Labour Economics 47 (2017) 35–47

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

How does education improve cognitive skills? Instructional time versus timing of instruction☆

Sarah C. Dahmann[a,b,*]

[a] School of Economics, The University of Sydney, Australia
[b] ARC Centre of Excellence for Children and Families over the Life Course, Australia

ARTICLE INFO

JEL:
I21
I24
I28
J24

Keywords:
Cognitive skills
Crystallized intelligence
Fluid intelligence
Skill formation
Education
High school reform
Gender skill gap

ABSTRACT

This paper investigates two mechanisms through which education may affect cognitive skills in adolescence, exploiting a school reform carried out at the state level in Germany as a quasi-natural experiment to identify causal effects: between 2001 and 2007, years at academic-track high school were reduced by one, leaving the overall curriculum unchanged. First, I exploit the variation over time and across states to identify the effect of an increase in class hours on same-aged students' intelligence scores, using data on seventeen year-olds from the German Socio-Economic Panel. Second, I investigate the influence of earlier instruction at younger ages, using data from the German National Educational Panel Study on high school graduates' competences. The results show that, on average, neither instructional time nor age-distinct timing of instruction significantly improves students' crystallized cognitive skills in adolescence. Yet, there is suggestive evidence that increasing instructional time may benefit male students exacerbating gender differences in numeracy.

1. Introduction

Cognitive skills are important determinants of many economic and social outcomes. At a macro level, cognitive skills in a population are strongly related to a country's economic growth (Hanushek, 2008). At a micro level, higher cognitive skills are associated with, among others, increased health and better old-age functioning mental abilities, and they are also linked to higher wages (see e.g. Heckman et al., 2006, or Heineck and Anger, 2010) and better education. The latter association is, however, a two-way relationship. On the one hand, individuals with higher cognitive abilities are likely to be better educated as they choose more often to continue education or easier meet access requirements. On the other hand, education itself also improves cognitive skills. Most studies use changes in compulsory schooling laws as an exogenous variation to identify causal positive effects of an additional year of schooling on cognition (e.g. Banks and Mazzonna, 2012). However, they do not provide evidence on the underlying mechanisms. This paper, therefore, investigates the roles of instructional time and timing of instruction as two potentially important, but not exclusive, channels through which secondary education may affect cognitive skills.

My research question is two-fold, which will be addressed in two analyses: First, I assess the impact of an increase in instructional time – dedicated to corresponding additional curriculum – on cognitive skills of adolescents in Germany. Second, I investigate whether the timing of instruction influences cognitive skill development, i.e. whether the allocation of class hours at a younger age changes cognitive skills, keeping the level of education constant.

To address these research questions, I exploit a reform in German high schools implemented between 2001 and 2007 that shortened total years of schooling from thirteen to twelve, leaving the overall curriculum unchanged. As a result, the number of weekly class hours significantly increased. Hence, while still in school, affected students

☆ This paper was largely written while I was at the German Institute for Economic Research (DIW Berlin) and I gratefully acknowledge funding from the German Academic Scholarship Foundation. I thank two anonymous reviewers as well as Silke Anger, Stijn Baert, Deborah Cobb-Clark, Friederike von Haaren, Susanne Kuger, Adam Lederer, Henning Lohmann, Bettina Siflinger, C. Katharina Spieß, Stefan C. Wolter, and seminar and conference participants at DIW Berlin, the Annual Conference of the European Association of Labour Economists, the Annual Conference of the European Economic Association, the Annual Conference of the European Society for Population Economics, the Annual Conference of the Verein für Socialpolitik, the Annual Conference of the Scottish Economic Society, the International Workshop of Applied Economics of Education, the IZA European Summer School in Labor Economics, the Trondheim Workshop on "Education, Skills, and Labor Market Outcomes", the International Young Scholar SOEP Symposium, the Spring Meeting of Young Economists, and the Essen Health Conference for helpful comments and discussions. I declare that I have no conflict of interest.
* Correspondence address: School of Economics, Merewether Building (H04), Faculty of Arts and Social Sciences, The University of Sydney, NSW 2006, Australia.
E-mail address: sarah.dahmann@sydney.edu.au.

http://dx.doi.org/10.1016/j.labeco.2017.04.008
Received 27 October 2016; Received in revised form 26 April 2017; Accepted 27 April 2017
Available online 03 May 2017

have covered a greater share of the overall curriculum than non-affected students of the same age. I use this intensified curriculum as an exogenous increase in the instructional quantity received up to the age of seventeen and exploit the variation over time and region in the implementation of the reform to identify its causal effect on adolescents' cognitive skills. Using rich data on adolescents from the German Socio-Economic Panel (SOEP) study, difference-in-differences estimates exhibit positive effects of the reform on crystallized and fluid intelligence, which, however, are not statistically significant on average. Tentative results suggest the impact may differ by gender: whereas male students' scores improved especially in numerical skills, female students' skills did not improve at all. These results may indicate the potential importance of instructional time as a mechanism in education improving cognitive skills, but also reveal its aggravating role in gender skill differences. I further use the variation in the age at which students received instruction as a quasi natural-experiment to investigate the impact of educational timing on students' competences. Using extensive data from the German National Educational Panel Study (NEPS) for the federal state of Baden-Wuerttemberg on students in their final grade, estimations suggest that the earlier knowledge transfer did not significantly alter the development of competences among students affected by the reform. Here, the potential benefit of early investment and age effects seem to offset each other. As a result, students affected by the reform catch up with their non-affected counterparts in terms of their competences by the time of graduation, apart from potential age effects resulting in slightly decreased fluid intelligence scores.

Next, I describe the theoretical background and existing literature. When explaining the high school reform in more detail, I elaborate on potential channels and anticipated effects. After a description of data and empirical strategy, I present the results. I test the robustness of the findings in several sensitivity analyses, before I conclude discussing the implications.

2. Theoretical basis and previous literature

Theoretical basis. Cognitive skills shape a variety of later-life outcomes. Together with non-cognitive skills, they form an important part of an individual's human capital as they constitute personal skills. A common approach to describe the formation and development of such skills is proposed by Cunha and Heckman (2007). They argue that an individual's present stock of skills depends on his or her past stock of skills, previous investment, and environmental factors. More specifically, they suggest the following model:

$$\theta_{t+1} = f_t(\theta_t, I_t, h), \tag{1}$$

where a vector of skill stocks at age $t+1$, θ_{t+1}, depends in some positive functional form $f(\cdot)$ on the past vector of skills (with initial endowment θ_1), on the investment in period t, I_t, and on parental, or more generally environmental, characteristics, h. In this model, Cunha and Heckman propose a multiplier effect driven by two mechanisms, *self-productivity* and *dynamic complementarity*. Self-productivity occurs whenever $\partial f_t(\theta_t, I_t, h)/\partial\theta_t > 0$. This implies that skills persist such that higher skills at one point in time create higher skills in the subsequent period, and is not restricted to one and the same skill but also includes cross effects between different skills. Dynamic complementarity occurs whenever $\partial^2 f_t(\theta_t, I_t, h)/\partial\theta_t\partial I_t' > 0$ implying that the productivity of investment is increasing with higher existing skills. Cunha and Heckman (2008) test and verify both propositions empirically. Hence, the resulting multiplier effect suggests that investments are most productive in early stages in life, making childhood the critical period for skill formation. Skills may, therefore, be malleable through e.g. educational interventions, especially at an early stage in life. However, there are important differences across dimensions of skills to distinguish.

Cognitive skills are usually distinguished into different facets. Two major ones in the empirical literature are fluid intelligence and crystallized intelligence.[1] Fluid intelligence relates to innate abilities that people are genetically endowed with. These include, for example, the ability to reason, the level of comprehension, or the capability of processing information, and are usually not influenced to a great extent by environmental factors. Crystallized intelligence, in contrast, denotes explicitly or implicitly learned knowledge or behavior. Therefore, it covers any specific knowledge of facts, for example, as well as learned behavioral traits such as the ability to read or calculate. Unlike fluid intelligence, crystallized intelligence is determined through environmental factors like education or upbringing. Several studies show that education indeed improves the crystallized component of cognitive skills, both in the short- and long-run.

Previous literature. For Scandinavian countries, several studies use data on males between the ages 18 and 20 from military cognitive assessment tests to identify short-term effects: Brinch and Galloway (2012) use an increase in compulsory schooling from seven to nine years in Norway between 1955 and 1972. Their difference-in-differences estimates and their instrumental variable results suggest positive returns and translate an additional year of schooling into an increase of 3.7 IQ points[2]. For Sweden, Carlsson et al. (2015) exploit a random variation in test dates to find that one additional year of schooling leads to an increase in crystallized intelligence of up to 0.21 standard deviations. Fluid intelligence does not seem to be affected by schooling, but rather positively by age. Instrumenting schooling and initial IQ, Falch and Massih (2011) find cognitive returns to one additional year of schooling between 2.9 and 3.8 IQ points for the Swedish population in Malmö that enrolled in the military in 1947 and 1948. Cascio and Lewis (2006) use data from the 1979 National Longitudinal Survey of Youth (NLSY79) to estimate returns to schooling on the Armed Forces Qualifying Test (AFQT) scores of males and females aged 15 to 19 years in the United States. Exploiting variation in the date of birth and school entry regulations, they find positive effects of 0.32 standard deviations; however only for racial and ethnic minorities. Setting up a regression discontinuity design to analyze the long-term effects of a compulsory schooling reform in England, Banks and Mazzonna (2012) find an increase in memory functioning of between 0.35 and 0.6 standard deviations among males and females older than 50. In turn, executive functioning only increased for males, with effect sizes ranging from 0.37 to 0.63 of a standard deviation. Using SHARE data on Austria, Czech Republic, Denmark, France, Germany, and Italy, Schneeweis et al. (2014) exploit the variation in compulsory schooling across the different countries to investigate cognitive ability of individuals older than 50. They find positive effects of 0.1 standard deviations of one additional year of schooling on memory functioning as well as some evidence on the reduction in cognitive decline in terms of verbal fluency through schooling. Furthermore, their effect sizes are generally larger in magnitude for males. Lastly, Kamhöfer and Schmitz (2016) investigate the long-term impact of education in Germany on word fluency among males and females born between 1940 and 1970 using data from the German Socio-Economic Panel (SOEP) study in 2006. They use different instruments for schooling to estimate local average treatment effects but find no effects. However, their outcome is limited in the sense that it is a single-edged view on cognition as it does not cover further dimensions of cognition next to word fluency, and that it is based on an ultra-short test that is conducted in only 90 seconds. Furthermore, weak instruments may be a threat to their identification, while age and cohort specific effects cannot be disentangled, which may confound their results.

Hence, with the exception of the study on Germany by Kamhöfer and Schmitz (2016), all studies clearly find substantial positive effects of an additional year of schooling on cognitive abilities. To establish effect causality, most of these analyses exploit a change in overall

[1] For a more detailed overview, see for example Baltes (1993) who describes fluid abilities as the *fluid-like mechanics of intelligence* and crystallized abilities as the *crystallized pragmatics of intelligence*.

[2] Note that generally IQ scales are defined to have mean 100 and standard deviation 15. An increase of 3.7 IQ points thus roughly corresponds to a 0.25 standard deviation increase.

school duration by one year. Still, the underlying mechanisms remain unresolved. However, for policy conclusions, it is critical to understand whether there are driving forces beyond overall school duration behind this relationship. While school duration, *per se*, cannot be changed infinitely, the existence of underlying channels would open new possibilities for decision makers to target cognitive ability when designing educational policies.

A change in school duration may have different consequences related to skill formation. On the one hand, an additional year of schooling may induce a larger curriculum to cover, i.e. constitute a direct increase in time and material of instruction. On the other hand, a change in the overall years of schooling may as well only lead to a redistribution of covered material and instruction over the different grades, i.e. over different age spans of the students. While the former constitutes a direct increase in investment I_t in Eq. (1), the latter implies a shift in the timing of investment I_t. Both may, therefore, impact cognitive skills: on the one hand, keeping age and past skills constant, an increase in investment, i.e. an increase in instructional quantity, may directly improve cognitive abilities. On the other hand, keeping overall instruction quantity constant, the age at which instruction for a given topic is received may influence cognitive abilities as well. Here, controversial mechanisms could interact, where earlier instruction is assumed to increase returns from later investments according to Cunha and Heckman (2007) and thereby improve cognitive skills, and because skills are more malleable at younger ages, but later instruction could benefit from maturity or time required to digest instruction.[3]

It therefore still remains to investigate whether either instructional time or timing of instruction drive the positive relationship between schooling and cognitive skills or whether it is both.[4] To the best of my knowledge, this study is the first to investigate and disentangle these two mechanisms. To identify causal effects, I use a unique variation in the German schooling system that allows me to conduct two separate analyses to provide a complete picture. The two analyses are complementary with each investigating a different mechanism and, therefore, distinct implications for policy makers. First, keeping age constant, the causal effect of an increase in instructional time is identified. Second, keeping the educational level constant, the role of instructional timing and age is analyzed. In addition, this study extends the literature on Germany and thereby may help to reconcile the zero findings of Kamhöfer and Schmitz (2016) with estimates from other countries that demonstrate positive cognitive returns to education. Furthermore, the rich datasets contain extensive tests of cognitive ability allowing for different cognitive dimensions to be distinguished. In addition, the inclusion of female respondents enables to investigate whether results in existing studies on adolescents found mainly among males extend to the overall population including both male and female students. Furthermore, the investigation of gender heterogeneity and understanding the influence of schooling in gender skill differences may help to develop adequate educational policies when for example aiming at promoting female participation in STEM – Science, Technology, Engineering and Mathematics – subjects (see e.g. OECD, 2014).

3. The German high school reform

3.1. Institutional background and change

In Germany, educational policy is the responsibility of the federal states. In all cases, however, children enter elementary school at the age of six and continue on to secondary education usually after four years. Secondary education in Germany is provided at three different levels, listed in ascending order by their level of education provided: *Hauptschule* (basic track), *Realschule* (intermediate track) and *Gymnasium* (upper track). Of these three, only successful completion of *Gymnasium* (henceforth referred to as academic-track or simply high school) leads to the *Abitur*, the university entrance qualification.[5] With a share of 34.4% of all German secondary students in the 2012/13 academic year attending *Gymnasium*, it is the most attended type of secondary school (Malecki et al., 2014).

Typically, high school lasted nine years, implying a total of thirteen years of schooling. Starting in 2001, several German states reduced this time at high school by one year, enabling graduation after completing only twelve years of schooling which was desirable for three reasons.[6] First, the earlier graduation aligned with school duration in other countries, therefore making German graduates more competitive on the international labor market. Second, the reform was to increase efficiency in the German education system and hence reduce the costs per student. Third, it allowed earlier labor market entry expanding the labor force by one birth cohort. As Western labor markets face major challenges caused by demographic changes with an increasing disparity between the group of active workers and the rising share of an older population, this relieves pressure on the pension schemes. In contrast, opponents of the reform feared that the reduced time in high school would hurt the quality of education. However, the overall curriculum remained unchanged: from grade five through receiving the *Abitur*, 265 year-week hours must be completed (KMK, 2013).[7] As a result, weekly class hours significantly increased and school days prolonged. The increase of, on average, 3.7 class hours per week constitutes an increase of 12.5% of overall week hours. The allocation of this increase in workload to different grades is determined on a state and school level, but grades seven to nine are usually most affected. Although the reform was implemented across almost the entire country, the timing of the introduction differs by state. An overview of the implementation of the reform by federal state is given in Table 1.

3.2. Anticipated effects of the reform on cognitive skills

The reform may affect students' cognitive skills through several channels. I aim to disentangle two important mechanisms: the effect of an increase in instructional time (keeping age constant) and the role of earlier instruction (keeping reached educational level constant).[8] According to the Cunha and Heckman (2007) skill formation model, both cases should be assumed to lead to higher cognitive skills. Still, it is an empirical question whether and to what extent these mechanisms can be verified to lead to higher cognition, especially during adolescence. Even if they hold, further aspects may hinder or offset their positive effects. Therefore, I conduct two separate analyses; the results

[3] Existing literature related to any of these particular mechanisms will be elaborated on in Section 3.2 when discussing anticipated effects of the reform.

[4] The two are not exclusive as, naturally, there may be a variety of further potential channels in the relationship between education and cognitive skills. However, these are not considered here.

[5] In some states, comprehensive schools that provide all three tracks at one school exist in addition.

[6] A similar educational policy change took place in Ontario, Canada, in 1999: Krashinsky (2014) finds that students with one year less of high school perform significantly worse at university than their counterparts in terms of grades. Unlike the German high school reform, however, this change effectively reduced the curriculum taught as the number of years was reduced along with the number of courses available to students. The German setting is therefore unique in the sense that school duration was altered but the overall curriculum was not.

[7] Year-week hours are the number of weekly class hours in each year that are summed up over all years. This requirement was kept even while reducing high school duration from nine to eight years. In Berlin, Brandenburg, and Mecklenburg-West Pomerania the assignment to different secondary school tracks takes place only at grade seven. Here, the reform reduced time at high school from seven to six years but the year-weak hours requirement holds equally, counting class hours from grade five onwards.

[8] An increase in instructional time when filled with additional content but keeping age constant, naturally leads to the introduction of certain parts of the curriculum at younger ages. The first mechanism is therefore not perfectly to disentangle from earlier timing. However, the results reveal that the earlier timing hardly yields effects. Further, I consider a substantial increase in instructional time by more than 800 class hours, which can therefore be expected to clearly dominate in the analysis on this mechanism.

Table 1
Introduction of the reform by state.

State	Implementation of the reform	Graduation of first cohort affected
Saxony[a]	–	–
Thuringia[a]	–	–
Saarland	2001	2009
Hamburg	2002	2010
Saxony-Anhalt	2003	2007
Mecklenburg-West Pomerania	2004	2008
Bavaria	2004	2011
Lower Saxony	2004	2011
Baden-Wuerttemberg	2004	2012
Bremen	2004	2012
Hesse[b]	2004	2012-2014
North Rhine-Westphalia	2005	2013
Berlin	2006	2012
Brandenburg	2006	2012
Schleswig-Holstein	2007	2016
Rhineland-Palatinate[c]	2007	–

Source: Autorengruppe Bildungsberichterstattung (2010).
[a] Saxony and Thuringia kept the 12-year school system after reunification.
[b] Gradual introduction: school year 2004/05 (10% of all schools); 2005/06 (60%); 2006/07 (30%).
[c] In Rhineland-Palatinate, the reform has only been introduced in selected schools so far.

of which shed light on the mechanisms behind the relation between schooling and cognitive abilities.

Instructional time. The first analysis compares same-aged, i.e. seventeen-year-old students, where the students affected by the reform have accumulated significantly more class hours, which were filled with corresponding additional curriculum. This increase in instructional quantity should especially raise crystallized measures of intelligence, while fluid intelligence is generally assumed unaffected.[9] Few studies similarly investigate the impact of class hours, but on more curriculum-oriented achievement tests: Andrietti (2015) and Huebener et al. (2017) estimate the same reform's effects on PISA test scores of ninth-graders. Therefore, these students have been affected by the reform for less time compared to my sample of investigation of seventeen year-olds. Still, Andrietti (2015) finds positive effects in reading, mathematical, and science literacy skills, with the first being driven by female students, while Huebener et al. (2017) find that the reform widens the gap in student performance as the high-performing students benefit the most. Using PISA data as well, two other studies find that differences in instructional time explain variation in student performance across countries (Lavy, 2015), and within Switzerland (Cattaneo et al., 2016), respectively. Machin and McNally (2008) employ difference-in-differences estimation to evaluate the introduction of a literacy hour in English elementary schools. They find that devoting one hour per day on English literacy along with changing the structure and content of teaching increases students' rank in reading and English skills by 2 to 3 percentage points. Taylor (2014) uses a fuzzy regression discontinuity design to investigate the effect of increasing the share of class hours spent in math classes in sixth grade at Miami-Dade County Public Schools. He finds that math achievement rises by 0.16 to 0.18 standard deviations, but that effects fade with time passed since the remediation course. Contrary to the German high school reform, these latter policy changes do not constitute an increase in *overall* instructional time. Keeping the length of school days constant, increases in instructional time in one subject may therefore come at the cost of other subjects. Cortes et al. (2015) analyze a policy change in Chicago Public Schools that doubled the amount of time devoted to algebra for low skilled ninth-graders. Using a regression discontinuity design, they find positive effects on achievement test scores and further outcomes. Using heavy snowfall as an exogenous variation in the number of school days that students in Maryland could attend, Marcotte (2007) finds that students with less instructional time performed significantly worse on the Maryland School Performance Assessment Program exams. Different to the German high school reform, these two studies investigate the effect of an increase in instructional time keeping the curriculum constant. The increased (or decreased) time therefore serves for more (or less) repetition and practice of the same content, i.e. decelerates (or accelerates) the speed of learning during each class hour. In contrast, the reform analyzed in this study provides a unique setting in which an increase in instructional time implies both an increase in class hours along with the corresponding increase in the curriculum taught. In this case, it may be that either the additional knowledge taught cannot be absorbed by the students[10] or simply that cognitive skills are no longer malleable at this age in adolescence, bringing no particular change in cognition at all. Lastly, the increase in formal instructional time may substitute informal cognitively stimulating activities or come at the cost of further, e.g. non-cognitive, skills or extracurricular activities important for skill development, offsetting the positive effects on cognition or even negatively impacting them.[11]

Timing of instruction. The second analysis compares students in their final year of high school, although at different biological ages. Students affected by the reform are therefore in grade twelve, while students not affected by the reform are attending grade thirteen. At this point in time both students affected by the reform and students not affected have reached the same educational level, accumulating the same number of class hours. However, students affected by the reform have received this instruction at a relatively younger age. According to Cunha and Heckman (2007) this earlier investment – presumably leading to higher cognitive skills at an earlier stage in life, which is tested in the first analysis – increases a person's stock of skills at an earlier stage making any investment thereafter even more productive.[12] As a result, students affected by the reform may have acquired higher cognitive skills through this multiplier effect of early investment, at least in crystallized dimensions of cognition. Again, if the instruction only substitutes already present learning, thereby not altering the timing of learning, there should be no effect. However, while attending their final year of high school, these students are one year younger than those students not affected by the reform, which may have negative consequences for cognition, including both crystallized and fluid dimensions.[13] Furthermore, students may lack the maturity to digest particular subjects at a younger age making instruction less productive, as found by Clotfelter et al. (2015). Exploiting a policy shift in some school districts of North Carolina, they find that accelerating the

[9] It is not possible to completely separate these two dimensions of intelligence in a test environment, where fluid and crystallized skills are often required simultaneously, whenever, for example, speed is introduced to give specific knowledge. Furthermore, (Baltes, 1993, p. 581) notes that in practice, crystallized and fluid skills interact and that, in addition "the pragmatics [crystallized intelligence] always build on the mechanics [fluid intelligence]". Further, Cunha and Heckman (2007) assume that skills are cross-fertilizing, i.e. that changes in one domain of skills foster changes in another domain.

[10] Whether this is the case may especially differ between distinct types of students, as e.g. students with lower initial skills may have more difficulties with keeping up at the new pace.

[11] Dahmann and Anger (2014) show that the reform indeed had an effect on some personality traits. The participation in extracurricular activities seems, however, not to be affected (see Table 3).

[12] As the overall curriculum was not altered, students are expected to have acquired the same knowledge, not necessarily the same level of underlying of skills. Any positive reform effects are therefore consistent with the Cunha and Heckman (2007) skill formation model, even if total investment has not changed, as long as the productivity of investment decreases with age.

[13] See, for example, Baltes (1987), who illustrates the life-span development of cognitive abilities: Both crystallized and fluid intelligence peak close to the age of 25; however, crystallized ability remains relatively stable thereafter, whereas fluid ability decreases with age. Importantly, up to the early 20-years, both domains of intelligence increase with age, mostly irrespective of the environment.

introduction of algebra coursework into eighth grade has significant negative impacts on students' performances in algebra and the follow-up geometry course. Furthermore, they find that low performing students are harmed the most, further increasing inequality. Unlike this policy change, the German high school reform is not bound to any particular subject, rather applying to the complete high school curriculum.

4. Data

To investigate both potential mechanisms, I conduct two analyses which consider different samples. By nature, these pose different requirements to the underlying dataset to enable identification and are, therefore, based on two sources.

4.1. The German Socio-Economic Panel (SOEP)

The first analysis is based on a sample of same-aged students taken from the German Socio-Economic Panel (SOEP) study, which is a representative household panel surveyed annually (Wagner et al., 2007) with information on around 30,000 individuals in almost 15,000 households in the 2013 wave.[14] In addition to various individual and household characteristics, including family background and childhood environment, the SOEP includes cognitive potential measures for different subsamples since 2006. The cognitive abilities of adolescents, who respond to the SOEP youth questionnaire in the year they turn seventeen, are assessed in every wave starting in 2006. Thus, I use the 2006 through 2013 waves, including all adolescent respondents aged seventeen[15] who attend *Gymnasium* in my sample. To identify whether a student is affected by the reform, I use the information on the federal state of residence and the year of school entry. In case information on the latter is not provided, the year of school entry is imputed from the date of birth. As Saxony and Thuringia established a twelve-year-school system before Germany's reunification, I consider all students in these two states as affected.[16] I exclude students from Rhineland-Palatinate where the reform has not been implemented state-wide, and students from Hesse who entered high school in 2004 or 2005 when schools operated under both schemes. To avoid adding noise to the amount and level of education received by the subjects, I exclude all students who have repeated any grade.[17] Lastly, I restrict the sample to those who successfully completed the cognitive assessment test[18] and have valid information on their background and family characteristics. The final sample consists of 723 students, of whom 288 are affected by the reform.

Cognitive skill measures. In the SOEP adolescent questionnaire, cognitive skills are measured through a short form of the I-S-T 2000 R (see Amthauer et al., 2001) that takes 30 minutes. This test consists of three parts, each with 20 questions (for details see Schupp and Herrmann, 2009; Richter et al., 2013). The first part consists of word analogies and measures verbal skills: participants are asked to find a matching word according to a specific rule. In the second part numerical skills are measured, where the respondent has to fill in the correct arithmetic operators in incomplete equations. Together, these two (verbal and numerical) tasks record crystallized intelligence as they reflect an individual's explicitly learned competences. In contrast, the third task serves to measure fluid intelligence: here, three abstract figures are displayed according to a specific rule with participants asked to pick a fourth figure from five proposed figures. On each of these three test components adolescents answer as many questions as possible, in the given amount of time. The scores then measure the number of correct answers (out of 20 possible questions). To facilitate the interpretation of results, I standardize all scores to mean zero and variance one. Summary statistics are given in Tables B.1 and B.2 in the appendix and the development of cognitive skills over time is graphically illustrated in Fig. C.1 in the appendix.

Other variables. To account for individual characteristics that may also influence cognitive abilities, I control for several pre-reform characteristics in my preferred specification. These include socio-economic and demographic variables like gender, migration background, and when they were born[19]. Furthermore I capture a student's previous performance by the teacher's recommendation after elementary school.[20] Family variables include parental characteristics based on education, work status and occupational status, and also capture whether a student grew up with only one parent.[21] An overview of all variables is given in Table A.1 in the appendix.

4.2. The German National Educational Panel Study (NEPS)

The second analysis is based on data from the German National Educational Panel Study (NEPS), which is a longitudinal dataset aimed at mapping competence development and learning environment over the life cycle. It follows a multicohort sequence design starting with more than 60,000 target persons from six cohorts (Blossfeld et al., 2011). In addition to these six original cohorts, it includes a cross-sectional additional study in the German federal state of Baden-Wuerttemberg, which targeted students at academic-track high school in their final year.[22] Baden-Wuerttemberg is the third largest federal state in Germany, both in terms of area and population, with a share of 34% of students at *Gymnasium* in 2012/2013 (Malecki et al., 2014) which corresponds almost perfectly to the German average of 34.4%. Per student expenditures at general schools in Baden-Wuerttemberg (6,900 EUR) were very close to the German average (7,100 EUR) in 2013, and even perfectly coincided with the national average of 7,500 EUR when considering high schools only (Schmidt et al., 2016). In

[14] This paper uses data from the Socio-Economic Panel (SOEP), data for years 1984-2013, version 30, SOEP, 2014, doi:10.5684/soep.v30.

[15] In 2006, when the test of cognitive abilities was conducted for the first time, adolescent respondents from the 2004 and 2005 waves were also tested. I include these individuals (aged eighteen and nineteen) in my preferred specification. Birth year dummies control for potential age effects. Still, a robustness check including only seventeen year-olds is conducted to confirm the results.

[16] In a robustness check, students from these two states are excluded.

[17] Huebener and Marcus (2017) find that repetition rates up to grade nine remained unchanged by this reform. Indeed, only 55 students drop from my sample due to grade repetition; of which 29 are affected by the reform and 26 are not. In a robustness check I include these grade repeaters to confirm the results.

[18] The share of students who specifically refused to take the cognitive test does not significantly differ between the treatment group (3.6%) and the control group (5.8%).

[19] Specifically, this measures whether students were born in the first or in the second half of the year and, thereby, controls for the grade in school they attend at the date of the interview: As the vast majority of the interviews (77% of my sample) are administered during the first quarter of the year in which adolescents turn seventeen, students born between July and December usually attend grade ten. In contrast, students born between January and June usually enter school comparatively young and are, therefore, on average, one grade more advanced at the time of the interview.

[20] At the end of elementary school, teachers recommend one of the different secondary school tracks to the student's parents based on their perception of the student's performance and potential. I classify students as low-performing who were recommended to attend either *Realschule* or *Hauptschule*, i.e. the intermediate and lower secondary tracks. Even though this recommendation is not equally binding across all federal states, the number of these students attending *Gymnasium* nevertheless, is naturally very low.

[21] Additional estimations control for being the oldest child and being an only child, which both may be relevant for cognitive skills. These are not reported, as the measures are available for a subsample only, but confirm the findings. The same holds true for household size, which is excluded from the main specification as it is not measured prior to students being affected by the reform.

[22] This paper uses data from the National Educational Panel Study (NEPS): Additional Study Baden-Wuerttemberg, doi:10.5157/NEPS:BW:3.0.0. From 2008 to 2013, NEPS data were collected as part of the Framework Programme for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, the NEPS survey is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

Baden-Wuerttemberg, the last cohort not affected by the high school reform and the first affected cohort both graduated in 2012. Therefore, in 2012, the NEPS target population consisted of this double sized graduation class. Hence, I use this wave, including all respondents who attend the final grade of *Gymnasium* in my sample. Again, I exclude all students who repeated a grade. The final sample consists of 2,128 students, of whom 1,113 are affected by the reform.

Cognitive skill measures. Cognitive abilities are measured through an extensive 2 h 40 min test covering different educational dimensions. Reflecting explicitly learned knowledge, a 30 min achievement test in mathematics constitutes a measure of crystallized intelligence. On this test, students are given a set of 21 questions. Most questions are multiple choice, with others partly answered in an open format. Therefore, a weighted maximum likelihood estimate (WLE; Warm, 1989) based on the test items constitutes an individual's measure of mathematical ability. Fluid intelligence is covered by measures of general cognitive abilities, i.e. perceptual speed and reasoning. Perceptual speed is assessed by a picture symbol test where respondents are required to enter correct figures for the preset symbols according to an answer key (see Lang et al., 2007), with a total of 3×31 items to be solved in 3×30 s. Reasoning is measured in the same way as figural skills in the SOEP adolescent questionnaire: based on Raven's matrices, students fill in a missing geometrical element that fits the other elements of the matrix, in a total of 3×4 cases with 3×3 min time. For both of these fluid cognitive skill measures, the total score is calculated as the sum of correctly solved items. Again, for both crystallized and fluid measures of cognitive ability, I standardize all scores to mean zero and variance one. Summary statistics are given in Tables B.4 and B.5 in the appendix.

Other variables. In addition to achievement tests, the survey also includes further individual and school characteristics. Whether or not a student is affected by the reform is given in the survey. Individual characteristics include demographics as gender and migration background. Furthermore, the number of books at home, parental education, the father's work classification and the mother's occupational status characterize a student's socio-economic background and home environment. In addition, a survey conducted at the school level allows me to further control for school characteristics in some specifications, including school size, the share of students and teachers with a migration background, as well as stress factors caused by the implementation of the reform. Stress factors are areas such as resources and organization where the headmaster reported to have had particular difficulty when implementing the reform. An overview of all variables is given in Table A.2.

5. Empirical strategy

I exploit the German high school reform introduced in almost all federal states between 2001 and 2007 as a quasi-natural experiment to identify a causal effect of education on cognition. The control group consists of students who entered high school before the reform was introduced and, therefore, graduate after nine years of high school. In contrast, the treatment group consists of students entering high school after the implementation of the reform and thus graduating after only eight years.

5.1. Estimation using SOEP

For the first analysis, all students in the selected SOEP sample usually attend either grade ten or grade eleven at the time of the interview, thus having spent about either 5.5 or 6.5 years in high school at the time of the interview. However, the amount of education received during this time differs between the control group and the treatment group, as the reform provides an exogenous variation in the number of class hours attended at the time of the interview. Students affected by the reform should have accumulated at least between 800 and 945 class hours of education more, on average, than their non-affected counterparts at the same age.[23]

In a difference-in-differences framework, I assess the impact of this intensified curriculum on cognitive skills. I exploit the variation over time and region to identify causal effects estimating the following equation:

$$y_{ist,17} = \alpha \mathrm{REFORM}_{st} + X_i\beta + \sum_s \gamma_s \mathrm{STATE}_s + \sum_t \delta_t \mathrm{YEAR}_t + \varepsilon_{ist}, \quad (2)$$

where $y_{ist,17}$ is a measure of cognitive ability at age 17 of student i living in state s and born in year t. The variable of interest, $REFORM_{st}$ indicates whether students belong to the treatment or the control group. It equals 1 if students entering school in state s in year $t+6$ (or $t+7$ respectively, depending on their month of birth)[24] are affected by the reform when entering high school (i.e. belong to the treatment group) and 0 otherwise (i.e. belong to the control group). $STATE_s$ is a set of state dummies and $YEAR_t$ dummies indicate the year of birth. X_i is a vector of pre-reform individual characteristics, including the student's own demographic characteristics as well as childhood and family variables. The error terms, ε_{ist}, are clustered at the state level with 15 different states.[25]

Crucial to the identification of the prime parameter of interest, α, as a causal impact of education on cognitive skills, is the assumption that in absence of the reform, cognitive skills of students from the treatment group and of students from the control group do not differ significantly, i.e. $\alpha = 0$. This implies that cognitive skills develop similarly among students across states. While this so-called common trend assumption is not testable, it should be reasonable and not too restrictive in this case: here, students of the same school type are compared across different states. Since students likely select into different school types based on initial abilities and socio-economic background, I assume that students differ severely across different types of secondary schools and, therefore, also in their development of competences. In contrast, students *at high school*, but living in different states, can be expected to possess similar initial characteristics. Still, I allow for state-specific linear time trends in a robustness check. Furthermore, self-selection should not be possible, thus enabling a causal interpretation of the results. This is discussed extensively in Section 7 and addressed in several sensitivity analyses which clearly support the plausibility of this assumption. Lastly, the timing of the implementation of the reform may be related to certain state specific characteristics. According to Black et al. (2005) it is not crucial for my identification for the reform to be unrelated to these as I control for state fixed effects in the analysis. Nevertheless, see Dahmann and Anger (2014) for an investigation into the reform's implementation. They find that it is unrelated to the percentage of high school students in a state's population, to whether the government is conservative, to whether the next state elections were scheduled for 2001/2002, or to the state's GDP per capita. There is suggestive evidence that states with a higher median age of residents implemented the reform slightly earlier; an artifact related to the older population in East German states.

[23] The numbers are calculated as follows: (265/8-265/9)[average weekly increase in class hours due to the reform]*39.5[weeks of school per year]*5.5[years in high school so far] (or *6.5 years respectively).

[24] For the sample under consideration, the cutoff date is equal among all federal states: June 30. Hence, students born between January and June entered first grade six years after their year of birth, and students born between July and December entered first grade seven years after their year of birth.

[25] To account for the small number of clusters, it may be necessary to use wild cluster bootstrapped standard errors (see Cameron et al., 2008). The estimations, however, show that the wild cluster bootstrap leads to even slightly lower standard errors if different at all. Therefore, I report the usual standard errors without bootstrapping, as it is the more conservative estimation method in this case. Bootstrapped estimation results can be found in Table D.1 in the appendix.

5.2. Estimation using NEPS

For the second analysis, all students in the selected NEPS sample are in their final grade of high school and either in the first cohort affected by the reform or the last cohort not affected. At each school these groups attend the classes together during their final year.[26] It is an empirical question whether and to what extent the multiplier effect suggested by Cunha and Heckman (2007) can be found or is offset by potential biological age effects. To estimate this relationship between the timing of education and students' cognitive skills, I estimate a reduced version of Eq. (2) as there is no variation over time and across states in this sample:

$$y_{ij} = \alpha \text{REFORM}_i + X_i\beta + \varepsilon_{ij}, \tag{3}$$

where y_{ij} is a measure of cognitive ability of person i at school j, $REFORM_i$ is a dummy indicating whether person i is affected by the reform ($REFORM_i$=1) or not ($REFORM_i$=0), and X_i is a vector of individual characteristics. The error terms, ε_{ij}, are clustered at the school level.[27] The prime parameter of interest, α, indicates the role of the timing of the instruction received in students' cognitive skill development: a positive α could prove the existence of the multiplier effect, proposed by Cunha and Heckman (2007) for early life interventions, even in adolescence. A negative α in turn, could stem from potential age effects. Of course both effects may not be present or may offset each other, thus yielding inconclusive results.

To interpret this relationship as causal, it is crucial that the reform indeed constitutes a quasi-natural experiment. For this to hold, no selection should be possible while the treatment and control groups should be comparable in terms of both observable and unobservable characteristics. As the reform was introduced state-wide at the same time, students did not have a choice on whether to be affected by the reform or not; hence selection to treatment or control group *within* this sample can be ruled out.[28] However, selection *out of* the sample may have been possible, but is assumed to be unlikely.[29] Furthermore, Table B.6 in the appendix shows that treatment and control group are comparable with respect to the selected observable characteristics. While the comparability of unobserved characteristics in turn cannot be tested formally, the well-balanced observed characteristics together with the assumption of no self-selection point to the validity of this assumption.

6. Results

6.1. The impact of instructional time on cognitive skills

Estimation results of Eq. (2) are presented in Table 2.[30] Female students score significantly lower on both verbal and numerical skills, but show no differences to male students in terms of figural skills. The few other salient effects of the individual control variables that reveal statistical significance[31] are in line with expectations: in particular students who received the recommendation *not* to follow onto high school after elementary school, possess significantly less skills throughout all domains. Furthermore, students with a migration background show comparably less fluid skills and students with a working-class family background lack behind in both verbal and figural skills.[32]

Table 2
Average effects of the reform.

	Outcome Variables: Cognitive Skills		
	Crystallized		Fluid
	Verbal	Numerical	Figural
Reform	0.068	0.138	0.089
	(0.094)	(0.131)	(0.095)
Female	-0.255***	-0.412***	0.052
	(0.055)	(0.077)	(0.053)
Migration background	-0.173	-0.196	-0.298***
	(0.170)	(0.154)	(0.061)
Born January–June	0.069	0.008	0.087
	(0.060)	(0.087)	(0.098)
Low-performing student	-0.412***	-0.380***	-0.259**
	(0.079)	(0.092)	(0.096)
Rural area	-0.003	0.013	0.042
	(0.097)	(0.062)	(0.074)
High parental education	-0.008	-0.099	-0.019
	(0.097)	(0.100)	(0.077)
Working-class father	-0.275***	0.076	-0.131*
	(0.082)	(0.095)	(0.074)
Working mother	0.038	0.075	0.017
	(0.070)	(0.093)	(0.080)
Single parent	0.017	0.055	0.106
	(0.105)	(0.070)	(0.123)
R^2	0.114	0.121	0.101
Observations	723	723	723

Notes: SOEPv30 waves 2006 to 2013. OLS regressions. Further, a maximum set of state dummies, year of birth dummies, dummies for the different SOEP subsamples, and a constant are included. Standard errors, reported in parentheses, are clustered at the state level. *p < 0.1, **p < 0.05, ***p < 0.01.

When looking at the impact of the reform, it can be seen that, on average, the coefficients exhibit positive signs across all dimensions of cognitive skills, which is in line with theory. Yet, none of them reveals to be statistically significant, indicating that the reform did not significantly improve students' cognitive abilities, on average. Also, their magnitude between 0.07 and 0.14 standard deviations is smaller than the effect sizes of one additional year of schooling between 0.2 and 0.3 standard deviations found in other studies for this age group (c.f. Brinch and Galloway, 2012; Carlsson et al., 2015; Falch and Massih, 2011; Cascio and Lewis, 2006). To further investigate potential driving factors of why students' cognition did not improve as much as elsewhere, I estimate the reform's impact on further outcomes to uncover how students reacted to the reform and whether other stimulating activities were crowded out (see Table 3).

[26] Along with the reform of shortening high school, Baden-Wuerttemberg revised the curriculum to move from an input-oriented teaching to an output-oriented teaching focusing on achieving educational standards. In the two final years, however, both affected and non-affected students attend the same classes and are, therefore, subject to the exact same curriculum and type of teaching.

[27] The number of different schools is 48 which suffices for standard inference (Cameron et al., 2008).

[28] Changing from the treatment to the control group (or vice versa) would only be possible by skipping (repeating) a class. However, this is ineffective as both groups end up in the same graduating classes.

[29] This could happen if students drop out of this double cohort either by repeating or skipping a class. Besides the rare cases, this would not have changed the fact of being affected by the reform; hence it would be selection unrelated to the implementation of the reform and, therefore, not pose a threat to the identification. An alternative is moving to a different state where the reform had not yet been introduced. Involving high moving costs for the entire family, this option seems highly unlikely. So does choosing an alternative secondary school track instead of high school. Nonetheless, see Section 7 for a more extensive discussion of potential selection based on the SOEP sample supporting the non-selectivity assumption.

[30] As the data on adolescents in the SOEP is a pooled cross-section over several waves, no appropriate weights exist. To account for possible over- or underrepresentation of certain demographic groups, I include dummies for each SOEP subsample in this, as well as all following, estimations instead. The different SOEP subsamples correspond to newly entering groups in the survey, partly with a demographic focus as target.

[31] The absence of statistical significance does not necessarily imply a zero effect. Given the relatively small sample size, a lack of statistical power can naturally be expected.

[32] The sample is special in the sense that it only comprises high school students, and this selection into the highest track may differ between socio-economic groups. This could explain the negative but insignificant coefficients of high parental education in case a much larger share of students in this group make it into high school while it is only the very best among students with low parental education. The effects of the individual characteristics on cognitive skills may, therefore, not be representative for the average

Table 3
Channels – effects of the reform on leisure-time activities and paid tutor lessons.

	Outcome Variables				
	Participation in Activity				Tutor
	Music	Sport	Reading	Tech. work	lessons
Reform	0.098	-0.043	0.017	-0.028	-0.034
	(0.081)	(0.052)	(0.063)	(0.084)	(0.073)
Observations	723	723	723	721	722

Notes: SOEPv30 waves 2006 to 2013. OLS regressions. Further, a maximum set of state dummies, year of birth dummies, dummies for the different SOEP subsamples, and a constant are included. Individual characteristics controlled for include female, migration background, born January–June, low-performing student, rural area, high parental education, working-class father, working mother, and single parent. Standard errors, reported in parentheses, are clustered at the state level. *p < 0.1, **p < 0.05, ***p < 0.01.

Table 4
Heterogeneous effects of the reform by gender.

	Outcome Variables: Cognitive Skills		
	Crystallized		Fluid
	Verbal	Numerical	Figural
Reform	0.094	0.289**	0.141
	(0.106)	(0.121)	(0.108)
Reform*Female	-0.052	-0.290***	-0.099
	(0.156)	(0.092)	(0.104)
R^2	0.114	0.125	0.102
Observations	723	723	723

Notes: SOEPv30 waves 2006 to 2013. OLS regressions. Further, a maximum set of state dummies, year of birth dummies, dummies for the different SOEP subsamples, and a constant are included. Individual characteristics controlled for include female, migration background, born January–June, low-performing student, rural area, high parental education, working-class father, working mother, and single parent. Standard errors, reported in parentheses, are clustered at the state level. *p < 0.1, **p < 0.05, ***p < 0.01.

First, I analyze changes in leisure-time activities, as in an experimental study, Schellenberg (2004) shows that e.g. music lessons increase IQ among children. I therefore analyze whether such activities which may be related to cognitive skill development, were crowded out by the reform. The estimates show that there were no such effects neither on music, sports, reading, or technical work.[33] Hence, although the increase in instructional time came with longer school days, relevant after-school activities do not seem to have been crowded out.

Second, I analyze whether outcomes related to additional investment in the students' performance at school changed following the reform. These may hint at whether students cope with the accelerated learning through the increased instruction by the same age or rely on additional resources. In particular I look at the use of paid tutor lessons, which seems not to be altered by the reform. In unreported regressions I find that also parental involvement in the student's educational outcomes remains unaltered, as measured by general interest of parents in their child's school performance, their help with homework, as well as problems arising between children and parents as a result of disagreements over studies.[34] Overall, students therefore seem to cope well with the increased learning intensity, which supports the potential for positive effects of the reform on cognitive skills despite the coefficients lacking statistical significance on average.

Heterogeneous effects. So far, the estimates show no average effects of the increased instructional time for the overall population of high school students, which is consistent with the zero findings on long-term effects in Germany by Kamhöfer and Schmitz (2016). However, effects may differ for distinct subgroups which could be masked in these estimations. In particular given that the studies investigating only male adolescents find positive short-term effects (cf. Brinch and Galloway, 2012; Carlsson et al., 2015; Falch and Massih, 2011) and that also long-term effects are partly found among males only (cf. Banks and Mazzonna, 2012) or to be larger among males (cf. Schneeweis et al., 2014), investigating potential effect heterogeneity by gender is of crucial interest. Indeed, there is suggestive evidence for such gender differences in the effects (see Table 4): While there is virtually no effect among female students, male students' numerical skills improve by more than a quarter of a standard deviation following the reform. Given that male students in this sample outperformed female students in numerical abilities even before the introduction of the reform (see Table B.2 in the appendix),[35] it is notable that here the increased instructional time could be an aggravating, rather than mitigating, factor for gender skill differences. In addition, the magnitude of this effect coincides with the estimates found for male adolescents in the literature. Furthermore, using estimates on wage returns to cognitive ability in Germany by Heineck and Anger (2010), the increase by 0.3 of a standard deviation found here could translate into a wage increase for males by up to 1 percent, thereby not only raising gender skill differences but also the gender wage gap as a consequence.

However, these results should be interpreted cautiously as the problem of multiple testing arises when testing for the presence of effects in multiple subgroups, because the probability of falsely detecting significant effects simply by chance rises with the number of subgroups tested.[36] In my case, the average results together with four distinctions of subgroups[37] result in a total of nine coefficients to be tested. Therefore, I additionally consider whether the gender specific effects still hold after a Bonferonni correction of p-values based on the nine different hypotheses. While the reform effect of males results in a statistically not significant p-value at conventional levels anymore, the interaction term still remains significant at the ten percent significance level with a Bonferroni corrected p-value of 6.3%. Hence, this conservative approach suggests that despite investigating multiple subgroups, there seem to be gender differences in effects. In addition, the sensitivity analyses presented in Section 7 consistently find them to be robust across different sample specifications and methodological alterations suggesting that the increase in instructional time widens the gender gap in numerical skills. The initial dominance of male students in this skill domain also provides two potential reasons for why instructional

(*footnote continued*)
adolescent in Germany but are included here merely for control purposes.

[33] The outcome variable for music and sports in each case refers to participating in this activity *at all*. Investigating the frequency of the activity (at least daily or at least once a week) instead, does not alter the results. The outcome variable for reading and technical work or programming in each case refers to participating in this activity *at least once a week*. Here as well, investigating the reform's effects on daily participation does not show a different pattern.

[34] Estimation results are available from the author upon request.

[35] A potential reason for this observed gender difference could be a greater variability among males compared to females, i.e. a larger share of male adolescents scoring particularly low and high on the skill assessment (see e.g. Hedges and Nowell, 1995). In this case, including only students at academic-track high school in the analysis, who presumably possess higher skills, mechanically raises skill averages among male compared to female students. In my case however, the inspection of students at *all* types of secondary school shows that the greater variability hypothesis is not supported by this data but rather that males outperform females in verbal and numerical skills across all types of secondary school.

[36] The problem of multiple testing also arises when investigating multiple outcomes. This is addressed in Section 7 where results prove to be robust to combining the three measures of cognitive ability to one general summary index.

[37] In addition to the average sample and the distinction by gender, distinctions by socio-economic status, migration background, and calendar half of birth are considered later in this paragraph.

time may be more beneficial for male students than for female students. On the one hand, the initially higher numerical skills can shape male preferences for choosing mathematically demanding subjects.[38] On the other hand, the high initial numerical skills especially among male students constitute a higher stock of already existing skills. According to the assumption of Cunha and Heckman (2007) that investment is more productive when existing skills are higher, the increase in instructional quantity especially benefits those with already higher skills, i.e. male students in the domain of numerical skills. Like this, education seems to improve especially domains of skills with comparative advantages among the respective group of students. This is also in line with the finding by Huebener et al. (2017) that the reform widened student achievement gaps as high-performing students benefit the most.[39] In addition, female students seem to cope less well with the increased learning intensity as investigating potential channels reveals that, although on average the utilization of paid tutor lessons did not change, it significantly increased for females compared to their male counterparts by 12.6 percentage points. Again, this should be regarded with caution due to the potential problem of multiple testing. Still, the magnitude of this effect is sizeable given a level of paid tutor lessons at 36.8 percent in the overall sample. Furthermore, it is well in line with Quis (2015) who consistently finds inferior outcomes for female students when investigating students coping with the reform in terms of stress and mental health.

Following similar lines of reasoning as for gender differences, further effect heterogeneities could enhance inequality. With a particular focus on disadvantaged students, defined by socio-economic or migration background, no such differences are found, however. Students born in the first half of the calender year are not affected differently than those born in the second half of the year either, even though they differ by construction in the grade in school they attend at the time of the interview.[40]

6.2. The impact of timing of instruction on cognitive skills

Estimation results of Eq. (3) are presented in Table 5.[41] Again significant gender differences can be observed in that male students outperform females in mathematics and reasoning ability, while females score higher on perceptual speed. The effects of the remaining individual control variables on cognitive abilities are in line with expectations: whereas children with a migration background score relatively lower on the competence measures, students from high socio-economic backgrounds, measured by parental education or the number of books at home, have acquired higher skills especially in the crystallized domain of competences. Further, students born in the first half of the year slightly outperform students born in the second half of the year in mathematics and perceptual speed. Entering elementary school at a relatively younger age, these students may have benefited from the earlier instruction from first grade onwards and a stimulating environment with more older classmates.

Table 5
Average effects of the reform.

	Outcome Variables: Cognitive Skills		
	Crystallized	Fluid	
	Mathematics	Speed	Reasoning
Reform	0.008	-0.049	-0.080*
	(0.043)	(0.066)	(0.046)
Female	-0.584***	0.268***	-0.231***
	(0.054)	(0.054)	(0.048)
Migration background	-0.169***	-0.105*	-0.172***
	(0.052)	(0.061)	(0.055)
Born January–June	0.061*	0.082*	-0.020
	(0.031)	(0.047)	(0.050)
High parental education	0.129**	-0.072	-0.003
	(0.049)	(0.046)	(0.056)
Working-class father	-0.041	0.054	0.056
	(0.061)	(0.077)	(0.069)
Working mother	-0.025	-0.035	0.053
	(0.059)	(0.062)	(0.055)
Books at home	0.187***	0.009	0.013
	(0.046)	(0.053)	(0.046)
R^2	0.108	0.023	0.021
Observations	2125	2128	2128

Notes: NEPS:BW:3.0.0 wave 2011/2012. OLS regressions. Further, a constant is included. Standard errors, reported in parentheses, are clustered at the school level. *p < 0.1, **p < 0.05, ***p < 0.01.

However, the reform exhibits no statistically significant effects on mathematical competence, as one dimension of crystallized intelligence (see Table 5).[42] This indicates that students affected by the reform have caught up with their non-affected counterparts in terms of mathematical competences. However, these estimates also reveal that the age-respective timing of instruction during adolescence does not influence skill formation in this crystallized domain; for two potential reasons: on the one hand, there may be neither a positive multiplier effect of earlier investment present nor a biological age effect benefiting older students' competence. On the other hand, both effects may be present, but offset each other. While it is not possible to disentangle these two mechanisms in this setting, the effects on fluid intelligence may give additional valuable insights.

Investigating the reform's effects on fluid measures of intelligence, the estimates show that there is no significant impact on processing speed; but reasoning ability, as measured by Raven's matrices test, is significantly lower for students affected by the reform. As fluid intelligence is assumed not to be directly affected by any type of investment, no positive multiplier effect of earlier instruction could be expected. Nonetheless, fluid intelligence *is* assumed and found to change with biological age where it is increasing during childhood and adolescence. The estimated decrease of eight percent of a standard deviation, therefore, most likely stems from the age difference of one year, on average, between affected and non-affected students, but should not be related to the curriculum covered at any particular age. These results on fluid intelligence can be taken into account when interpreting the zero effects on the crystallized dimension: given the students' performance in the tasks to assess reasoning ability, age effects in cognitive skill formation seem still to be present in late adolescence benefiting older students. If this was true for all dimensions of cognition, age effects can be expected to also influence crystallized dimensions. Hence, the zero effect of the reform on mathematical ability among students of the same educational level likely is the result of an interaction of this age effect offset by a positive

[38] The choice of the major fields of study is only possible in the last two years of high school. Although after the reform this choice takes place one year earlier than before, most students in my sample are not able to choose major fields yet (or did so only very recently), as they are interviewed the year they turn seventeen. Even if they were, note that as mathematics belongs to the core subjects it cannot be eliminated by any of the students. The same holds true for German literature.

[39] In my sample, unreported quantile regressions at the 0.25, the 0.5, and the 0.75 cutoffs largely exhibit coefficients increasing in magnitude along these cutoffs. Although imprecisely estimated given the sample size, these results are consistent with the theory of cognitive returns to the reform being increasing with the level of skills. Estimation results are available from the author upon request.

[40] Estimation results are available from the author upon request.

[41] Due to a resulting loss in the number of observations, I do not include school characteristics in this and the following estimations. Table E.1 in the appendix shows however that, when further controlling for school characteristics, results are not altered.

[42] Even though a lack of statistical significance could be a consequence of the sample size, the 95% confidence intervals range between -0.14 and +0.14 (Males) and -0.09 and +0.11 (Females). Hence, in any case coefficient sizes are smaller than for the increase in instructional time.

multiplier effect of earlier investment.

Analyzing whether the reform's effects differ by the characteristics of its implementation, reveals that the results are insensitive to the school's assessment of how smooth the implementation of the reform went.[43] In particular, whether certain aspects were regarded as strong stress factors in the transition or not, does not yield higher costs in terms of the school's students' cognitive abilities. In contrast, the perception of negative consequences of the reform did: in the schools where the reform was assessed to have a negative impact in general, students scored lower on the mathematics ability test. However, this may reflect the reverse pathway as low performing students may induce the headmaster to evaluate the reform negatively.

In general, I find that crystallized intelligence is not affected by the earlier timing of instruction. One reason could be that generally the earlier instruction could simply have replaced cognitively stimulating activities at home. Unfortunately the data do not allow for the investigation of responses in leisure-time allocation to activities that are especially mathematically stimulating. Computer usage refers in particular to playing computer games and chatting, instead of programming, and is, therefore, not as mathematically challenging. Still, there is no reform effect on computer usage (see Table 6). In contrast, there are responses to the reform in terms of leisure-time reading which decreased following the reform. Furthermore, I find an increase in the take-up of paid tutor lessons following the reform by more than six percentage points. This suggests that some students may have had difficulties coping with the earlier instruction and the induced increased learning intensity. These adverse effects could also explain the zero finding, on average, of why a presumably positive multiplier effect of earlier instruction did not lead to higher mathematical ability. The participation in sports during leisure-time seems to increase as well, at least on average, but the change amounts to only less than three percentage points.

Heterogeneous effects. Although, on average, I find no significant effects on crystallized intelligence, again the earlier timing of instruction may have impacted particular groups of students differently. If differential effects enlarge or reduce existing inequalities, they are of particular interest for policy makers seeking to decrease prevalent skill gaps. Contrary to the analysis on instructional time, however, the estimation results do not reveal any gender differences (see Table 7) indicating that the tentative male advantage in numerical skills at the age of seventeen did not translate into higher mathematical ability at the time of graduation.[44] When investigating potential channels, unreported estimation results reveal that it is especially male students who reduce reading in their leisure-time following the reform. This may explain why the advantage in crystallized intelligence males potentially obtained at the age of seventeen is offset until graduation as male students reduce their cognitively stimulating leisure-time activities compared to females. With respect to further individual characteristics, again hardly any such differences exist with respect to demographic and socio-economic variables, the only exception hereto being socio-economic status when defined by the number of books at home where disadvantaged students face an improvement in mathematical ability.[45]

7. Sensitivity analyses

To confirm the positive, yet statistically insignificant, coefficients of the increased instructional time, I conduct several robustness checks with the corresponding tables provided in the appendix. Table D.1 presents methodological alterations: wild cluster bootstrap to account for the small number of clusters yields even lower standard errors. Including a linear trend for each state to allow for state-specific developments over time, I also find that results are not altered. To estimate a more general effect of the reform, I construct a summary index of cognitive ability following Anderson (2008). This summary index has the further advantage of being potentially more powerful than considering each outcome individually. In addition, it accounts for multiple hypothesis testing[46], as tests are robust to overtesting since additional outcomes do not increase the probability of a false rejection. Indeed, the estimation reveals a positive effect of the reform on cognitive ability in general of 0.1 of a standard deviation which is estimated more precisely than each of the individual-level tests. However, it does not suffice for establishing statistical significance for the average sample. Yet, the estimation still reveals significant positive effects for male and zero effects for female students in this overall measure of cognitive ability.

To deal with the measurement of cognitive skills (see Table D.2), I first exclude all students aged eighteen or nineteen years from my sample to rule out potential age effects.[47] The average results are preserved. Second, I control for the month in which the assessment of cognitive skills took place as those interviewed later in the year have additional knowledge over students interviewed earlier, which has no

Table 6
Channels – effects of the reform on leisure-time activities and paid tutor lessons.

	Outcome Variables				
	Participation in Activity				Tutor
	Music	Sport	Reading	Computer	lessons
Reform	-0.010	0.029*	-0.034*	0.012	0.067**
	(0.021)	(0.015)	(0.017)	(0.010)	(0.026)
Observations	2114	2096	2094	2086	2122

Notes: NEPS:BW:3.0.0 wave 2011/2012. OLS regressions. Individual characteristics controlled for include female, migration background, born January–June, high parental education, working-class father, working mother, and books at home. Further, a constant is included. Standard errors, reported in parentheses, are clustered at the school level. *p < 0.1, **p < 0.05, ***p < 0.01.

Table 7
Heterogeneous effects of the reform by gender.

	Outcome Variables: Cognitive Skills		
	Crystallized	Fluid	
	Mathematics	Speed	Reasoning
Reform	0.003	-0.072	-0.090
	(0.070)	(0.094)	(0.061)
Reform*Female	0.009	0.040	0.017
	(0.082)	(0.097)	(0.088)
R^2	0.108	0.024	0.021
Observations	2125	2128	2128

Notes: NEPS:BW:3.0.0 wave 2011/2012. OLS regressions. Individual characteristics controlled for include female, migration background, born January–June, high parental education, working-class father, working mother, and books at home. Further, a constant is included. Standard errors, reported in parentheses, are clustered at the school level. *p < 0.1, **p < 0.05, ***p < 0.01.

[43] Estimation results are available from the author upon request.

[44] Numerical ability measured in SOEP is not directly comparable to mathematical ability measured in NEPS. Although both address the same or similar dimensions of crystallized intelligence, SOEP only tests basic numerical ability, independent of the educational curriculum covered. In contrast, the achievement test in NEPS explicitly asks for knowledge covered at this stage in high school including analysis, linear algebra, and statistics.

[45] Estimation results are available from the author upon request.

[46] The concern of multiple hypotheses testing arises when investigating several outcomes. Therefore, summary indexes weight each dimension to maximize the amount of information captured.

[47] In all other specifications, I include birth year dummies to account for this.

impact though. Lastly, I investigate the persistence of effects in this sample. For a small subsample, an additional measure of cognitive skills is available via the adult questionnaire in 2012. At this point, the individuals are of age 18 to 25, with most of them, thus, having completed high school. The very small number of observations of 80 or less, naturally yields imprecise estimates. Still, the magnitude of the coefficient in particular in the dimension of crystallized intelligence remains consistent.

Next, I consider the composition of students to validate that the reform can indeed by regarded as a quasi-natural experiment (see Table D.3). First, I omit all individual characteristics in the specification. As this does not alter the results qualitatively, I conclude that they are not biased by the omission of these observables. Second, I add grade repeaters to my analysis which is important to consider in case there are systematic differences in repetition rates following the reform e.g. by gender. This seems however not to be the case, as the results are preserved.

Crucial for the interpretation of a causal effect in specification (2) is that there is no selectivity (see Table D.4). First, I consider selection within the sample. As the reform was introduced state-wide at the same time,[48] students did not have a choice on whether to be affected by the reform or not.[49] However, students could attend high school in a different state that had not yet introduced the reform. This, however, imposes high moving costs for the entire family and, therefore, seems unlikely. Still, I consider two samples in which this is not possible or did not occur: (i) federal states that adopted the reform relatively late, i.e. with the first affected cohort graduating in 2012 or later, and neighbor states where the reform was already implemented and (ii) students who never moved from their place of childhood.[50] The estimations reveal largely even greater coefficients for crystallized intelligence. In particular with respect to verbal skills, affected students in late-adopter states seem to significantly score 0.4 of a standard deviation higher than non-affected students, on average. This implies that the reform has stronger effects when being less avoidable. At the same time this illustrates that selectivity, if at all a problem, leads to a downward bias in estimates suggesting that the original estimates are conservative. Second, I consider selectivity out of the sample by students attending a different type of secondary school instead of high school. The graduation from the lower and intermediate secondary school does not lead to the *Abitur*. Given the ever growing importance of educational certificates on the labor market, this is a far-reaching decision and can, therefore, be assumed to be relatively rarely a direct implication of the newly introduced reform. Indeed, Huebener and Marcus (2017) find that high school entry and graduation rates are not affected at all by the reform. In some states, however, the university entrance qualification can still be obtained after a total of 13 years of schooling at comprehensive schools, which combine all three – academic, intermediate and lower – tracks of secondary school.[51] As comprehensive schools are not equally common across all federal states, I include only students from states where no or only few comprehensive schools exist to further rule out selection. With the exception of the effect on verbal skills, which remains insignificant but turns negative, coefficients remain of similar size.

Institutional aspects are considered in Tables D.5 and D.6. First, even though there is a large discrepancy between former East and West Germany and their educational systems historically differed, the reform effects do not differ between these regions. Second, the exclusion of students from Saxony and Thuringia, where the reform was not introduced but an established twelve-year-schooling system continued, shows that the magnitude of the effect slightly decreases but that males' numerical skills still significantly improve and in particular benefit compared to females' skills. Third, I consider central exit examinations which existed since the 1990s or earlier in some states[52] and were introduced between 2005 and 2008 in most of the remaining states. Importantly, the introduction of these examinations did not coincide with the implementation of the high school reform, but took place earlier. Hence, the entire sample under analysis, both the treatment *and* control group, are subject to final exit examinations to obtain their *Abitur*. Moreover, central exit examinations are only relevant at the very end of high school and, hence, are unlikely to systematically affect seventeen year-olds at their current educational stage. Still, to rule out any interplay between the effects of both educational changes, I consider the subsample of states only where this tradition was long-standing existent. Coefficients are altered to some extent, possibly also due to a substantial reduction of the sample, but still keep indicating that numerical skills may be the most affected. Fourth, I exclude students from the double graduating cohort which might have been impacted differently, as the first cohort affected by the reform and the last one not affected were to merge into one class for the last two final years. Again, the coefficient on numerical skills remains largest, while all coefficients decrease but continue statistically insignificant. Fifth, I use dummies based on the academic years instead of calender years to account for the time dimension in the estimation of Eq. (2), as the academic year ranges from July to June, as defined by school-entry cut-off dates. The results are consistent with previous estimations. Sixth, I consider a potential heterogeneity that can arise due to institutional differences. The allocation of the increase in class hours to the different grades is at the discretion of the federal states and even schools. Yet, in the states of Berlin, Brandenburg, and Mecklenburg-West Pomerania there is less scope as high school begins only with grade seven. Excluding these states from the analysis, the coefficients become more pronounced, though still statistically insignificant. This could suggest that especially the earlier years in high school, i.e. grades five and six, may be critical, and that a larger scope to allocate increases in class hours may be more beneficial.

Lastly, to verify the empirical strategy, I use the sample of students who follow secondary school tracks *other* than high school in a Placebo estimation where the reform indeed does not show effects (see Table D.7).

Hence, overall, the sensitivity analyses are consistent with the average results in that they reveal mostly positive, but statistically not significant, coefficients when investigating the impact of instructional time on cognitive skills. However, at the same time, all sensitivity analyses prove the gender-specific results to be robust: estimations accounting for methodological alterations (see Table D.1), for the measurement of cognitive skills (see Table D.2), for the composition of students (see Table D.3), for selectivity (see Table D.4), and for institutional aspects (see Tables D.5 and D.6) all consistently reveal that males benefit from the reform in terms of their numerical skills, in particular compared to

[48] Exceptions to this are Rhineland Palatinate and Hesse, from where students are excluded accordingly.

[49] As the reform has only been announced and implemented *after* these students had entered elementary school already, students could hardly change their grade in response to the introduction. Only in the first cohort affected and the last cohort not affected, repeating and skipping classes could have impacted the treatment status. Besides this rarely happening, estimations without these two grades, which constitute the double cohort, are provided in Table D.5 and confirm the findings.

[50] These constitute 94.5% of my original sample, hence demonstrating very low regional mobility. Place of childhood is defined as the town or area where they lived the majority of their life until age 15.

[51] Students from comprehensive schools are excluded from my sample. Still, the option could have affected selection into high school.

[52] Specifically Baden-Wuerttemberg, Bavaria, Mecklenburg-West Pomerania, Saarland, Saxony, Saxony-Anhalt and Thuringia.

females. When considering only states with high school entry in grade five already, the positive effect on males' numerical skills is even more pronounced while the gender heterogeneity is slightly reduced (see Table D.6) suggesting not only that more flexibility in allocating the increases in class hours may be more beneficial for males but also less detrimental for equality. In contrast, no such gender effects can be found in the Placebo estimation (see Table D.7), which once more underlines their robustness.

To confirm the validity of the findings on the timing of instruction, which exhibit zero effects in mathematics and slight decreases in fluid intelligence, I also conduct several sensitivity analyses. As this analysis is bound to one federal state, Baden-Wuerttemberg, no regional variation can be exploited to account for state-specific factors questioning complete external validity. Still, to prove robustness of the effects within Baden-Wuerttemberg, I focus on the composition of students (see Tables E.2 and E.3): First, I omit all individual characteristics from the specification which indicates that results are not biased by the omission of these observables as effects are not altered. Second, a weighted regression to allow generalization of the results with respect to high school graduates in Baden-Wuerttemberg confirms the coefficient sizes, while statistical significance naturally drops (Chambers and Skinner, 2003). Third, I add school fixed effects to the estimation to eliminate student sorting across schools and simultaneously control for all school-specific factors. Again, this reduces statistical significance but the results remain qualitatively similar. Fourth, I add grade repeaters to the sample, as Huebener and Marcus (2017) find no effects of the reform on repetition rates up to grade nine, but do find that rates doubled in the final years of high school. Results are not altered, as scores on mathematics are still unaffected while reasoning ability does decrease following the reform. Lastly, I add further waves to the sample to disentangle the effect potentially specific to the double-graduating cohort. For this I include one wave with students prior to the reform (2010/2011) and one with students post reform (2012/2013). The results show that the zero effect on mathematics is not altered, while the negative impact on reasoning ability slightly drops. Hence, it is the students in the double-graduating cohort who specifically face a short-term decline in fluid scores. However, the results have to be taken cautiously, as the institutional-specific effect of being in the double-graduating cohort cannot be disentangled from time-specific effects[53].

8. Conclusion

As cognitive skills are important determinants of many economic and social outcomes, higher cognitive skills are often correlated with higher education. However, it is not only that individuals with higher cognitive abilities are likely to be better educated, but also that education improves cognitive skills. Most studies use changes in compulsory schooling laws as an exogenous variation to identify causal positive effects of an additional year of schooling. However, there is not much evidence on the underlying mechanisms in the economic literature.

This paper provides first evidence on disentangling two mechanisms through which education may improve cognitive skills in adolescence. I exploit a German high school reform carried out at the state-level between 2001 and 2007 as a quasi-natural experiment to estimate causal effects of this educational change on adolescents' cognitive abilities. Based on two separate analyses using SOEP and NEPS data, this study successfully disentangles the differential effects of instruction by focusing on quantity, on the one hand, and allocation with respect to age, on the other hand.

An increase in instructional time does, on average, not statistically significantly improve seventeen-year old adolescents' cognitive skills which is consistent with Kamhöfer and Schmitz (2016). Yet, tentative results suggest that male students may benefit in terms of their numerical skills. The estimated improvement by up to 0.3 standard deviations is comparable to the effect sizes of one additional year of schooling found for young men in Scandinavian countries (cf. Brinch and Galloway, 2012; Carlsson et al., 2015; Falch and Massih, 2011). To the best of my knowledge, this is the first study pointing at potential heterogeneous effects by gender as a source for different findings in the existing literature and indicating the necessity for further research to uncover whether educational quantity may indeed aggravate, instead of mitigate, gender skill differences.

In contrast, a positive multiplier effect that could result from this skill acquisition at younger ages, at least for males, does not seem to outweigh potential biological age effects until graduation. The differential age-respective timing of educational instruction during adolescence does therefore not significantly alter cognitive skill development when comparing crystallized measures of competences of students affected by the reform and students not affected by the reform at the end of high school. As fluid intelligence is generally not assumed to change over the life cycle in response to factors other than age, no positive multiplier effect can be expected for the reform to increase fluid components of intelligence. The age gap therefore yields even lower scores for students affected by the reform compared to their non-affected counterparts. Lastly, these results can be drawn onto for the evaluation of the reform: they may justify the maintenance of the curriculum while shortening high school duration as students seem to absorb the higher load of subject matters taught. Apart from lower reasoning scores, which may be attributable to the age difference, the results suggest that high school graduates are just as equally well off before and after the reform in terms of acquired competences.

I conclude from these analyses that in the positive impact education has on cognitive skills the relevant factor is not (only) school duration. The relationship may work through the amount of content taught; yet this requires further research to establish reliable effects. For decision makers this could open up new possibilities to target cognitive ability, other than simply changing overall school duration when designing educational policies. However, differential effects should be further investigated to avoid increasing inequality.

Appendix A-E. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.labeco.2017.04.008

[53] These could be e.g. transition rates into high school changing over time, inducing a different composition of students.

References

Amthauer, R., Brocke, B., Liepmann, D., Beauducel, A., 2001. Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R). Hogrefe-Verlag, Göttingen.

Anderson, M.L., 2008. Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. J. Am. Stat. Assoc. 103 (484), 1481–1495.

Andrietti, V., 2015. The causal effects of increased learning intensity on student achievement: evidence from a natural experiment. Universidad Carlos III de Madrid Working Paper Economic Series 15-06, Universidad Carlos III de Madrid.

Autorengruppe Bildungsberichterstattung, 2010. Bildung in Deutschland 2010. W. Bertelsmann Verlag GmbH & Co, Bielefeld.

Baltes, P.B., 1987. Theoretical propositions of life-span developmental psychology: on the dynamics between growth and decline. Dev. Psychol. 23 (5), 611–626.

Baltes, P.B., 1993. The aging mind: potential and limits. Gerontologist 33 (5), 580–594.

Banks, J., Mazzonna, F., 2012. The effect of education on old age cognitive abilities: evidence from a regression discontinuity design. Econ. J. 122 (560), 418–448.

Black, S.E., Devereux, P.J., Salvanes, K.G., 2005. Why the apple doesn't fall far: understanding intergenerational transmission of human capital. Am. Econ. Rev. 95 (1), 437–449.

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

Brinch, C.N., Galloway, T.A., 2012. Schooling in adolescence raises IQ scores. Proc. Natl. Acad. Sci. 109 (2), 425–430.

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2008. Bootstrap-based improvements for inference with clustered errors. Rev. Econ. Stat. 90 (3), 414–427.

Carlsson, M., Dahl, G.B., Öckert, B., Rooth, D.-O., 2015. The effect of schooling on cognitive skills. Rev. Econ. Stat. 97 (3), 533–547.

Cascio, E.U., Lewis, E.G., 2006. Schooling and the armed forces qualifying test: evidence from school-entry laws. J. Hum. Resour. 41 (2), 294–318.

Cattaneo, M.A., Oggenfuss, C., Wolter, S.C., 2016. The more, the better? The impact of instructional time on student performance. IZA Discussion Papers 9797, Institute for the Study of Labor (IZA).

Chambers, R., Skinner, C., 2003. Analysis of Survey Data. Wiley Series in Survey Methodology, Wiley.

Clotfelter, C.T., Ladd, H.F., Vigdor, J.L., 2015. The aftermath of accelerating algebra: evidence from district policy initiatives. J. Hum. Resour. 50 (1), 159–188.

Cortes, K.E., Goodman, J.S., Nomi, T., 2015. Intensive math instruction and educational attainment: long-run impacts of double-dose algebra. J. Hum. Resour. 50 (1), 108–158.

Cunha, F., Heckman, J.J., 2007. The technology of skill formation. Am. Econ. Rev. 97 (2), 31–47.

Cunha, F., Heckman, J.J., 2008. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. J. Hum. Resour. 43 (4), 738–782.

Dahmann, S., Anger, S., 2014. The impact of education on personality: evidence from a German high school reform. IZA Discussion Papers 8139, Institute for the Study of Labor (IZA).

Falch, T., Massih, S.S., 2011. The effect of education on cognitive ability. Econ. Inq. 49 (3), 838–856.

Hanushek, E.A., Wößmann, L., 2008. The role of cognitive skills in economic development. J. Econ. Lit. 46 (3), 607–668.

Heckman, J.J., Stixrud, J., Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. J. Labor Econ. 24 (3), 411–482.

Hedges, L.V., Nowell, A., 1995. Sex differences in mental test scores, variability, and numbers of high-scoring individuals. Science 269 (5220), 41–45.

Heineck, G., Anger, S., 2010. The returns to cognitive abilities and personality traits in Germany. Labour Econ. 17 (3), 535–546.

Huebener, M., Kuger, S., Marcus, J., 2017. Increased instruction hours and the widening gap in student performance. Labour Econ., forthcoming.

Huebener, M., Marcus, J., 2017. Compressing instruction time into fewer years of schooling and the impact on student performance. Econ. Educ. Rev. 58, 1–14.

Kamhöfer, D.A., Schmitz, H., 2016. Reanalyzing zero returns to education in Germany. J. Appl. Econm. 31 (5), 912–919.

KMK, 2013. Vereinbarung zur Gestaltung der gymnasialen Oberstufe in der Sekundarstufe II. Beschluss der Kultusministerkonferenz (Secretariat of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany). Original July 7, 1972, revised February 7, 2013.

Krashinsky, H., 2014. How would one extra year of high school affect academic performance in university? Evidence from an educational policy change. Can. J. Econ. 47 (1), 70–97.

Lang, F., Weiss, D., Stocker, A., von Rosenbladt, B., 2007. Assessing cognitive capacities in computer-assisted survey research: two ultra-short tests of intellectual ability in the German Socio-Economic Panel (SOEP). Schmollers Jahrbuch: J. Appl. Soc. Sci. Stud. / Zeitschrift für Wirtschafts- und Sozialwissenschaften 127 (1), 183–192.

Lavy, V., 2015. Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. Econ. J. 125 (588), F397–F424.

Machin, S., McNally, S., 2008. The literacy hour. J. Public Econ. 92 (5-6), 1441–1462.

Malecki, A., Schneider, C., Vogel, S., Wolters, M., 2014. Schulen auf einen Blick. Statistisches Bundesamt, Wiesbaden.

Marcotte, D.E., 2007. Schooling and test scores: a mother-natural experiment. Econ. Educ. Rev. 26 (5), 629–640.

OECD, 2014. Promoting Female Participation in STEM. What it means for higher education: quarterly briefs 2014. Available on ⟨http://www.oecd.org/edu/imhe/Promoting-female-participation-in-STEM.pdf⟩. Higher Education Programme, Organisation for Economic Co-operation and Development (OECD), Paris.

Quis, J.S., 2015. Does higher learning intensity affect student well-being? Evidence from the National Educational Panel Study. BERG Working Paper Series 94, Bamberg University, Bamberg Economic Research Group.

Richter, D., Metzing, M., Weinhardt, M., Schupp, J., 2013. SOEP scales manual. SOEP Survey Papers 138, DIW Berlin, German Institute for Economic Research.

Schellenberg, E.G., 2004. Music lessons enhance IQ. Psychol. Sci. 15 (8), 511–514.

Schmidt, A., Eichstädt, H., Fußmann, M., 2016. Bildungsausgaben - Ausgaben je Schülerin und Schüler 2013. Statistisches Bundesamt, Wiesbaden, Bielefeld.

Schneeweis, N., Skirbekk, V., Winter-Ebmer, R., 2014. Does education improve cognitive performance four decades after school completion? Demography 51 (2), 619–643.

Schupp, J., Herrmann, S., 2009. Kognitionspotenziale Jugendlicher: Ergänzung zum Jugendfragebogen der Längsschnittstudie Sozio-oekonomisches Panel (SOEP); Methodenbericht und Variablendokumentation. Data Documentation 43, DIW Berlin, German Institute for Economic Research.

Taylor, E., 2014. Spending more of the school day in math class: evidence from a regression discontinuity in middle school. J.Public Econ. 117(C), 162–181.

Wagner, G.G., Frick, J.R., Schupp, J., 2007. The German Socio-Economic Panel Study (SOEP) – scope, evolution and enhancements. Schmollers Jahrbuch: J. Appl. Soc. Sci. Stud./Zeitschrift für Wirtschafts- und Sozialwissenschaften 127 (1), 139–169.

Warm, T., 1989. Weighted likelihood estimation of ability in item response theory. Psychometrika 54 (3), 427–450.