies revealed student achievement effects (Linberg et al., 2018; Steinmann & Strietholt, in print; Strietholt, Manitius, Berkemeyer, & Bos, 2015). Two of the three studies investigated effects on educational inequality and found null results (Steinmann & Strietholt, in print; Strietholt et al., 2015). They examined for example inequalities in achievement scores between students with a high and a low social status or between students with and without German as first language. However, the longest investigated time span was two years. Furthermore, all these studies failed to take into account the ongoing changes from half- to all-day schools because they only determined the schools' organization form at one time point. This lack of precision likely led to biased effect estimates.

In order to contextualize the findings for Germany, we briefly summarize findings on the circumstances under which non-formal extended education programs are found to foster learning. In the US, extended education programs (e.g., afterschool programs, summer schools) have been studied extensively in experimental and quasi-experimental studies. Meta-analyses and literature reviews of this research suggest that some programs had no effects while others showed positive effects on student achievement. In contrast, positive effects were observed for programs with the following characteristics: They were designed to promote specific competences, they were closely linked to the regular curriculum, or they employed evidence-based educational approaches (Apsler, 2009; Cooper, Charlton, Valentine, Muhlenbruck, & Borman, 2000; Durlak, Weissberg, & Pachan, 2010; Lauer et al., 2006). Programs targeting at-risk students showed particularly positive effects (Durlak et al., 2010; Patall, Cooper, & Allen, 2010). Effective programs also employed highly qualified staff (Feldman & Matjasko, 2005; Lauer et al., 2006). However, programs that did not meet these quality characteristics showed smaller and often no effects on student outcomes (Apsler, 2009; Durlak et al., 2010; Roth, Malone, & Brooks-Gunn, 2010). In summary, programs that were located at the formal rather than the informal end of the learning continuum—i.e., that were more comparable to regular schooling—showed the most promising results. Additionally, some studies indicated that disadvantaged groups like students with a low socioeconomic status profit more from extended education, which implies that all-day schools reduce social achievement gaps (Lauer et al., 2006; Patall et al., 2010).

Hypotheses

The present study aimed to compare achievement levels and social inequalities in achievement in half- and all-day schools in Germany. Specifically, we investigated three research questions: Did the school's organization form have an impact on (a) reading and mathematics achievement, (b) achievement inequality regarding social status, and (c) achievement gaps between students who did and did not learn German as a first language? The political sphere expects all-day schooling to boost achievement levels and reduce educational inequality in achievement, while the scientific debate and previous findings cast doubt on these optimistic expectations. We used representative longitudinal large-scale data with test information for grades 5, 7, and 9 to investigate the three research questions.

Methods

Data

This paper used data from the National Educational Panel Study (NEPS; Blossfeld, Roßbach, & von Maurice, 2011).[2] NEPS collected longitudinal data on achievement, educational processes, and educational organizations, as well as on returns to education for different age cohorts. In the present study, we focused on a sample of secondary school students who were first tested and surveyed in grade 5 (Frahm et al., 2011; Strietholt et al., 2013). We limited our analyses to the *N*=3444 fifth graders who remained in the survey, i.e. they stayed at the same 164 schools until grade 9. Further, we excluded 20 schools whose principals' did not specify their schools' organization form at any measurement point (for more information, see section 4.2.5).In total, we investigated a sample of *N*=3024 students at 144 regular schools in 15 federal states[3] who were followed from grade 5 (school year 2010/11) until grade 9 (school year 2014/15). They were on average 10.9 years old in grade 5 and 49% were female.

Variables

Student Achievement

Mathematics and reading achievement scores are available for grades 5, 7, and 9. The achievement scales are comparable over time because the paper-pencil-based tests were linked by the anchor-item design they employ (Fischer, Rohm, Gnambs, & Carstensen, 2016). We used longitudinally linked weighted maximum likelihood estimates.[4] Table 1 shows roughly linear achievement increases between the measurement points. The overall increases between grades 5–9 corresponded to a bit more than one standard deviation. While there was some variance across measurement points in the reliability of the tests, the EAP/PV reliability was constantly high, ranging between .76 and .81 (Duchhardt & Gerdes, 2012; Krannich et al., 2017; Pohl, Haberkorn, Hardt, & Wiegand, 2012; Scharl, Fischer, Gnambs, & Rohm, 2017; Schnittjer & Gerken, 2017).[5]

2 NEPS drew a pepresentative random sample of fifth graders in the school year 2010/11 (*Starting Cohort 3,* doi:10.5157/NEPS:SC3:7.0.1). From 2008 to 2013, NEPS data was collected a spart of the Framework Program fort he Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LfBi) at the University of Bamberg in cooperation with a nationwide network.

3 Berlin was excluded, because no school entirely covered grades 5–9. This is because the transition from primary to secondary schools in Berlin typically takes place between grades 6 and 7.

4 We rescaled all achievement scores by multiplying the original by 100 to improve the readability of the estimates in the main analyses.

5 Reliability information on the mathematics test in grade 9 has not yet been published.

Table 1. Descriptive statistics for student and school characteristics

	All schools	Divided by school organization form		
		Half-day school	All-day school	Changed status
Student characteristics	*N*=3024	*N*=731	*N*=1326	*N*=967
Student achievement				
Mathematics grade 5 (*M*(*SD*))	-1.67	-35.12	14.87	6.50
	(113.40)	(120.51)	(110.92)	(104.94)
Mathematics grade 7 (*M*(*SD*))	74.83	37.35	92.81	84.53
	(122.68)	(129.96)	(118.08)	(115.48)
Mathematics grade 9 (*M*(*SD*))	151.27	114.26	166.25	163.67
	(116.96)	(118.47)	(117.66)	(108.84)
Reading grade 5 (*M*(*SD*))	0.19	-31.94	12.63	11.54
	(122.64)	(126.29)	(121.10)	(117.01)
Reading grade 7 (*M*(*SD*))	70.44	39.40	77.84	86.07
	(132.19)	(132.08)	(136.33)	(123.92)
Reading grade 9 (*M*(*SD*))	125.25	98.64	136.65	133.53
	(107.92)	(107.20)	(110.62)	(102.17)
Student background				
Parental education grade 5 (*M*(*SD*))	2.68	2.47	2.77	2.69
	(1.13)	(1.11)	(1.10)	(1.16)
Language of origin German grade 5	88.0%	78.9%	91.3%	90.4%
School characteristics	*N*=144	*N*=43	*N*=56	*N*=45
School type				
Hauptschule/Volksschule	27.8%	43.4%	17.7%	21.2%
School with several courses of education	13.8%	12.6%	12.0%	16.3%
Realschule	17.3%	16.6%	13.6%	20.9%
Gesamtschule	8.5%	7.6%	11.7%	6.9%
Gymnasium	32.6%	19.8%	45.0%	34.6%
School composition				
Social composition grade 5 (*M*(*SD*))	2.60	2.35	2.77	2.71
	(0.54)	(0.49)	(0.54)	(0.49)
Language composition grade 5 (*M*(*SD*))	0.85	0.76	0.92	0.87
	(0.21)	(0.25)	(0.12)	(0.19)

Note. The descriptive analyses were based on imputed data of *N*=3024 students at 144 schools; "w_t_cal" was used as a sampling weight. Unstandardized social and language composition variables were used.

Parental Education and Language of Origin

We investigated two facets of student background, which were both assessed in computer-assisted telephone interviews with parents in grade 5 (see Table 1). We operationalized social background as parental education level in line with the International Standard Classification of Education. Information on parents' highest general educational qualification were categorized from 0 ("level 0/1A: Inadequately completed general education") to 5 ("level 6: Doctoral degree and postdoctoral lecture qualification"). We treated this variable as continuous. We operationalized immigrant background by the student's language of origin. Answers to the question regarding the language students learned in the first three years in the family were dichotomized to 0 ("other than German") and 1 ("German").

School Organization: Half- and All-Day Schools

The main explanatory variable was school organization form. In questionnaires in grades 5, 7, and 9, principals were asked whether their school was a half- or all-day school. In the questionnaires, all-day schools were further categorized into nonmandatory, partially mandatory, and fully mandatory all-day schools. The sample sizes were, however, too small to estimate the effects of schools that remained fully mandatory over time, for example. For this reason, we subsume them as all-day schools. Due to the massive investment in all-day schooling, several schools converted from half- to all-day schools in the period of investigation, while some also changed from all- to half-day schools (see also KMK, 2018; Steiner, 2011). For a clearer interpretation of our main explanatory variable, we categorized such schools as a separate group in our analyses. The sample covered 43 half-day schools, 56 all-day schools, and 45 schools with a mixed status over time. The organization form was analyzed as two dummy variables, with half-day schools as the reference category. Table 1 depicts the characteristics of these groups and shows that half-day school students constituted a less privileged group than those attending all-day schools.

Covariates

Germany has a stratified secondary school system with different ability tracks. The school types corresponding to those tracks were part of the explicit strata in the NEPS sampling design and were used as covariates in the form of four dummy variables ("school with several courses of education", "Realschule", "Gesamtschule", and "Gymnasium", with "Hauptschule/Volksschule" as reference). To control for differences in student composition, we aggregated the student background information on parental education (social composition) and the language of origin (language composition) at school level. Both variables were treated as continuous variables in the analyses (see Table 1).

Missing Value Imputation

The dataset used in this analysis contained missing data (see Appendices A1 and A2). Twenty schools whose principals did not respond to the organization form question at any of the three measurement points were excluded from the analyses because we regarded the variable base as too poor for sound imputation. Apart from this, we imputed missing values five times by using two-level predictive mean matching, which is a simple extension of ordinary multiple imputation technique for non-clustered data (van Buuren & Groothuis-Oudshoorn, 2011). In order to replicate the data structure, we included the sampling weight as an imputation predictor as well as a rich set of further variables (see Appendix A2): all aforementioned variables, complementary variables, their repetitions in other waves, and counterparts in other instruments. The Appendix A2 also depicts which variables lied on school and student levels in the two-level imputation. We reran all analyses for the five imputed datasets and combined the estimates using Rubin's rules (1987).

Analyses

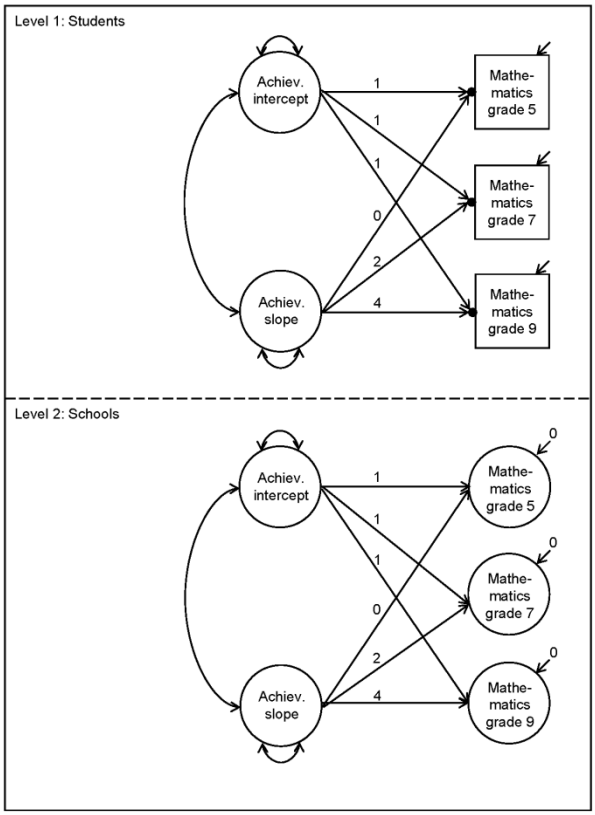
To explain our general analytical approach, we will describe how we modeled achievement growth and changes in inequality in achievement before outlining how we tested for differ-

ences between half- and all-day schools. All analyses were replicated for mathematics and reading in separate models.

Modeling Effects on Student Achievement

Two-level linear latent growth modeling was used to investigate achievement growth. At the student level, achievement scores in grades 5, 7, and 9 were used to model an achievement intercept for achievement in grade 5 and an achievement slope for the annual growth in achievement (see Figure 1). As in the student-level models, for the school-level models the schools' achievement scores in grades 5, 7, and 9 were used to model an achievement intercept and slope. The school level achievement intercept reflects that schools could vary in their initial achievement in grade 5 and the achievement slope indicates that the schools could differ in their growth rates.

Figure 1. Measurement model for latent growth in mathematics achievement



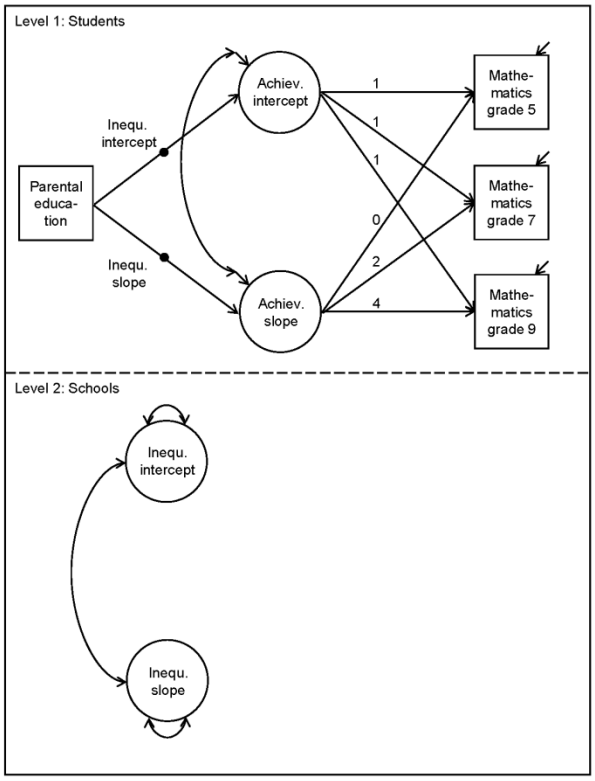
Note. The two-level latent growth model shown in this figure was estimated; the dots in the individual-level graph represent random intercepts; the random intercepts are shown in circles in the school-level graph because they are continuous latent variables that vary across schools; school-level residuals are constrained to zero to avoid negative variance. The school-level achievement intercept and slope serve as dependent variables in further analyses.

To answer the first research question on the effects of organization form on student achievement, we regressed both school achievement intercept and slope on the dummies for the organization form. The key parameter of interests is the achievement slope, because we were primarily interested in effects on learning progress. The main advantage of using longitudinal modeling is that it reduced the risk of confounding variables biasing the effect estimation of interest. Even with longitudinal data, confounding variables may bias the analyses. To further minimize the risk of confounding variables, we also controlled for school type and composition.

Modeling Effects on Inequality Regarding Parental Education

To model social inequalities within schools, we extended the previous student level model by regressing the achievement intercept and slope on parental education (see Figure 2). We call the within-school association between the achievement intercept and parental education the inequality intercept, and the within-school association between the achievement slope and parental education the inequality slope. The inequality intercept and slope were modeled as random parameters.

Figure 2. Measurement model for latent growth in social inequality in mathematics achievement



Note. The two-level latent growth model shown in this figure was estimated; the dots in the individual-level graph represent random parameters that vary across schools; the inequality intercept represents the achievement gap be-

tween students from more and less educated parents in grade 5 and it varies across schools; the inequality slope is the gap in achievement growth between students from highly and little educated parents and it varies across schools. The school-level achievement intercept and slope serve as dependent variables in further analyses.

To answer the second research question on the impact of half- versus all-day schools on educational inequality, we regressed the school level inequality intercept and inequality slope on the school organization dummy variables. The first parameter reflects differences between half- and all-day schools in the initial degree of social inequality in achievement in grade 5. The second parameter is the estimate for differences between school organization forms in terms of changes in social inequality in achievement. We were mainly interested in the second parameter. Again, we also included school controls (see previous section).

Modeling Effects on Inequality Regarding Language of Origin

In order to answer the third research question, we replaced parental education with language of origin and replicated the analyses described in the previous section. The resulting model investigated initial achievement gaps and gaps in achievement growth between students with and without German as a first language.

Results

Effects on Student Achievement

Table 2 summarizes the results of the latent growth analyses on student achievement for models with and without control variables for mathematics and reading. The column 1 shows that the mathematics achievement level in grade 5 was significantly higher in all-day schools than in half-day schools. We also observed higher performance levels in schools that changed organization form, but the difference was not significant and should not be over-interpreted because of the difficulty in interpreting the status of these mixed schools. However, the observed differences at the beginning of secondary school may be due to selection effects. The more important estimate of the effect of the organization form on achievement is the parameter of the regression of the achievement slope on the organization form (column 2). The key result is that there is no evidence of differing achievement growth rates in half- and all-day schools.

The observed large difference in the achievement intercept vanished after controlling for key school covariates (column 3). This finding confirms the presumption that differences in initial achievement levels were mainly due to differences in school intake. Importantly, however, the effects of organization form on the achievement slope did not change after controlling for key covariates (column 4).

The findings for reading replicated those for mathematics. Initially observed differences in performances levels in grade 5 were no longer significant when further school characteristics were controlled for. Furthermore, the achievement slopes were unrelated to the organization form.

Table 2. Predictors of growth in achievement at half- and all-day schools

	Mathematics				Reading			
	Model 1a		Model 1b		Model 1c		Model 1d	
	Achiev. intercept (1)	Achiev. slope (2)	Achiev. intercept (3)	Achiev. slope (4)	Achiev. intercept (5)	Achiev. slope (6)	Achiev. intercept (7)	Achiev. slope (8)
All-day school[1]	52.32*	-0.16	12.81	-0.28	46.30*	-1.76	6.86	-0.34
	(19.58)	(1.85)	(9.11)	(1.88)	(20.15)	(1.65)	(11.24)	(1.80)
Changed status[1]	47.83	0.82	13.47	1.16	48.99	-2.00	17.22	-0.57
	(27.86)	(1.91)	(11.34)	(2.05)	(28.80)	(1.85)	(12.08)	(1.97)
School with several courses of education[2]	–	–	44.06*	-2.85	–	–	28.55	1.96
			(16.26)	(3.16)			(22.69)	(3.72)
Realschule[2]	–	–	68.89*	-5.59*	–	–	58.16*	-2.46
			(13.19)	(2.54)			(19.04)	(2.95)
Gesamtschule[2]	–	–	37.17	-0.78	–	–	30.13	3.28
			(23.93)	(3.86)			(26.66)	(3.97)
Gymnasium[2]	–	–	135.64*	-3.15	–	–	109.10*	-3.57
			(25.09)	(3.10)			(29.45)	(4.06)
Social composition	–	–	19.83	0.66	–	–	26.73*	-0.77
			(10.51)	(1.19)			(9.42)	(1.40)
Language composition	–	–	7.89*	-1.43*	–	–	3.64	-0.79
			(3.85)	(0.70)			(3.59)	(0.66)

Note. The two-level analyses were based on imputed data of N=3024 students at 144 schools; "w_t_cal" was used as a sampling weight; maximum likelihood estimation with robust standard errors; unstandardized parameters with standard errors in parentheses; * statistically significant at $p<0.05$. [1] The reference category was "half-day school". [2] The reference category was "Hauptschule/Volksschule". Standardized social and language composition variables were used in order to improve the interpretability of the results.

Effects on Inequality Regarding Parental Education

Table 3 summarizes the results of the analyses on social inequality in student achievement for models with and without control variables for mathematics and reading. The column 1 shows that the degree of social inequality in grade 5 was significantly higher at all-day schools than at half-day schools. However, this difference at the beginning of secondary school may be due to selection effects. The more important estimate for effects on inequality is the parameter of the regression of the inequality slope on the organization form (column 2). The key result is that there is no evidence of differing inequality growth rates in half- and all-day schools.

The observed difference in the inequality intercept was no longer significant when controlling for school covariates (column 3), which supports the assumption that initial inequality differences between schools were mainly due to differences in school intake. Importantly, the null effects of organization form on the inequality slope remained stable when controlling for school covariates (column 4).

The reading analyses produced equivalent findings. Differences between inequality levels in grade 5 for half- and all-day schools were no longer significant when school controls were included in the analyses. The inequality slopes did not significantly differ between half- and all-day schools.

Table 3. Predictors of growth in achievement inequality related to parental education at half- and all-day schools

	Mathematics				Reading			
	Model 2a		Model 2b		Model 2c		Model 2d	
	Inequ. intercept (1)	Inequ. slope (2)	Inequ. intercept (3)	Inequ. slope (4)	Inequ. intercept (5)	Inequ. slope (6)	Inequ. intercept (7)	Inequ. slope (8)
All-day school[1]	19.43*	-1.25	4.39	-0.08	17.25*	-1.58	2.36	-0.17
	(7.61)	(0.90)	(3.00)	(0.82)	(7.91)	(0.82)	(3.90)	(0.75)
Changed status[1]	18.86	-1.01	5.06	0.14	18.41	-1.55	5.49	-0.41
	(10.30)	(0.86)	(4.26)	(0.73)	(11.04)	(1.01)	(4.71)	(0.76)
School with several courses of education[2]	–	–	20.92*	-1.73	–	–	15.06*	0.17
			(5.99)	(1.28)			(5.91)	(1.15)
Realschule[2]	–	–	29.79*	-2.74*	–	–	26.58*	-1.39
			(5.27)	(1.20)			(4.88)	(0.96)
Gesamtschule[2]	–	–	18.23*	-0.66	–	–	18.23*	0.45
			(7.71)	(1.66)			(7.72)	(1.29)
Gymnasium[2]	–	–	48.97*	-1.82	–	–	41.45*	-1.65
			(6.07)	(1.14)			(6.54)	(1.13)
Social composition	–	–	7.30*	-1.49*	–	–	9.08*	1.48*
			(2.01)	(0.38)			(2.23)	(0.43)
Language composition	–	–	3.43*	-0.15	–	–	3.45*	-0.30
			(1.54)	(0.34)			(1.64)	(0.29)

Note. The two-level analyses were based on imputed data of *N*=3024 students at 144 schools; "w_t_cal" was used as a sampling weight; maximum likelihood estimation with robust standard errors; unstandardized parameters with standard errors in parentheses; * statistically significant at $p<0.05$. [1] The reference category was "half-day school". [2] The reference category was "Hauptschule/Volksschule". Standardized social and language composition variables were used in order to improve the interpretability of the results.

Effects on Inequality Regarding Language of Origin

Table 4 summarizes the findings on language-related inequalities in mathematics and reading from models with and without controls. The analyses on the effects of school organization form on inequalities related to students' language of origin were identical to the analyses on social inequalities except that parental education was replaced with language of origin. The analyses basically replicated the results for parental education (see previous section). The findings did not support the assumption that all-day schools would differ from half-day schools with respect to the relationship between achievement growth and language of origin (see Table 4). There were no significant differences in language-related inequality levels in grade 5 between half- and all-day schools, when controlling for school type and composition.

Table 4. Predictors of growth in achievement inequality related to language of origin at half- and all-day schools

	Mathematics				Reading			
	Model 3a		Model 3b		Model 3c		Model 3d	
	Inequ. intercept (1)	Inequ. slope (2)	Inequ. intercept (3)	Inequ. slope (4)	Inequ. intercept (5)	Inequ. slope (6)	Inequ. intercept (7)	Inequ. slope (8)
All-day school[1]	51.47*	-0.33	11.70	-0.43	42.90*	-1.39	4.22	-0.10
	(19.10)	(2.02)	(9.33)	(1.99)	(20.70)	(1.84)	(11.49)	(1.92)
Changed status[1]	47.61	0.54	11.86	1.12	44.90	-1.32	11.57	-0.16
	(27.41)	(2.03)	(11.55)	(2.07)	(28.70)	(1.90)	(12.60)	(2.00)
School with several	–	–	44.61*	-3.21	–	–	30.00	2.08
courses of education[2]			(15.32)	(3.16)			(15.48)	(3.06)
Realschule[2]	–	–	68.08*	-5.09	–	–	58.99*	-1.96
			(12.04)	(2.63)			(13.10)	(2.49)
Gesamtschule[2]	–	–	32.87	1.90	–	–	37.03	4.58
			(19.88)	(3.62)			(21.75)	(3.44)
Gymnasium[2]	–	–	132.39*	-2.34	–	–	108.63*	-1.97
			(17.09)	(2.92)			(18.89)	(3.04)
Social composition	–	–	21.16*	0.14	–	–	26.36*	-1.34
			(5.75)	(1.07)			(5.88)	(0.97)
Language composition	–	–	5.60	-0.66	–	–	4.26	-0.43
			(3.90)	(0.78)			(3.50)	(0.74)

Note. The two-level analyses were based on imputed data of N=3024 students at 144 schools; "w_t_cal" was used as a sampling weight; maximum likelihood estimation with robust standard errors; unstandardized parameters with standard errors in parentheses; * statistically significant at $p<0.05$. [1] The reference category was "half-day school". [2] The reference category was "Hauptschule/Volksschule". Standardized social and language composition variables were used in order to improve the interpretability of the results.

Robustness Analyses

Student Participation Profiles

Previous research showed that student participation in all-day programs varies in many regards. For example, not all all-day school students attend programs—especially not academic programs—and if they do, they may not spend substantial amounts of time in them (e.g., StEG, 2016). It could be assumed that all-day schools with more preferable participation characteristics had achievement-increasing or inequality-decreasing effects. Therefore, using ranking and median-splitting, we divided the 56 all-day schools into two groups based upon eight participation variables (see Appendix A3). In additional analyses, we reran all models for mathematics and reading with the new dummy variables for low- and high-profile all-day schools and mixed schools, using half-day schools as reference. The results showed that high-profile all-day schools did not show higher achievement growth or lower inequality gradients than half-day schools.

Alternative Social Inequality Indicators

In the main analyses, we chose parental education and language of origin (both assessed in parent interviews) as student background indicators. In order to increase the generalizability of our findings, we reran models 2b, 2d, 3b, and 3d for mathematics and reading with alternative indicators. These indicators were books at home and countries of birth, and both were assessed in the student questionnaires in grade 5. The results were qualitatively identical to the main findings.

Federal States

Although previous research documents that the federal states conduct all-day schools differently (e.g., KMK, 2015), we did not control for this in the main analyses because the samples from some states included only few schools. In additional analyses, we extended models 1b, 1d, 2b, 2d, 3b, and 3d for mathematics and reading by adding dummy variables for federal states. The results were qualitatively the same as the main findings.

Schools with Missing Information on Organization Form

We excluded 20 schools whose principals did not provide information on the organization form in the main analyses because we only had limited information to impute the missing data (see section 4.2.5). At the same time, we acknowledge that the exclusion of schools limits the generalizability of our findings. To address this issue, we replicated all analyses for the full sample of N=3444 students at 164 schools with imputed missing data on the organization form based on the limited available data. Again, the results were qualitatively the same as the main findings.

Discussion

The present study investigated whether the school organization form—i.e., half- and all-day schools—affected the development of student achievement and educational inequalities. The longitudinal comparison of schools that continued to offer half-day schooling versus those that offered all-day schooling between 2010 and 2015 in two-level latent growth curve models showed no evidence supporting the assumption that organization form influenced mathematics or reading achievement growth and changes in educational inequality in the course of secondary schooling. Neither the achievement slopes nor the inequality slopes (related to social or language background) differed significantly between half- and all-day schools.

We found, however, that half- and all-day schools operated under different conditions. Simple models indicated higher initial levels of achievement and educational inequalities at all-day schools than at half-day schools. These differences were no longer significant when controlling for further school characteristics. Therefore, the differences in the achievement and inequality intercepts seemed to result from selection effects. For example, in the tracked German school system, students are allocated to different school types after primary school. This results in both performance-related and social segregation between school types (e.g., Maaz, Trautwein, Lüdtke, & Baumert, 2008). In the investigated sample, the most selective school type "Gymnasium" was overrepresented among all-day schools.

The fact that our study did not find significant differences between half- and all-day schools in terms of achievement and inequality development is in line with previous studies (Linberg et al., 2018; Steinmann & Strietholt, in print; Strietholt et al., 2015). One possible explanation for the absence of an effect is the quality of all-day schools in Germany. The literature suggests that international extended education programs should, for example, be aligned with learning goals and supervised by teachers to be effective, i.e. they should be rather formal than non-formal learning opportunities (e.g., Durlak et al., 2010; Miller & Truong, 2009). Previous research on the quality of all-day schools suggests that this is not the case in Germany (e.g., StEG, 2013, 2015, 2016). Indeed, the programs offered by all-day schools in the present sample were on average attended for less than three hours per week. The average shares of students who attended the academic all-day programs offering homework support, remedial education, or subject-specific programs were 15% or lower.

Scope

The study extends the existing research in three important regards. Most importantly, this is the first study that compared the developments of half- and all-day schools over a period of four years. The first of the three previous robust studies used cross-sectional data (Strietholt et al., 2015), the second had a one-year longitudinal section (Steinmann & Strietholt, in print), and the third had a two-year longitudinal section (Linberg et al., 2018). A second important benefit is that the present study compared schools that remained half- or all-day schools over four years and therefore took into account the changes in organizational structure undertaken by many schools in the investigated time span. Indeed, 31% of the schools reported a mixed organization structure over time. In the previous studies, organization form was assessed only once (Linberg et al., 2018; Steinmann & Strietholt, in print; Strietholt et al., 2015). Therefore, the all-day schools might have changed their organization structure immediately before or after this single point of assessment. Such schools might not have made enough progress in the school development process of becoming all-day schools in order to be fully effective (cf. StEG, 2013, 2015). This explanation for the null effects of all-day schools is less likely in the present study. Third, the study minimizes the risk of the regression-towards-the-mean phenomenon because it uses three measurement points for student achievement.

Apart from these advantages, the scope of the present study has certain limitations. Importantly, we only investigated effects on student achievement development and social disparities in this development. Needless to say, these are not the only important outcomes of extended education. For example, all-day schooling has been found to have beneficial effects on students' psychosocial development (StEG, 2016). Additionally, our analytical sample consisted of students who did not change schools between grades 5–9, which limits the generalizability of our results. Another major issue is that we simply compared half- and all-day schools, although previous research has shown that the quality of actual extended education at all-day schools is rather low. The data we used only allowed a further investigation of mediators to a limited extent. In additional analyses, we found that all-day schools with favorable student participation characteristics did not have significantly different achievement or inequality developments than half-day schools either. It was, however, not possible to investigate the effects of all-day schools with mandatory all-day participation for all students.

Future Research

We derive some conclusions for future research from the discussed scope of the present and previous studies. Given that U.S. studies showed that high-quality extended education can increase student achievement, especially among disadvantaged students (e.g., Durlak et al., 2010; Lauer et al., 2006), more research is needed on the circumstances under which all-day schools in Germany can be effective. For example, one current project aims to develop a reading support program that can also be effectively delivered by staff without teacher training (DIPF, 2018). Generally, one reason why international findings on effective extended education programs cannot be directly transferred to all-day schools is that all-day schooling is a rather nonspecific intervention with diverse implications, as the schools are highly autonomous in how they organize their all-day program. Causal relationships between all-day schools' setups and resources, all-day program features, participation characteristics, and student outcomes should be examined in greater detail in future studies (cf. Fischer & Klieme, 2013; Steinmann, 2018; Vest et al., 2013).

From an international perspective, German all-day schooling is an interesting example of an up scaling of non-formal extended education. However, in line with previous research, our study suggests that all-day schools rather provide childcare than extended formal learning opportunities to support student achievement. The study therefore contributes to the international state of research by illustrating that the extension of time at school is not sufficient to extend formal learning opportunities. In fact, the lack of standardization and highly qualified staff for example seem to limit the effectiveness of German all-day schools. This conforms to findings on comparable afterschool programs and summer schools in the US (cf. Apsler, 2009; Durlak et al., 2010; Roth et al., 2010). From an overall perspective, these findings highlight that international comparisons need to pay particular attention to qualitative characteristics of extended education programs.

Conclusion

Our findings do not support the assumption that all-day schools make a difference for student achievement or educational inequality development in Germany. Given the discussed literature, we draw two tentative conclusions: First, all-day schools do not yet seem to be of sufficient quality to increase student achievement or to decrease social inequalities, at least not on a large scale (see StEG, 2015, 2016; Steinmann, 2018). Second, in order to attain the educational goals of higher achievement gains and lower social inequalities, policy makers may be advised to consider investing in the quality of all-day schools (cf. Fischer et al., 2014; Lossen et al., 2016; StEG, 2016).

References

Apsler, R. (2009). After-school programs for adolescents: A review of evaluation research. *Adolescence*, *44*, 1-19.

Bellin, N., & Tamke, F. (2010). Bessere Leistungen durch Teilnahme am offenen Ganztagsbetrieb? *Empirische Pädagogik*, *24*(2), 93-112.

Blossfeld, H.-P., Roßbach, H.-G., & von Maurice, J. (Eds.) (2011). *Education as a lifelong process – The German National Educational Panel Study (NEPS). Zeitschrift für Erziehungswissenschaft: Sonderheft 14.*

BMBF (2003). *Investitionsprogramm "Zukunft Bildung und Betreuung": Ganztagsschulen. Zeit für mehr*. Bonn: BMBF.

Cooper, H., Charlton, K., Valentine, J. C., Muhlenbruck, L., & Borman, G. D. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development*, *65*(1), 1-127.

DIPF (2018). *StEG-Lesen: Förderung der Lesekompetenz an Grundschulen durch innovative Arbeitsgemeinschaft im Ganztag*. Retrieved from https://www.projekt-steg.de/content/steg-lesen

Duchhardt, D., & Gerdes, A. (2012). *NEPS technical report for mathematics – scaling results of starting cohort 3 in fifth grade (NEPS Working Paper No. 19)*. Bamberg: Otto-Friedrich-Universität, NEPS.

Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American journal of community psychology*, *45*(3-4), 294-309.

Feldman, A. F., & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research*, *75*(2), 159-210.

Fischer, L., Rohm, T., Gnambs, T., & Carstensen, C. H. (2016). *Linking the data of the competence tests (NEPS Survey Paper No. 1)*. Bamberg: LIfBi, NEPS.

Fischer, N., & Klieme, E. (2013). Quality and effectiveness of German all-day schools. Results of the Study on the Development of All-Day Schools in Germany. In J. Ecarius, E. Klieme, L. Stecher, & J. Woods (Eds.), *Extended education – an international perspective: Proceedings of the international conference on extracurricular and out-of-school time educational research* (pp. 27-52). Opladen [i.a.]: Barbara Budrich Publishers.

Fischer, N., Sauerwein, M. N., Theis, D., & Wolgast, A. (2016). Vom Lesenlernen in der Ganztagsschule: Leisten Ganztagsangebote einen Beitrag zur Leseförderung am Beginn der Sekundarstufe I? *Zeitschrift für Pädagogik*, *62*(6), 780-796.

Fischer, N., Theis, D., & Züchner, I. (2014). Narrowing the gap? The role of all-day schools in reducing educational inequality in Germany. *International Journal for Research on Extended Education*, *2*(1), 79-96.

Frahm, S., Goy, M., Kowalski, K., Sixt, M., Strietholt, R., Blatt, I., Bos, W., & Kanders, M. (2011). Transition and development from lower secondary to upper secondary school. *Zeitschrift für Erziehungswissenschaft, 14*(special issue, supplement 2), 217-232.

KMK (2008). *Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2002 bis 2006*. Bonn: KMK.

KMK (2015). *Ganztagsschulen in Deutschland: Bericht der Kultusministerkonferenz vom 03.12.2015*. Berlin, Bonn: KMK.

KMK (2018). *Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2012 bis 2016*. Berlin: KMK.

Krannich, M., Jost, O., Rohm, T., Koller, I., Pohl, S., Haberkorn, K.,. . . Gnambs, T. (2017). *NEPS technical report for reading: Scaling results of starting cohort 3 for grade 7 (Update NEPS Survey Paper No. 14)*. Bamberg: *LIfBi*, NEPS.

Kuger, S. (2016). Curriculum and learning time in international school achievement studies. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Assessing Contexts of Learning* (pp. 395-422). Cham: Springer.

Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research*, *76*(2), 275-313.

Linberg, T., Struck, O., & Bäumer, T. (2018). Vorzug Ganztagsschule? Zusammenhänge mit der Kompetenzentwicklung im Bereich Lesen und Mathematik. *Zeitschrift für Erziehungswissenschaft*, *14*(2), 51.

Lossen, K., Tillmann, K., Holtappels, H. G., Rollett, W., & Hannemann, J. (2016). Entwicklung der naturwissenschaftlichen Kompetenzen und des sachunterrichtsbezogenen Selbstkonzepts bei Schüler/-innen in Ganztagsgrundschulen: Ergebnisse der Längsschnittstudie StEG-P zu Effekten der Schülerteilnahme und der Angebotsqualität. *Zeitschrift für Pädagogik*, *62*(6), 760-779.

Maaz, K., Trautwein, U., Lüdtke, O. & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives, 2*(2), 99-106.

Miller, B. M., & Truong, K. A. (2009). The role of afterschool and summer in achievement. The untapped power of afterschool and summer to advance student achievement. *Zeitschrift für Pädagogik*, *55*(54th supplement), 124-142.

Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the school day or school year: A systematic review of research (1985–2009). *Review of Educational Research*, *80*(3), 401-436.

Plantenga, J., & Remery, C. (2013). *Childcare services for school age children: A comparative review of 33 countries*. Luxembourg: Publications Office of the European Union.

Pohl, S., Haberkorn, K., Hardt, K., & Wiegand, E. (2012). *NEPS technical report for reading – scaling results of starting cohort 3 in fifth grade (NEPS Working Paper No. 15).* Bamberg: Otto-Friedrich-Universität, NEPS.

Roth, J. L., Malone, L. M., & Brooks-Gunn, J. (2010). Does the amount of participation in afterschool programs relate to developmental outcomes? A review of the literature. *American journal of community psychology*, *45*(3-4), 310-324.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

Scharl, A., Fischer, L., Gnambs, T., & Rohm, T. (2017). *NEPS technical report for reading: Scaling results of starting cohort 3 for grade 9 (NEPS Survey Paper No. 20).* Bamberg: *LIfBi*, NEPS.

Schnittjer, I., & Gerken, A.-L. (2017). *NEPS technical report for mathematics: Scaling results of starting cohort 3 in grade 7 (NEPS Survey Paper No. 16).* Bamberg: *LIfBi*, NEPS.

Stecher, L., & Maschke, S. (2013). Research on extended education in Germany – A general model with all-day schooling and private tutoring as two examples. *International Journal for Research on Extended Education, 1*(1), 31-52.

StEG (2013). *Ganztagsschule 2012/2013: Deskriptive Befunde einer bundesweiten Befragung*. Frankfurt a. M. [i.a.]: DIPF [i.a.].

StEG (2015). *Ganztagsschule 2014/2015: Deskriptive Befunde einer bundesweiten Befragung*. Frankfurt a. M. [i.a.]: DIPF [i.a.].

StEG (2016). *Ganztagsschule: Bildungsqualität und Wirkungen außerunterrichtlicher Angebote: Ergebnisse der Studie zur Entwicklung von Ganztagsschulen 2012–2015*. Frankfurt a. M. [i.a.]: DIPF [i.a.].

Steiner, C. (2009). Mehr Chancengleichheit durch die Ganztagsschule? *Zeitschrift für Pädagogik*, *55*(54th supplement), 81-105.

Steiner, C. (2011). Teilnahme am Ganztagsbetrieb. Zeitliche Entwicklung und mögliche Selektionseffekte. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher, & I. Züchner (Eds.), *Ganztagsschule: Entwicklung, Qualität, Wirkungen: Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG)* (pp. 57-75). Weinheim [i.a.]: Beltz Juventa.

Steinmann, I. (2018). Ganztagsschulen als neue Organisationsformen – Entwicklungen und Evaluation. *Enzyklopädie Erziehungswissenschaft Online*. doi: 10.3262/EEO08180405.

Steinmann, I., & Strietholt, R. (in print). Effekte nachmittäglicher Bildungsangebote auf sprachliche Kompetenzentwicklungen. *Zeitschrift für Pädagogik.*

Steinmann, I., Strietholt, R., & Caro, D. (2018). Participation in extracurricular activities and student achievement: Evidence from German all-day schools. *School Effectiveness and School Improvement.* doi 10.1080/09243453.2018.1540435.

Strietholt, R., Manitius, V., Berkemeyer, N., & Bos, W. (2015). Bildung und Bildungsungleichheit an Halb- und Ganztagsschulen. *Zeitschrift für Erziehungswissenschaft, 18*(4), 737-761.

Strietholt, R., Naujokat, K., Mai, T., Kretschmer, S., Jarsinski, S., Goy, M., Frahm, S., Kanders, M., Bos, W., & Blatt, I. (2013). The National Educational Panel Study (NEPS) in Germany. An overview of design, research options, and access, with a focus on lower secondary school. *European Educational Research Journal, 12*(4), 568-579.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1-67.

Vest, A. E., Mahoney, J. L., & Simpkins, S. D. (2013). Patterns of out-of-school time use around the world: Do they help to explain international differences in mathematics and science achievement. *International Journal for Research on Extended Education*, *1*(1), 71-85.

Werquin, P. (2010). *Recognising non-formal and informal learning: Outcomes, policies and practices*. Paris: OECD.

Züchner, I., Fischer, N. (2014). Kompensatorische Wirkungen von Ganztagsschulen – Ist die Ganztagsschule ein Instrument zur Entkopplung des Zusammenhangs von sozialer Herkunft und Bildungserfolg? *Zeitschrift für Erziehungswissenschaft, 17*, 349-367.

Appendix A1. Descriptive statistics of student and school characteristics prior to imputation

	All schools	**Divided by school organization form**			
		Half-day school	**All-day school**	**Changed status**	**Missing**[1]
Student characteristics	*N*=3024	*N*=297	*N*=744	*N*=433	*N*=1550
Student achievement					
Mathematics grade 5 (*M*(*SD*))	-1.35	2.86	23.34	11.74	-14.51
	(114.89)	(119.56)	(107.78)	(100.68)	(118.33)
Mathematics grade 7 (*M*(*SD*))	75.32	75.23	106.10	89.47	59.90
	(122.51)	(132.76)	(112.20)	(119.53)	(122.44)
Mathematics grade 9 (*M*(*SD*))	152.13	148.21	176.72	176.41	136.05
	(117.12)	(118.74)	(113.08)	(101.60)	(120.15)
Reading grade 5 (*M*(*SD*))	1.82	1.71	20.61	13.19	-8.02
	(122.81)	(129.18)	(114.26)	(110.75)	(126.73)
Reading grade 7 (*M*(*SD*))	70.96	62.93	82.02	85.41	63.89
	(131.81)	(131.71)	(134.78)	(118.14)	(134.18)
Reading grade 9 (*M*(*SD*))	127.09	124.14	138.55	133.02	121.65
	(107.76)	(106.72)	(106.52)	(95.82)	(111.62)
Student background					
Parental education grade 5 (*M*(*SD*))	2.81	2.67	3.02	2.87	2.73
	(1.10)	(1.02)	(1.06)	(1.11)	(1.11)
Language of origin German grade 5	92.1%	90.2%	93.9%	95.4%	90.5%
School characteristics	*N*=144	*N*=18	*N*=29	*N*=19	*N*=78
School type					
Hauptschule/Volksschule	27.8%	22.3%	12.7%	3.7%	37.5%
School with several courses of education	13.8%	4.9%	6.6%	9.0%	17.8%
Realschule	17.3%	47.5%	14.8%	39.2%	8.0%
Gesamtschule	8.5%	0.0%	17.0%	15.4%	6.4%
Gymnasium	32.6%	25.3%	49.0%	32.6%	30.2%
School composition					
Social composition grade 5 (*M*(*SD*))	2.59	2.66	2.92	2.89	2.43
	(0.61)	(0.45)	(0.51)	(0.34)	(0.65)
Language composition grade 5 (*M*(*SD*))	0.79	0.90	0.94	0.94	0.71
	(0.34)	(0.18)	(0.07)	(0.12)	(0.40)

Note. The descriptive analyses were based on non-imputed data of *N*=3024 students at 144 schools; "w_t_cal" was used as a sampling weight. [1] Students at schools where school organization form information is missing for at least one out of three measurement points. Unstandardized social and language composition variables were used.

Appendix A2. Variables in the imputation model on the student and school level: Sources, measurement points, and percentages of missing values

	Source	2010/11	2011/12	2012/13	2013/14	2014/15
Replication weight	cohort profile	0.0%				
School level variables						
School type	cohort profile	0.0%				
Federal state	cohort profile	0.0%				
Social composition	aggregated student data	0.6%				
Language composition	aggregated student data	0.6%				
Cross-sectional organization form variable	principal quest.	35.9%		24.4%		39.3%
School offer of						
Homework supervision	principal quest.		15.5%		36.3%	
Enrichment groups: students with high grades	principal quest.		18.5%		36.6%	
Remedial instruction: students with low grades	principal quest.		16.6%		35.9%	
Remedial instruction: non-native speakers	principal quest.		17.4%		37.7%	
Language of origin instruction	principal quest.		17.4%		37.9%	
Subject-specific offering in mathematics	principal quest.		19.5%		39.5%	
Subject-specific offers in science	principal quest.		17.8%		38.1%	
Subject-specific offers in German	principal quest.		19.8%		41.1%	
Subject-specific offers in foreign languages	principal quest.		18.2%		40.2%	
Sports offers	principal quest.		17.6%		38.8%	
Music/art offers	principal quest.		16.0%		37.7%	
Religion offers	principal quest.		24.0%		43.0%	
Trades and home economics offers	principal quest.		16.7%		38.6%	
Technology/new media offers	principal quest.		16.4%		36.0%	
Community activities/student government	principal quest.		15.5%		37.8%	
Forms of social learning	principal quest.		15.9%		38.1%	
Forms of intercultural learning	principal quest.		15.3%		43.4%	
Required free-time activities	principal quest.		16.0%		38.7%	
Voluntary free-time activities	principal quest.		15.7%		38.9%	
All-day provision for grade 8	principal quest.				35.5%	
Number of staff in all-day program						
Without definite university degree	principal quest.				56.2%	
With university degree	principal quest.				61.0%	
No. days all-day progr. in 8th and/or 12th grade	principal quest.				38.9%	
Student level variables						
Student achievement						
Mathematics	student test	8.9%		4.8%		7.1%
Reading	student test	9.0%		4.9%		15.1%
Orthography	student test	8.7%		4.8%		7.4%
Perceptual speed	student test	8.9%				11.8%
Cognitive reasoning	student test	9.1%				11.5%
Declarative metacognition	student test		4.4%			11.2%
ICT literacy	student test		5.3%			7.3%
Scientific competence	student test		5.3%			7.3%
Reading speed	student test	8.7%				7.1%

	Source	2010/11	2011/12	2012/13	2013/14	2014/15
Time between tests	student test			13.5%		11.2%
Student participation in						
Homework support	student quest.		10.8%		13.4%	
Remedial education	student quest.		11.4%		13.7%	
Subject-specific programs	student quest.		12.2%		14.7%	
Subject-unrelated projects	student quest.		10.9%		14.4%	
Leisure facilities	student quest.		12.1%		14.9%	
Participation frequency	student quest.		16.8%		24.4%	
Student rating of all-day program						
Enjoyment	student quest.		35.3%		55.4%	
Useful in class	student quest.		36.4%		56.1%	
Wish for more offers	student quest.		36.4%		56.2%	
Learning new things	student quest.		37.1%		56.5%	
Preferred free time	student quest.		37.3%		56.7%	
Made new friends	student quest.		37.3%		56.7%	
Improving grades	student quest.		37.4%		56.7%	
Not alone in the afternoons	student quest.		37.5%		56.8%	
Student background						
Books at home	student quest.	14.2%		7.9%		8.0%
Parental education	parent interv.	37.9%		43.3%		55.0%
Partner parental education	parent interv.	50.3%		49.6%		60.2%
Family native-born	student quest.	8.2%		8.2%		
Language of origin	parent interv.	37.1%		37.1%		
Gender	cohort profile	4.8%		1.3%		
Age in years	cohort profile	12.1%		7.3%		

Appendix A3. Descriptive statistics for the subsamples of all-day schools with high and low student participation profiles for robustness analyses

	All-day schools	Divided by student participation characteristics	
		Low profile all-day school	High profile all-day school
School characteristics	*N*=56	*N*=28	*N*=28
Average student participation frequency[1]			
Average frequency grade 6 (*M*(*SD*))	2.47	1.79	2.97
	(1.48)	(0.78)	(1.66)
Average frequency grade 8 (*M*(*SD*))	1.61	1.19	1.92
	(1.38)	(0.72)	(1.64)
Percentage of students participating in programs[2]			
Homework support grade 6 (*M*(*SD*))	0.15	0.07	0.21
	(0.15)	(0.07)	(0.16)
Homework support grade 8 (*M*(*SD*))	0.09	0.04	0.12
	(0.08)	(0.04)	(0.09)
Remedial education grade 6 (*M*(*SD*))	0.13	0.08	0.17
	(0.11)	(0.08)	(0.12)
Remedial education grade 8 (*M*(*SD*))	0.10	0.05	0.14
	(0.08)	(0.05)	(0.09)
Subject-specific learning offers grade 6 (*M*(*SD*))	0.06	0.04	0.08
	(0.07)	(0.04)	(0.08)
Subject-specific learning offers grade 8 (*M*(*SD*))	0.09	0.05	0.11
	(0.08)	(0.05)	(0.09)

Note. The descriptive analyses were based on imputed data of *N*=1326 students at 56 all-day schools; "w_t_cal" was used as a sampling weight. All variables were taken from student questionnaires. [1] Aggregated variable on hours per week students made use of all-day offers overall; open answering format. [2] Aggregated variables on the participation in specific all-day programs; binary answering format with 0 ("no participation") and 1 ("participation").