Intensity and content of private tutoring lessons during German secondary schooling: effects on students' grades and test achievement

Karin Guill¹ · Melike Ömeroğulları^{1,2} · Olaf Köller¹

Received: 11 June 2021 / Revised: 25 August 2021 / Accepted: 7 September 2021 /

Published online: 12 October 2021

© The Author(s) 2021

Abstract

Private supplementary tutoring is a widespread phenomenon. However, evidence that private tutoring has positive effects on academic achievement or about the specific conditions of successful private tutoring is rare. Adapting Carroll's (1963) model for school learning to private tutoring, we expected to find positive effects of tutoring duration, tutoring intensity, and students' motivation to attend private tutoring. In a sample of eighth-grade students in German secondary schools ($N=8510$, 18.6% currently being tutored), we conducted regression analyses with multiple covariates and did not find a positive main effect of private tutoring attendance in any of the school subjects examined. Moreover, within the subsamples of tutored students, we were not able to identify positive effects of tutoring duration, tutoring intensity, tutoring content (such as a focus on homework completion, test preparation, or study behavior), or students' motivation to attend private tutoring. Given these disillusioning findings, we primarily derive suggestions for future research.

Keywords Shadow education · Private tutoring · Learning time · Lesson content · Secondary school

✉ Karin Guill
guill@leibniz-ipn.de
Melike Ömeroğulları
melike.oemerogullari@fau.de
Olaf Köller
koeller@leibniz-ipn.de

¹ IPN - Leibniz Institute for Science and Mathematics Education, Olshausenstraße 62, 24118 Kiel, Germany

² FAU – Friedrich-Alexander-Universität Erlangen-Nürnberg, Regensburgerstraße 160, 90478 Nürnberg, Germany

Introduction

In line with the rising demand for private tutoring lessons (Byun et al., 2018), research on the spread and effectiveness of private tutoring lessons has greatly increased in the past 20 years, resulting in many articles, several special issues (Guill & Spinath, 2014; Manzon & Areepattamannil, 2014; Zhang & Bray, 2019), and various books on the topic (Bray et al., 2015; Jokic, 2013). In Germany, about half of the student population receives private tutoring at some point in their school career (Hille et al., 2016). However, looking at the effectiveness of private tutoring, one can still agree with the conclusion of the review by Park et al. (2016): Research is inconclusive. After systematic differences between tutored and nontutored students (Byun et al., 2018) were controlled for, some studies identified positive effects of private tutoring, but others found none at all. Given the enormous amount of economic resources families invest in private tutoring classes (Germany: Klemm & Hollenbach-Biele, 2016; South Korea: Lee, 2005), it is important to gain further insights into the circumstances under which private tutoring attendance improves academic achievement.

Private tutoring refers to tutoring in academic subjects (e.g., languages, mathematics, or science) to improve students' academic achievement. It is provided by tutors for financial gain and does not include extracurricular activities, such as sports, or remedial lessons in school (Bray, 2014). Despite this quite clear definition, many empirical studies rely on an intuitive understanding of the term "private tutoring" in their questionnaires and fail, for example, to differentiate between paid and fee-free private tutoring (Bray & Kobakhidze, 2014). However, when looking at the effects of private tutoring, as in our study, tutoring costs are not the most important characteristic. More relevant characteristics are the amount of private tutoring that students receive and the content of the tutoring lessons. At an international level, tutored students in East and Southeast Asia tend to attend much more intensive forms of private tutoring than students in Western countries (Guill & Wendt, 2016; Entrich, 2014). However, in both regions, if studies attempted to obtain more details on private tutoring lessons, they considered either the duration of private tutoring attendance, that is, how long students had already been attending (Ömeroğulları et al., 2020; Heinrich & Nisar, 2013) or the number of hours per week spent attending private tutoring lessons (Choi & Park, 2016 in South Korea; Guill et al., 2020a, in Germany); they did not consider the combination of both factors.

There are not many quantitative studies that provide data on tutoring content. In contrast to this, we were able to rely on a large and rich longitudinal data set from a metropolitan area in Northern Germany. This data set had several strengths: (1) Duration and weekly hours of tutoring were measured separately for several tutoring subjects. (2) Teacher-assigned grades, as well as achievement test results, were available as outcomes of private tutoring attendance and as control variables before the start of tutoring lessons. (3) At least some students provided data on the content of their tutoring lessons, differentiating, for example, between homework completion, preparation for tests in schools, and the instruction of learning strategies. (4) The data made it possible to check interaction effects between these characteristics of tutoring lessons and the characteristics of the students, namely, their prior achievement and their motivation to engage in private tutoring lessons.

In the remainder of this paper, we first discuss the theoretical reasons for why tutoring duration and intensity and specific tutoring content can be expected to affect academic achievement, and we present empirical research on this topic. Afterward, we present our

data, methods, and empirical results and discuss the consequences of our results for parent counseling, education policy, and future research.

Effects of private tutoring

Private tutoring as additional learning time

Tutoring is generally believed to be effective because it gives students more time to learn (Kuan, 2011; Mischo & van Kessel, 2005). This argument is based on Carroll's (1963) model of school learning, which posits that learning is a function of the ratio of the time spent on learning to the time needed to learn. The time needed depends on cognitive factors and the quality of instruction, and the time spent depends on the amount of time that is available and the student's motivation to use that time.

Empirical studies have reached inconsistent conclusions regarding the amount of time available for learning during tutoring and the impact of that time on school achievement. There is some evidence of linear links (Choi & Park, 2016, in South Korea), missing links (Guill & Bos, 2014; Guill et al., 2020a, both in Germany; Smyth, 2008b, in Ireland), and a threshold amount of instruction time needed for successful learning (Cheo & Quah, 2005, in Singapore; Heinrich & Nisar, 2013, in the USA; Ömeroğulları et al., 2020, in Germany). Conversely, there appears to be a saturation threshold: The positive effect of private tutoring decreases when the hours spent on private tutoring increase (Liu, 2012).

However, in most studies, the amount of tutoring is measured either by the number of months and years (Heinrich & Nisar, 2013; Ömeroğulları et al., 2020) that students have already attended private tutoring lessons (tutoring duration) or—and in most cases—by the number of hours per week students spend with private tutoring (tutoring intensity). As both factors, tutoring duration and tutoring intensity, contribute to the additional learning time made available by private tutoring, they should both be considered when the effects of the amount of tutoring are modeled.

Furthermore, for students who have additional learning time in private tutoring classes, it is mostly unclear whether their total learning time (in and outside school) indeed increases. Some authors (Bray, 2009; Kenny & Faunce, 2004) suggested—based on anecdotal evidence—that tutored students pay less attention during regular school classes and thus spend less time actively learning in that context because they assume that tutoring will allow them to make up for what they have missed. Additionally, students might simply displace time that would otherwise be spent learning at home to private tutoring lessons. In both cases, private tutoring lessons do not increase the active learning time—or at least not as much as might be expected considering the amount of tutoring time. Only a few studies controlled for the learning time at home or at school when analyzing the effects of private tutoring. Kim and Hong (2018) found a positive relationship between private tutoring and academic achievement in South Korea even after controlling for self-study time but they did not control for students' prior academic achievement. Smyth (2008a) also controlled for homework and study time at home but did not find any effect of private tutoring.

Coming back to Carroll's model, not all students are expected to profit equally from private tutoring lessons. Students who are more motivated use the learning time available during tutoring lessons as active learning time and are therefore more likely to improve their academic achievement. Indeed, Kuan (2011) found a positive interaction effect between tutoring attendance and students' motivation to attend tutoring, while others

(Guill & Bonsen, 2010; Guill & Bos, 2014; Guill et al., 2020a) did not find evidence for an interaction between tutoring attendance and students' motivation to learn or their motivation for a specific subject. However, such a general measure of a student's motivation or interest in a specific subject might be too vague a measure of the student's motivation to engage in private tutoring lessons. In line with Carroll, we therefore expected to find a positive interaction between students' motivation to engage in private tutoring and the amount of learning time made available by private tutoring.

Furthermore, according to Carroll's model, more able students or students with higher prior knowledge should be able to profit more effectively from the learning time available as they have a more differentiated knowledge structure, which helps them to integrate new learning material. However, contrary to this claim, in some previous studies, a higher level of prior knowledge was associated with less success achieved through tutoring (Guill & Bonsen, 2010; Guill & Bos, 2014; Kuan, 2011). Guill et al. (2020a), who also found a negative interaction between private tutoring and prior grades, suggested that tutors and students invest more effort in the case of low-performing students if the risk of grade repetition is high.

Based on our adaptation of Carroll's model to private tutoring, we expected to find a positive relationship between academic achievement and private tutoring duration (Hypothesis 1a), between academic achievement and private tutoring intensity (Hypothesis 1b), and between academic achievement and the interaction of both private tutoring duration and intensity (Hypothesis 1c), at least when changes in out-of-school learning time were controlled for. We expected this effect to be stronger for students with higher prior academic achievement (Hypothesis 2a) and for students who were more motivated to attend private tutoring lessons (Hypothesis 2b).

Instructional foci of private tutoring lessons

More time-intensive or longer-term tutoring is not necessarily just "more of the same." Tutoring lessons may have very different instructional foci depending on the conceptions of the tutor and the tutees. Based on earlier descriptive findings on the content of tutoring lessons in Germany (Rudolph, 2002), we differentiated between four core areas of tutoring lessons.

- (1) Tutoring may focus on *homework completion*. Either the tutees do their homework during private tutoring with the tutor's assistance or the tutor checks the tutee's homework. In both cases, the tutor assists the student to fulfill the basic requirements of school lessons. Furthermore, the student can be assured that his or her homework is correct and given confidence to present it in school. Both factors can improve the student's grades. However, if the tutor corrects the student's homework instead of focusing on the student's understanding of the subject matter, private tutoring will not result in improved academic achievement as measured by achievement test scores. Homework support from tutors that can be classified as controlling and interfering even has negative effects on students' homework-related study behavior (Guill et al., 2020b).
- (2) Tutoring may focus on material currently being taught in the classroom and on *test preparation*. If the tutee lacks a deeper understanding of the subject matter, the tutor may train basic mathematical algorithms or grammatical rules, for example, which allow the tutee to master a substantial part of teacher-designed tests and, therefore, to reach a sufficient grade and to avoid grade repetition. In this case, tutoring may result

in better grades although the effects on students' achievement in academic tests—which focus on a deeper understanding and cover a broader range of the curriculum—may be small or even nonexistent.

- (3) In contrast, tutoring might also focus on filling gaps in *prior knowledge* (Thomas et al., 2006). Prior knowledge is essential for students to integrate new learning content, to revise, and to actively construct new knowledge structures. Prior knowledge is therefore a strong predictor of academic achievement (Helmke & Schrader, 2001). If a tutee struggles in school for a long time, his or her prior knowledge is no longer sufficiently elaborated to integrate new content. If tutoring focuses especially on improving prior knowledge, this might not immediately result in better grades but might become visible in achievement tests, which focus not only on current lesson content but also on a broader range of the curriculum.
- (4) Last but not least, in the long run, students should ideally be able to master their learning process independently. To reach this goal, tutoring lessons may focus on developing *study behavior and learning strategies* that are requirements for self-regulated learning processes (Helmke & Schrader, 2001; Ramdass & Zimmerman, 2011). In this case, short-term effects on grades and test achievement might be low but effects on students' cognitive and metacognitive learning strategies can be expected.

While short-term tutoring most probably mainly focuses on short-term goals such as improving the grade of the next and maybe decisive test, long-term tutoring offers more time for long-term goals such as improving the tutee's prior knowledge and study behavior. To the best of our knowledge, however, no research so far has systematically looked at correlations between these various thematic priorities and the duration and intensity of private tutoring, nor at whether such priorities have a differential effect on successful learning. According to qualitative analyses, tutors are aware that tutoring may be more effective in the long run if it focuses on closing gaps in prior knowledge (Mischo & van Kessel, 2005), and tutors therefore sometimes take this approach (Thomas et al., 2006).

Summing up our argumentation, we expected private tutoring that puts a focus on homework completion and test preparation to improve mainly the students' grades (Hypotheses 3a and 3b), private tutoring that puts a focus on prior knowledge to improve mainly the students' test achievement (Hypothesis 3c), and private tutoring that puts a focus on study behavior and learning strategies to improve the students' cognitive and metacognitive learning strategies (Hypothesis 3d).

We expected to find this general pattern of results across all tutoring subjects. Nevertheless, in line with research on the effects of classroom instruction, which is usually subject-specific (Baumert et al., 2010), we believed that it would be useful to analyze the effects of the instructional foci of private tutoring separately for each subject. Grading practices differ between subjects and might influence the instructional foci of tutoring lessons as well as their effects. Similarly, achievement tests, such as those used in large-scale assessments, differ between subjects. They do not capture every aspect of the curriculum and might be differentially sensitive to the instructional foci of private tutoring lessons. We therefore expected subject-specific analyses to be more capable of revealing the effects of the instructional foci of private tutoring.

Materials and methods

Sample

Our analyses relied on two consecutive waves of the KESS Study (Competencies and Attitudes of Students), namely KESS 7 at the beginning of grade 7 in 2005 (Bos et al., 2009) and KESS 8 at the end of grade 8 in 2007 (Bos & Gröhlich, 2010). A complete student cohort of about 14,000 students in the city of Hamburg, a metropolis in Northern Germany, took part in this study. Participation in the questionnaire was voluntary for students and their parents, whereas the achievement tests in various subjects were obligatory for the students. Students attending schools for children with special educational needs and students who did not provide information on private tutoring were excluded from further analyses. Therefore, our analytic sample consisted of $N=8510$ students.

Measures

Private tutoring

After a screening question on ongoing private tutoring (“Are you currently attending private tutoring [in German: *Nachhilfeunterricht*]?”), the eighth-grade students reported for each subject, mathematics (M), English (E), German (both as a native language and as a second language: G), and learning strategies (LS), for how long they had been attending private tutoring and for how many hours per week they were currently attending private tutoring. Tutoring duration was measured by four categories: *up to 3 months*, *4 to 6 months*, *7 to 9 months*, and *more than 9 months*. Tutoring intensity was classified into three categories: *up to 1 h per week*, *up to 2 h per week*, and *more than 2 h per week*. Information on tutoring in German or German as a second language was integrated into one variable by using the larger value on the original variables.

Parents stated in the KESS 8 parents’ questionnaire whether, and if so, for how long their child had been attending private tutoring in mathematics, English, German, or learning strategies in the past 2 years, namely in grade 7 or 8. Following Guill & Bos (2014), this information was dummy-coded to no (0) or any private tutoring (1) in the past for those students who did not mention any ongoing private tutoring.

The tutoring items did not explicitly focus on paid private tutoring but rather relied on an intuitive understanding of the German term “*Nachhilfeunterricht*”, which mostly refers to paid private tutoring. The typical setting in Germany is one-to-one tutoring or tutoring in very small groups of not more than five students. The tutoring setting was not further differentiated in our study.

Concerning activities during tutoring lessons, students stated for one tutoring subject whether each of the following activities was done *not at all*, *rarely*, *sometimes*, or *often*: doing homework, checking homework (homework scale: $\alpha_M=0.51$; $\alpha_E=0.73$; $\alpha_G=0.65$), concentration training and training of study techniques (scale on learning techniques: $\alpha_M=0.76$; $\alpha_E=0.70$; $\alpha_G=0.58$), going over previous tests, preparing for future tests, training content from current lessons and (in the case of language tutoring) vocabulary training (scale on current subject material: $\alpha_M=0.72$; $\alpha_E=0.55$; $\alpha_G=0.80$), and repetition of earlier lessons (item on prior knowledge). If the tutoring content could not be attributed to one subject, it was recoded to missing; unfortunately, this resulted in quite high missing

rates (see 3.4). Furthermore, students stated their motivation to attend tutoring classes on a 3-item scale (e.g., “I get private tutoring lessons because I am not satisfied with my school achievement”; $\alpha=0.80$). Responses were given on a 4-point Likert scale and were reverse-coded (1 = *does not apply at all* to 4 = *fully applies*). Students estimated changes in their learning time at home since the start of their tutoring lessons (e.g., “Since I started to attend private tutoring, I have spent *less* [1]/ *an equal amount of* [2]/ *more* [3] time doing my homework at home”; $\alpha_M=0.77$; $\alpha_E=0.79$; $\alpha_G=0.73$).

Outcome variables

Academic achievement was measured through subject-specific school grades and achievement tests. These two measures differ in their representation of academic achievement. Grades are based on students’ performance in teacher-designed tests and during oral discussions in the classroom. They are influenced by student characteristics such as study habits, effort, or classroom behavior, but standardized test scores are not and are therefore more objective (Marsh et al., 2005). However, grades are the only feedback students get regarding their achievement and they are a decisive factor for grade repetition. The grades in mathematics, German, and English at the end of grade 8 were reported by the teachers and reverse-coded for analyses (1 = *fail* to 6 = *very good*). The mathematics achievement test was composed of different subdomains, for example, arithmetic, algebra, and geometry; for German, a reading test was administered. In English, student’s knowledge of vocabulary, spelling, syntax, and grammar was assessed through a cloze test in which students had to replace the missing half of every fourth word in three different texts (Bos & Gröhllich, 2010). The achievement test scores were scaled via a two-parameter logistic IRT (item response theory) model. Weighted likelihood estimates (WLEs) were calculated for English, reading, and mathematics (Feddermann et al., 2019). To measure the students’ learning strategies, a short version of the WLST 7–12 (Würzburg reading strategies knowledge test for grades 7 to 12; Schlagmüller & Schneider, 2007) was used in KESS 8, measuring the students’ knowledge about the relative superiority of different reading strategies. Unfortunately, the instrument was not available in KESS 7.

Covariates

We controlled for students’ prior knowledge by using the subject-specific midterm grades in the same school year (reverse-coded: 1 = *fail* to 6 = *very good*). In Germany, the school year is subdivided into two terms. Students are graded for their performance during the first term with midterm grades. Insufficient midterm grades are therefore often a reason to start tutoring.

Furthermore, we controlled for students’ prior knowledge by using the test scores at the beginning of grade 7. We added the subject-specific average test score of each school as a covariate because students’ learning outcomes are influenced by the ability composition of their classroom and school (Boonen et al., 2014). Basic cognitive abilities were measured in grade 7 by the figural analogies subscale of the revised cognitive abilities test for grades 4 to 12 (KFT 4–12+R; Heller & Perleth, 2000). In the early 2000s, the Hamburg school system consisted of numerous tracks: Some students attended lower or intermediate secondary school tracks (*Hauptschule* or *Realschule*) or an integrated form of these two tracks (*Integrierte Haupt- und Realschule*, *IHR*). Others attended comprehensive schools

(*Gesamtschule*). In both tracks, IHR and comprehensive school, students attended courses based on their abilities in the different school subjects. The academic track (*Gymnasium*) qualified students for university admission. For each track, we used a dummy variable. Subject-specific academic self-concept, work habits, and interest in mathematics were assessed at the beginning of grade 7. Academic self-concept scales (Marsh, 1990) were adapted from PISA (Programme for International Student Assessment) 2000 for the subjects mathematics ($\alpha=0.92$), German ($\alpha=0.75$), and English ($\alpha=0.79$; e.g., “In the school subject English, I learn quickly”). Work habits in mathematics ($\alpha=0.85$), German ($\alpha=0.83$), and English ($\alpha=0.85$) were assessed with five items each (e.g., “I try to participate in mathematics classes”; Trautwein, 2003). Interest in mathematics was measured with six items (e.g., “I like to calculate using large numbers”; $\alpha=0.83$; Schiefele, 1996) and performance anxiety with four items ($\alpha=0.80$; e.g., “I get frightened when the teacher says that we will write a test”; Valtin et al., 2000). Responses were given on a 4-point Likert scale (1 = *completely disagree* to 4 = *completely agree*). We further controlled for gender (male = 0, female = 1). Concerning the migration background, two dummy variables were coded, indicating whether one or both parents were born abroad (no = 0, yes = 1). Educational background was measured via parents’ secondary school degree, that is, whether they had graduated from the academic track or not (no *Abitur* = 0, *Abitur* = 1) and via their higher education (no university degree = 0, university degree = 1). The socioeconomic status of the family was provided as the highest International Socio-Economic Index of Occupational Status (HISEI; Ganzeboom et al., 1992).

Analyses

Our analyses were conducted in two steps: (1) First, the effect of tutoring duration was analyzed with the data of all students, using the nontutored students as the reference group. (2) Second, we used the subsample of tutored students in each subject and analyzed the effect of tutoring duration, intensity, and content. Furthermore, we checked the effect of the interaction terms between students’ prior knowledge, tutoring duration, and tutoring intensity, as well as between students’ tutoring motivation, tutoring duration, and tutoring intensity. Grades and test achievement were centered at the sample mean of tutored and nontutored students and tutoring motivation was centered at the sample mean of tutored students to ease the interpretation of interaction effects. We analyzed tutoring effects on grades and test achievement.

Due to the multilevel structure of our data with students nested in classrooms and schools, we used the “type=complex” command in *Mplus* 8 (Muthén & Muthén, 1998–2017) to avoid an underestimation of the standard errors of the coefficients. We report standardized regression coefficients (stdyx for continuous predictors and stdy for binary predictors).

For each subject and each dependent variable, we tested the effect of 13 tutoring characteristics (nine for the tutoring of learning strategies), resulting in a substantial risk of alpha inflation. Therefore, we also calculated corrected *p* values by applying the Bonferroni-Holm method (Hemmerich, 2019; Holm, 1979).

As the covariate distribution of tutored and nontutored students showed a large area of common support (Ömeroğulları et al., 2020), propensity score matching was not conducted before the regression analyses.

Missing data

Using the procedure described in an earlier paper Ömeroğulları et al. (2020), missing data in the dependent and independent variables were mostly handled using multiple imputation with the mice package (van Buuren & Groothuis-Oudshoorn, 2011) and miceadds (Robitzsch et al., 2016) in R. We performed multilevel imputation on the school level for test scores and grades ($n = 10$ datasets with 50 iterations). The percentage of missing data varied from 0.01 to 81.5% (the latter percentage is for the variables on work habits), and the average missing rate was 21.3% in the sample of tutored and nontutored students. Missing values were due to the nonparticipation of students in the questionnaire, the multimatrix design of the items, for example, on work habits, and individual nonresponse. The results of the statistical analysis were averaged over 10 imputed data sets according to Rubin's rules (1987). Missing data on tutoring items (duration, intensity, content, and motivation) were not imputed but were handled using full information maximum likelihood estimation (FIML; Enders, 2010) integrated in *Mplus*. Rates were very high, ranging from a mean rate of 42.3% in mathematics to 53.9% in learning strategies, 60.1% in English, and 68.0% in German.

Results

Descriptive results

The portion of students that reported that they were currently attending private tutoring was 18.6%. On a descriptive level, in all subjects, nontutored students outperformed tutored students before and after tutoring (see Table 1) regarding grades, test achievement, and general cognitive abilities. Mean grades before and after tutoring remained stable in both groups. On average, tutored students showed lower motivation in terms of work habits and interest, were found to attend all school tracks, and had the same socioeconomic status in mathematics and English tutoring and a slightly lower socioeconomic status in German and learning strategy tutoring than the nontutored students. Except for tutoring in mathematics, on average, the parents of nontutored students had lower academic degrees than those of tutored students. Girls attended private tutoring in mathematics more often than boys, while boys attended tutoring in German and learning strategies more often than girls. Attendance of English tutoring was the same for boys and girls.

In most cases, students attended private tutoring for up to 6 months and for not more than 2 h per week. They described their motivation to attend tutoring lessons as rather high and, on average, they did not change their learning time at home. The instructional foci were mostly homework and current and earlier lesson content and less often learning strategies. For more details, see Table 2.

The correlations between most of the variables were rather low (see Tables A.1 to A.4 in the supplemental online material), but achievement-related variables such as grades and test achievement had moderate to high correlations. Correlations between the scales on tutoring content and tutoring duration or tutoring intensity were all below $r = .20$.

Table 1 Characteristics of tutored and nontutored students ($N=8510$)

	Mathematics						English						German/Reading						Reading strategy knowledge					
	tutored			nontutored			tutored			nontutored			tutored			nontutored			tutored			nontutored		
	n	SD	M	n	SD	M	n	SD	M	n	SD	M	n	SD	M	n	SD	M	n	SD	M	n	SD	M
End-of-year grade (G8)	3.28	0.85	3.80	1.00	3.44	0.81	3.85	0.95	3.58	0.78	3.93	0.88	3.66	0.77	3.92	0.88			3.11	0.77	3.92	0.77	3.92	0.88
Midterm grade (G8)	3.30	0.92	3.81	1.04	3.35	0.85	3.84	0.95	3.48	0.83	3.91	0.90	3.55	0.83	3.90	0.90								
Test achievement (G8)	-0.01	0.95	0.32	1.18	-0.38	1.19	0.10	1.31	-0.34	1.14	0.42	1.22	28.11	8.09	30.30	8.88								
Test achievement (G7)	-0.40	0.91	-0.07	1.14	-1.18	1.23	-0.76	1.29	-0.79	1.14	-0.08	1.20	-0.74	1.19	-0.10	1.20								
Mean test achievement school level (G7)	-0.07	0.72	-0.12	0.74	-0.88	0.90	-0.79	0.93	-0.43	0.69	-0.10	0.75	-0.40	0.72	-0.11	0.75								
General cognitive abilities (G7)	14.18	6.50	15.73	6.64	14.74	6.51	15.62	6.65	13.57	6.69	15.66	6.62	13.41	6.82	15.63	6.62								
Performance anxiety (G7)	2.24	0.85	2.11	0.83	2.23	0.85	2.12	0.83	2.30	0.88	2.12	0.83	2.28	0.89	2.12	0.83								
Ability self-concept (G7)	2.32	0.94	2.79	0.96	2.75	0.78	3.03	0.76	2.86	0.69	3.10	0.68	2.93	0.72	3.09	0.68								
Work habits (G7)	2.99	0.55	3.14	0.57	3.00	0.54	3.12	0.57	2.95	0.54	3.08	0.55	2.98	0.55	3.08	0.55								
Interest in mathematics/Reading habits (G7)	2.47	0.86	2.74	0.85					2.57	0.88	2.82	0.90												
IHR: lower level	0.01	0.11	0.02	0.14	0.02	0.13	0.02	0.14	0.03	0.17	0.02	0.14	0.03	0.17	0.02	0.14								
IHR: higher level	0.02	0.13	0.02	0.14	0.02	0.14	0.02	0.14	0.02	0.15	0.02	0.14	0.03	0.16	0.02	0.14								
Intermediate track	0.13	0.34	0.14	0.35	0.16	0.37	0.14	0.35	0.17	0.38	0.14	0.35	0.17	0.38	0.14	0.35								
Comprehensive school: lower level	0.17	0.38	0.14	0.34	0.21	0.41	0.14	0.34	0.27	0.45	0.14	0.35	0.26	0.44	0.14	0.35								
Comprehensive school: higher level	0.13	0.34	0.14	0.35	0.12	0.32	0.13	0.34	0.11	0.31	0.13	0.33	0.06	0.25	0.13	0.33								
Academic track	0.48	0.50	0.46	0.50	0.40	0.49	0.47	0.50	0.27	0.44	0.47	0.50	0.31	0.46	0.47	0.50								
Parent(s) qualified for university	0.52	0.50	0.53	0.50	0.48	0.50	0.53	0.50	0.43	0.49	0.53	0.50	0.44	0.50	0.53	0.50								
Parent(s) with university degree	0.32	0.47	0.35	0.48	0.29	0.45	0.35	0.48	0.26	0.44	0.35	0.48	0.27	0.45	0.34	0.48								
Migration: one parent born abroad	0.14	0.35	0.11	0.31	0.12	0.33	0.11	0.31	0.14	0.34	0.11	0.31	0.14	0.34	0.11	0.31								
Migration: both parents born abroad	0.20	0.40	0.22	0.42	0.21	0.41	0.22	0.42	0.26	0.44	0.22	0.41	0.23	0.42	0.22	0.42								
Highest ISEI in family	51.44	15.94	51.25	16.34	49.86	15.91	51.40	16.32	47.46	15.80	51.51	16.30	48.34	16.48	51.39	16.28								
Gender: female	0.61	0.49	0.50	0.50	0.50	0.50	0.52	0.50	0.44	0.50	0.52	0.50	0.42	0.49	0.52	0.50								

G8 grade 8, G7 grade 7, IHR integrated lower and middle secondary school (*Integrierte Haupt- und Realschule*). Grades are reverse coded (1 = fail to 6 = very good)

Regression models

Summary of the effects of private tutoring

Even after controlling for multiple covariates, tutored students showed significantly lower test achievement than nontutored students in all subjects (see Table 3). With regard to grades, students tutored for up to 6 months had lower grades in mathematics than nontutored students, but there was no relationship between tutoring and grades in the other subjects. Tutoring in learning strategies was not related to reading strategy knowledge.

Tutoring characteristics

Within the subsamples of tutored students, we found hardly any systematic effects (see Table 4 and Tables B.1 to B.6 in the supplemental online material). The models were sequenced as follows: We first tested for the main effects of private tutoring (PT) intensity and duration in separate models and in a common model (Model 1: Hypotheses 1a and 1b). Second, we tested for the main effects of instructional foci in separate models and a common model (Model 2: Hypotheses 3a-d). Third, we tested for the interaction between PT duration and PT intensity (Model 3: Hypothesis 1c), the interaction between PT duration/intensity and prior achievement (Model 4: Hypothesis 2a), and the interaction between PT duration/intensity and tutoring motivation (Model 5: Hypothesis 2b). In nearly all cases (for an exception see below) neither PT duration, nor PT intensity, nor PT motivation was related to grades or test achievement at the end of grade 8. No instructional focus of the tutoring lessons had any systematic relationship with the dependent variables.

Concerning mathematics test achievement (see Table B.2), we found a positive interaction between PT duration and PT intensity ($\beta=0.16$), which means that the combination of long-term and intensive tutoring was especially effective (Hypothesis 1c), but the main effect of PT intensity was negative ($\beta=-0.12$). Motivation to attend tutoring had a positive main effect ($\beta=0.15$) after the interaction of motivation and PT intensity was controlled for ($\beta=-0.11$, $p>0.05$). Concerning end-of-year grades in English (see Table 4), we found a positive interaction between prior achievement and PT intensity ($\beta=0.19$), which means that especially high-achieving students profited from more intensive private tutoring lessons.

Concerning private tutoring in German (see Table B.5), we found a negative main effect of tutoring intensity ($\beta=-0.10$) in Model 1. Most interestingly, focusing on homework during private tutoring lessons was also negatively related to achievement in the reading test ($\beta=-0.46$).

Concerning the control variables, as to be expected, midterm grades were the strongest predictor of end-of-year grades. Test achievement in grade 7 predicted grades and test achievement in grade 8. The effects of the motivational variables were rather small and, in many cases, not significant but this was in line with earlier findings. Family background did not have an effect either on grades or on test achievement. Girls had higher grades and test achievement in English and better reading strategy knowledge. There was no gender effect either on mathematics or on German.

Table 2 Duration, intensity, and instructional foci of private tutoring ($N=8510$)

	Mathematics ($N_{\max}=1025$)		English ($N_{\max}=603$)		German/Reading ($N_{\max}=491$)		Reading strategy knowledge ($N_{\max}=311$)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>PT duration</i>								
Up to 3 months	39%	-	38%	-	42%	-	69%	-
4 to 6 months	29%	-	27%	-	26%	-	11%	-
7 to 9 months	9%	-	12%	-	12%	-	8%	-
10 and more months	23%	-	24%	-	20%	-	12%	-
<i>PT intensity</i>								
Up to 1 h per week	43%	-	47%	-	39%	-	40%	-
1–2 h per week	28%	-	23%	-	24%	-	22%	-
More than 2 h per week	10%	-	11%	-	16%	-	14%	-
Motivation to attend PT ¹	3.23	0.75	3.11	0.79	3.08	0.80	3.07	0.80
Change in learning time ²	2.13	0.51	2.09	0.53	2.01	0.46	2.10	0.45
<i>Instructional foci</i>								
Homework ³	3.05	0.77	2.80	1.00	2.41	1.01	-	-
Current lesson content ³	3.42	0.67	3.26	0.58	2.67	0.94	-	-
Learning strategies ³	2.14	1.00	2.22	0.98	2.65	0.96	-	-
Earlier lesson content ³	3.08	0.96	3.10	1.07	2.71	1.16	-	-

¹Range from 1 = *does not apply at all* to 4 = *often fully applies*. ²Range from *less* [1] to *an equal amount of* [2] to *more* [3] time spent doing homework at home. ³Range from 1 = *not at all* to 4 = *often*

Discussion

Summary of findings

In our data set, 18.6% eighth-grade students were currently (i.e., in 2007) attending private tutoring. This proportion is in line with the findings of Hille et al. (2016), who reported that 19% of eighth-grade students attended paid private tutoring in the last 6 months of 2013. So, even though our tutoring items did not explicitly ask about paid private tutoring, this number implies that, in most cases, we captured paid private tutoring.

In line with several previous studies with the same and similar data sets (Guill & Bos 2014; Guill et al., 2020a, b; Ömeroğulları et al., 2020), we did not find any positive effects of private tutoring on academic achievement, either on grades or on test achievement. Nontutored students outperformed tutored students both before and after several months of tutoring. This was the case although we controlled for numerous achievement-related covariates, systematically differentiating between tutored and nontutored students. Furthermore, in detailed analyses, we did not find any effects of either tutoring intensity, tutoring duration, or students' motivation to attend tutoring. With the exception of a negative effect of a focus on homework during German tutoring on reading achievement, we did not find any effect of specific aspects of tutoring content. After alpha inflation was corrected, neither the homework effect nor any interaction effects between tutoring and students' characteristics reached statistical significance. Even before the correction of alpha inflation, the analyses of the interaction effects did not reveal a clear pattern and the pattern of results

Table 3 Effects of private tutoring on grades and test achievement in several subjects ($N = 8510$)

	Mathematics			English			German/Reading			Reading strategy knowledge		
	Grade		Test achievement		Grade		Test achievement		Grade		Test achievement	
	beta	SE	beta	SE	beta	SE	beta	SE	beta	SE	beta	SE
Tutoring^a												
Up to 3 months	-.11	.03	-.12	.02	-.04	.04	-.10	.04	-.03	.05	-.12	.04
4 to 6 months	-.16	.04	-.15	.03	.02	.05	-.10	.05	-.04	.06	-.27	.06
7 to 9 months	-.04	.07	.01	.06	-.10	.07	-.16	.07	-.04	.09	-.20	.07
10 and more months	-.08	.05	-.13	.04	.02	.06	-.14	.05	-.04	.07	-.21	.07
PT at an earlier time point	-.03	.04	-.09	.03	-.05	.04	-.13	.04	-.05	.05	-.04	.05
Midterm (G8) grade	.64	.02			.69	.02			.67	.02		
General cognitive abilities	.04	.01	.10	.01	.04	.01	.07	.01	.03	.01	.10	.01
Test achievement (G7)	.18	.02	.49	.01	.12	.02	.35	.02	.11	.01	.46	.01
Mean test achievement school level	.02	.03	.19	.03	.02	.03	.14	.03	.05	.03	.12	.03
Performance anxiety (G7)	-.03	.01	-.02	.01	-.01	.01	-.02	.01	-.01	.01	-.02	.01
Ability self-concept (G7)	.06	.01	.07	.01	.06	.01	.12	.01	.03	.01	.01	.01
Work habits (G7)	.03	.01	.03	.01	.02	.01	.01	.01	.02	.01	.00	.01
Interest	.00	.01	.01	.01					.02	.01	.09	.01
IHR: lower level ^b	-.10	.07	.08	.05	-.11	.07	.13	.11	-.04	.06	.07	.07
IHR: higher level ^b	-.02	.10	.31	.06	.05	.08	.40	.08	-.09	.09	.30	.06
Intermediate track ^b	-.12	.05	.24	.05	-.06	.05	.46	.05	-.20	.05	.31	.05
Comprehensive school: lower level ^b	.02	.05	.05	.03	.04	.05	.05	.05	-.04	.05	.09	.05
Comprehensive school: higher level ^b	-.21	.05	.27	.04	-.06	.05	.52	.06	-.27	.06	.33	.05
Academic track ^b	-.20	.06	.30	.07	-.10	.06	.74	.08	-.38	.07	.39	.07
Parent(s) qualified for university ^c	.01	.03	.01	.02	-.01	.03	.01	.02	.01	.03	.02	.02
Parent(s) with university degree ^d	.03	.03	.05	.02	.01	.03	.02	.03	.04	.03	.04	.03

Table 3 (continued)

	Mathematics			English			German/Reading			Reading strategy knowledge		
	Grade			Test achievement			Grade			Test achievement		
	beta	SE		beta	SE		beta	SE		beta	SE	
Migration: one parent born abroad ^e	-.05	.03		-.04	.02		.03	.03		.05	.03	
Migration: two parents born abroad ^e	-.07	.02		-.01	.02		.00	.03		-.01	.03	
Highest ISEI in family	.00	.01		.01	.01		.02	.01		.01	.01	
Gender ^f	.06	.02		.02	.02		.07	.02		.13	.02	

Significant coefficients are shown in bold ($p < .05$). Continuous variables are stdyx- and binary variables stdy-standardized. *PT* private tutoring, *G8* grade 8, *G7* grade 7, *IHR* integrated lower and middle secondary school (*Integrierte Haupt- und Realschule*). Grades are reverse coded (1 = *fail* to 6 = *very good*). Reference categories: ^aNo private tutoring, ^bLower secondary track (*Hauptschule*). ^cParents not qualified for university admission. ^dParents without university degree. ^eBoth parents born in Germany. ^fMale

Table 4 Effects of characteristics of private tutoring in English on grades in English at the end of grade 8 ($N=603$)

	Model 1		Model 2		Model 3		Model 4		Model 5	
	beta	SE	beta	SE	beta	SE	beta	SE	beta	SE
PT duration	.01	.04	.02	.04	.08	.09	-.01	.04	.01	.04
PT intensity	.02	.04	.04	.04	.09	.10	.05	.04	.02	.04
Change in learning time	.08	.08	.09	.08	.08	.08	.08	.08	.07	.08
Motivation to attend PT	-.04	.04	-.01	.04	-.04	.04	-.04	.04	-.12	.09
<i>Instructional foci</i>										
Homework			.07	.08						
Current lesson content			-.07	.09						
Learning strategies			-.12	.08						
Earlier lesson content			-.06	.09						
<i>Interaction terms</i>										
Duration $\times$ intensity					-.12	.13				
Duration $\times$ prior achievement							-.11	.07		
Intensity $\times$ prior achievement							.19	.08		
Duration $\times$ motivation									.06	.08
Intensity $\times$ motivation									.03	.09
Midterm (G8) grade English	.64	.03	.62	.04	.64	.03	.62	.03	.64	.03
General cognitive abilities (G7)	.03	.04	.03	.04	.03	.04	.01	.03	.03	.04
English test achievement (G7)	.14	.06	.14	.06	.14	.06	.14	.06	.14	.06
Mean test achievement school level	.03	.08	.03	.08	.03	.08	.04	.08	.03	.08
Performance anxiety (G7)	-.06	.04	-.05	.05	-.06	.04	-.06	.04	-.06	.04
English ability self-concept (G7)	.03	.04	-.01	.05	.03	.04	.04	.04	.04	.04
Work habits English (G7)	.03	.05	.04	.06	.03	.05	.03	.05	.02	.05
IHR: higher level ^a	-.03	.03	-.04	.04	-.03	.03	-.04	.03	-.03	.03
Intermediate track ^a	-.09	.06	-.10	.09	-.08	.06	-.09	.06	-.08	.06
Comprehensive school: lower level ^a	-.03	.07	-.04	.09	-.03	.07	-.03	.07	-.03	.07
Comprehensive school: higher level ^a	-.06	.06	-.07	.08	-.06	.06	-.08	.06	-.07	.06
Academic track ^a	-.16	.11	-.21	.14	-.16	.11	-.17	.10	-.16	.11
Parent(s) qualified for university ^b	.02	.06	.01	.07	.03	.06	.01	.06	.02	.06
Parent(s) with university degree ^c	-.01	.05	-.01	.06	-.01	.05	.00	.05	-.01	.05
Migration: one parent born abroad ^d	-.01	.04	-.01	.04	-.01	.04	-.01	.04	-.01	.04
Migration: two parents born abroad ^d	-.02	.05	-.02	.05	-.02	.05	-.02	.05	-.02	.04
Highest ISEI in family	.05	.05	.06	.05	.05	.05	.05	.04	.05	.04
Gender ^e	.09	.04	.07	.04	.09	.04	.10	.04	.09	.04

Significant coefficients are shown in bold ($p < .05$). Continuous variables are stdyx- and binary variables stdy-standardized. *PT* private tutoring, *G8* grade 8, *G7* grade 7, *IHR* integrated lower and middle secondary school (*Integrierte Haupt- und Realschule*). Grades are reverse coded (1 = *fail* to 6 = *very good*). Reference categories: ^aLower secondary track (*Hauptschule*). ^bParents not qualified for university admission. ^cParents without university degree. ^dBoth parents born in Germany. ^eMale

was only partly in line with our hypotheses. Our results indicate that a strong focus on homework completion during German tutoring does not foster students' reading skills.

Strengths and limitations

Given that, in sum, our findings are disillusioning, a detailed analysis of the specific strengths and limitations of our study and data analyses is important. A major strength of our study is that we were able to analyze a large longitudinal data set with substantial sample sizes of tutored students in several school subjects and even in learning strategies. We had information about tutoring duration and tutoring intensity. Hardly any studies up until now have measured the amount of private tutoring students receive in this much detail.

A further strength of our study is that the data set included variables on tutoring content. We are not aware of another large-scale study that has such a detailed description of tutoring content. However, we were unable to identify any effects of this tutoring content. On a methodological level, this might be because of extremely high missing data rates on tutoring content. Students were asked to choose only one tutoring subject to rate the tutoring content but often failed to follow these instructions and chose at least two subjects. In that case, we could not assign statements about tutoring content to any subject and therefore lost this data. In consequence, the estimates of the effects of tutoring content had high standard errors. Future studies using digital tools could technically force the students to restrict themselves to just one subject. Furthermore, the reliabilities of some tutoring content scales were rather low (<0.60). In future studies, this problem could be solved by adding more items to each scale. However, on a theoretical level, one could discuss whether tutoring content is sufficient to describe tutoring quality. For example, homework support during tutoring can vary a lot, being either autonomy-supportive or controlling, and this has differential effects on students' study behavior (Guill et al., 2020b). Classroom instruction can be described by three dimensions of instructional quality: structure, cognitive activation (or challenge), and support. This model has been successfully adapted to characterize tutoring lessons (Guill et al., 2020a; Bäumer et al., 2011), but a valid measure of cognitive activation during tutoring is still missing. Additionally, future studies should measure the tutoring setting and the tutor's qualification as both aspects might influence tutoring quality.

Based on Carroll's (1963) model of school learning, we identified students' motivation to engage in private tutoring as a moderator of tutoring effects, but we were unable to prove such an effect. Our measure on tutoring motivation focused on the decision to start private tutoring. However, a scale focusing on students' willingness to engage during tutoring lessons might be more helpful to capture the effects of student motivation. This willingness to engage should in turn be related to more general predictors of motivated behavior such as students' self-efficacy, their causal attributions, and their goals (Wigfield & Eccles, 2000).

As also implied by Carroll's model, we tried to control for changes in students' out-of-school learning time. However, this measure depends not only on the students' capacity to evaluate their current learning time but also on their capacity to evaluate the development of that time over a time span of several months and, therefore, might be of limited validity. Deriving changes in learning time from several measures of current learning time before and during tutoring would be a feasible alternative.

Furthermore, although we controlled for systematic differences in prior knowledge between tutored and nontutored students and also within the group of tutored students, there might also be systematic differences concerning cognitive and metacognitive learning

strategies in favor of nontutored or less tutored students. As long as these are not controlled for, tutoring effects might be masked.

Information on students' learning environments outside school and tutoring lessons was limited. Students might regularly be tutored by an older sibling or parent. Theoretically, this support within the family might have an impact on students' achievement similar to that of private tutoring. However, even after controlling for the quality of parental homework support (Guill et al., 2020b) were not able to detect positive effects of private tutoring.

Last but not least, as in all studies using large-scale assessments to analyze the effects of private tutoring attendance, no data are available on short-term declines in academic achievement or motivation just before the decision was made to start private tutoring. Furthermore, smaller improvements within one grade, for example, C (–) to C (+), cannot be measured and general achievement tests might not be sensitive enough to capture small improvements.

Future research outlook

A more detailed diagnosis of students' achievement development when they attend private tutoring is desirable. This was provided by, for example, Mischo and Haag (2002), who used grades in teacher-designed tests instead of end-of-term grades. We still believe that it is important to measure tutoring quantity precisely as tutoring can only have an effect if it is attended for a minimum amount of time.

Concerning the problem of systematic differences between tutored and nontutored students, which we did not control for, more experimental designs are desirable to evaluate the effect of private tutoring. However, students can hardly be randomly assigned to tutoring or nontutoring groups while ignoring the families' preferences. Therefore, given tutoring vouchers to the experimental group, which is a random encouragement design (West et al., 2008), might be a worthwhile alternative.

Research on tutoring content and quality is still in its infancy. As a next step, we propose that multiple perspectives should be integrated, that is, not only the students' but also the tutors'—or even observers'—reports of tutoring quality. Given these limitations to our measures of tutoring motivation and students' out-of-school learning time, we do not feel it is appropriate to discuss the implications of our results for Carroll's model of school learning. However, as there are students who reduced their learning time at home, private tutoring should not only be conceptualized as additional learning time but also as differently spent learning time, that is, learning time spent with a tutor instead of alone. This is an additional reason to focus more on tutoring quality.

Furthermore, the focus should be broadened to include more dependent variables. Although the improvement of academic achievement might be the main goal of attending private tutoring, tutoring may also have effects on students' motivation and their satisfaction with school and family life (Guill et al., 2020a). Given the existing variety of tutoring settings, all results should be validated with samples from different age cohorts, for additional school subjects, and concerning the international context, in different cultural contexts.

Educational implications for parent counseling and education policy

To date, neither the current state of research (Guill et al., 2020a; Park et al., 2016) nor our analyses provide enough evidence to support a recommendation of private tutoring as a generally effective strategy to improve academic achievement. Even after several months of private tutoring, tutored students did not catch up with nontutored students. If the aim of private tutoring is for students to improve at school, students and parents should closely monitor whether the expected improvement really occurs. The granting of public subsidies for private tutoring with voucher programs and similar measures (Bray, 2009) cannot be supported by our findings and should be weighed up carefully against structured, free-of-charge tutoring programs that take place within school as these have been shown to be an effective measure to improve reading and mathematics achievement (Pellegrini et al., 2018; Slavin et al., 2011) independent of the parents' willingness and potential to pay for private tutoring lessons.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10212-021-00581-x>.

Acknowledgements The authors would like to thank Gráinne Newcombe for language editing.

Funding Open Access funding enabled and organized by Projekt DEAL. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) –316567771. This paper uses data from the longitudinal study KESS. The data set was generated by the Free and Hanseatic City of Hamburg through the Ministry of Schools and Vocational Training between 2005 and 2012 and has been provided to the MILES (Methodological Issues in Longitudinal Educational Studies) scientific consortium for a limited period with the aim of conducting in-depth examinations of scientific questions. MILES is coordinated by the Leibniz Institute for Science and Mathematics Education (IPN).

Data will be made available to the Research Data Centre (FDZ) at the IQB presumably in 2022. Documentation of questionnaires is published.

Code availability On request.

Declarations

Ethics approval Ethics approval was not required for this study. But all questionnaires were checked by the Hamburg data security officer.

Consent to participate Achievement tests were mandatory. Questionnaires were voluntary and required the informed consent of the legal guardians.

Consent for publication The informed consent included the use of data for scientific purposes.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133–180. <https://doi.org/10.3102/0002831209345157>.
- Bäumer, T., Preis, N., Roßbach, H.-G., Stecher, L., & Klieme, E. (2011). Education processes in life-course-specific learning environments. *Zeitschrift Für Erziehungswissenschaft*, 14(2), 87–101. <https://doi.org/10.1007/s11618-011-0183-6>.
- Boonen, T., Speybroeck, S., de Bilde, J., Lamote, C., van Damme, J., & Onghena, P. (2014). Does it matter who your schoolmates are? An investigation of the association between school composition, school processes and mathematics achievement in the early years of primary education. *British Educational Research Journal*, 40(3), 441–466. <https://doi.org/10.1002/berj.3090>.
- Bos, W., Bonsen, M., & Gröhlich, C. (2009). *KESS 7 – Kompetenzen und Einstellungen von Schülerinnen und Schülern – Jahrgangsstufe 7* [KESS 7 – Competencies and attitudes of students – Grade 7]. Waxmann.
- Bos, W. & Gröhlich, C. (2010). *KESS 8 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* [KESS 8 – Competencies and attitudes of students at the end of Grade 8]. Waxmann.
- Bray, M. (2009). *Confronting the shadow education system: What government policies for what private tutoring?* UNESCO International Institute for Educational Planning.
- Bray, M. (2014). The impact of shadow education on student academic achievement: Why the research is inconclusive and what can be done about it. *Asia Pacific Education Review*, 15(3), 381–389. <https://doi.org/10.1007/s12564-014-9326-9>.
- Bray, M., & Kobakhidze, M. N. (2014). Measurement issues in research on shadow education: Challenges and pitfalls encountered in TIMSS and PISA. *Comparative Education Review*, 58(4), 590–620.
- Bray, M., Kwo, O., & Jokic, B. (Eds.). (2015). *Researching private supplementary tutoring: Methodological lessons from diverse cultures. CERC Studies in Comparative Education: (Vol. 32)*. Springer.
- Byun, S.-Y., Chung, H., & Baker, D. P. (2018). Global patterns of the use of shadow education: Student, family, and national influences. *Research in the Sociology of Education*, 20, 75–105. <https://doi.org/10.1108/S1479-353920180000020004>.
- Carroll, J. B. (1963). A model for school learning. *Teacher College Record*, 64(8), 723–733.
- Cheo, R., & Quah, E. (2005). Mothers, maids and tutors: An empirical evaluation of their effect on children's academic grades in Singapore. *Education Economics*, 13(3), 269–285. <https://doi.org/10.1080/09645290500073746>.
- Choi, Y., & Park, H. (2016). Shadow education and educational inequality in South Korea: Examining effect heterogeneity of shadow education on middle school seniors' achievement test scores. *Research in Social Stratification and Mobility*, 44, 22–32. <https://doi.org/10.1016/j.rssm.2016.01.002>.
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.
- Entrich, S. R. (2014). Effects of investments in out-of-school education in Germany and Japan. *Contemporary Japan*, 26(2), 71–102. <https://doi.org/10.1515/cj-2014-0004>.
- Feddermann, M., Guill, K., List, M. K., Matthießen, R., Ömeroğulları, M., Stallasch, S. E., et al. (2019). *KESS – Skalierung der Leistungstests [KESS – Scaling of the achievement tests]*. Kiel/Germany: IPN.
- Ganzeboom, H. B. G., de Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.
- Guill, K., & Bonsen, M. (2010). Leistungsvorteile durch Nachhilfeunterricht in Mathematik am Beginn der Sekundarstufe I ? – Ergebnisse der Hamburger Schulleistungsstudie KESS 7 [Achievement gains through private tutoring in mathematics at the beginning of secondary education? – Results from the German (Hamburg) School Achievement Study KESS]. *Unterrichtswissenschaft*, 38(2), 117–133.
- Guill, K., & Bos, W. (2014). Effectiveness of private tutoring in mathematics with regard to subjective and objective indicators of academic achievement: Evidence from German secondary school sample. *Journal for Educational Research Online*, 6(1), 34–67.
- Guill, K., & Spinath, B. (2014). Editorial: Special issue editorial: Effects of private tutoring. *Journal for Educational Research Online*, 6(1), 7–11. Available at: <http://www.j-e-r-o.com/index.php/jero/article/view/405/190>
- Guill, K. & Wendt, H. (2016). Außerschulischer Nachhilfeunterricht am Ende der Grundschulzeit [Private tutoring at the end of primary schooling]. In: H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwip-pert & D. Kasper (eds.). TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von

- Grundschulkindern in Deutschland im internationalen Vergleich (pp. 247–256) Münster: Waxmann. urn:nbn:de:0111-pedocs-140225
- Guill, K., Lüdtke, O., & Köller, O. (2020a). Assessing the instructional quality of private tutoring and effects on student outcomes: Analyses from the German National Educational Panel Study. *British Journal of Educational Psychology*, 90(2), 282–300. <https://doi.org/10.1111/bjep.12281>.
- Guill, K., Bahr, I. A., & Ömeroğulları, M. (2020b). Qualität der Hausaufgabenhilfe in Elternhaus und Nachhilfeunterricht als Prädiktoren des hausaufgabenbezogenen Arbeitsverhaltens [Quality of homework support in the parental home and private tutoring as predictors of homework-related study behavior]. *Psychologie in Erziehung und Unterricht*, 67(4), 313. <https://doi.org/10.2378/peu2020.art08d>.
- Heinrich, C. J., & Nisar, H. (2013). The efficacy of private sector providers in improving public educational outcomes. *American Educational Research Journal*, 50(5), 856–894. <https://doi.org/10.3102/0002831213486334>.
- Heller, K. A., & Perleth, C. (2000). *KFT 4–12+R. Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision* [KFT 4–12+R. Cognitive abilities test for Grade 4 to 12, Revision]. Beltz.
- Helmke, A., & Schrader, F.-W. (2001). School Achievement: Cognitive and Motivational Determinants. In D. L. Sills (Ed.), *International encyclopedia of the social and behavioral sciences / edited by Neil J. Smelser and Paul B. Baltes* (pp. 13552–13556). Elsevier Science. <https://doi.org/10.1016/B0-08-043076-7/02413-X>.
- Hemmerich, W. (2019). *Rechner zur Adjustierung des α -Niveaus: StatistikGuru* [Computer for the adjustment of the α level: StatistikGuru]. <https://statistikguru.de/rechner/adjustierung-des-alpha-niveaus.html>.
- Hille, A., Spieß, C. K., & Staneva, M. (2016). More and more students, especially those from middle-income households, are using private tutoring. *DIW Economic Bulletin*, 6(6). Retrieved from https://www.diw.de/documents/publikationen/73/diw_01.c.526666.de/diw_econ_bull_2016-06.pdf.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
- Jokic, B. (Ed.). (2013). *Emerging from the shadow: A comparative qualitative exploration of private tutoring in Eurasia. NEPC Comparative Studies in Education: Vol. 2*. NEPC - Network of Education Policy Centers.
- Kenny, D. T., & Faunce, G. (2004). Effects of academic coaching on elementary and secondary school students. *Journal of Educational Research*, 98(2), 115–126.
- Kim, S., & Hong, S. (2018). The effects of school contexts and student characteristics on cognitive and affective achievement in South Korea. *Asia Pacific Education Review*, 19(4), 557–572. <https://doi.org/10.1007/s12564-018-9557-2>.
- Klemm, K., & Hollenbach-Biele, N. (2016). *Nachhilfeunterricht in Deutschland: Ausmaß - Wirkung - Kosten* [Private tutoring in Germany: Extension - effect - costs]. Bertelsmann Stiftung.
- Kuan, P.-Y. (2011). Effects of cram schooling on mathematics performance: Evidence from Junior High students in Taiwan. *Comparative Education Review*, 55(3), 342–368.
- Lee, C. J. (2005). Korean education fever and private tutoring. *KEDI Journal of Educational Policy*, 2(1), 99–107.
- Liu, J. (2012). Does cram schooling matter? Who goes to cram schools? Evidence from Taiwan. *International Journal of Educational Development*, 32(1), 46–52. <https://doi.org/10.1016/j.ijedudev.2011.01.014>.
- Manzon, M., & Areepattamannil, S. (2014). Shadow educations: Mapping the global discourse. *Asia Pacific Journal of Education*, 34(4), 389–402. <https://doi.org/10.1080/02188791.2014.969194>.
- Marsh, H. W. (1990). *Self-Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescents' self-concept: An interim test manual and a research monograph*. The Psychological Corporation.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, 76(2), 397–416. <https://doi.org/10.1111/j.1467-8624.2005.00853.x>.
- Mischo, C., & Haag, L. (2002). Expansion and effectiveness of private tutoring. *European Journal of Psychology of Education*, 17(3), 263–273. <https://doi.org/10.1007/BF03173536>.
- Mischo, C., & van Kessel, M. (2005). Wie wirkt Nachhilfe? Mögliche Wirkfaktoren im Wissen der Nachhilfelehrer [How does private tutoring work? Possible effecting mechanisms in the knowledge base of tutoring teachers]. *Empirische Pädagogik*, 19(1), 28–46.
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide*. Eighth Edition. Muthén & Muthén.
- Ömeroğulları, M., Guill, K., & Köller, O. (2020). Effectiveness of private tutoring during secondary schooling in Germany: Do the duration of private tutoring and tutor qualification affect school achievement? *Learning and Instruction*, 66. <https://doi.org/10.1016/j.learninstruc.2020.101306>

- Park, H., Buchmann, C., Choi, J., & Merry, J. J. (2016). Learning beyond the school walls: Trends and implications. *Annual Review of Sociology*, 42(1), 231–252. <https://doi.org/10.1146/annur-ev-soc-081715-074341>.
- Pellegrini, M., Lake, C., Inns, A., & Slavin, R. E. (2018). *Effective programs in elementary mathematics: A best-evidence synthesis*. Online available from: http://www.bestevidence.org/math/elem/elem_math_2018.htm [28.03.2019].
- Ramdass, D., & Zimmerman, B. J. (2011). Developing self-regulation skills: The important role of homework. *Journal of Advanced Academics*, 22(2), 194–218. <https://doi.org/10.1177/1932202X1102200202>.
- Robitzsch, A., Grund, S., & Henke, T. (2016). Miceadds. Retrieved from <https://CRAN.R-project.org/package=miceadds>.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley.
- Rudolph, M. (2002). *Nachhilfe – gekaufte Bildung?* [Private tutoring – education bought?]. Klinkhardt.
- Schiefele, U. (1996). *Motivation und Lernen mit Texten* [Motivation and learning with texts]. Hogrefe.
- Schlagmüller, M., & Schneider, W. (2007). *Würzburger Lesestrategie-Wissenstest für die Klassen 7–12* [Würzburg reading strategies knowledge test for Grades 7 to 12]. Hogrefe.
- Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6(1), 1–26. <https://doi.org/10.1016/j.edurev.2010.07.002>.
- Smyth, E. (2008a). The more, the better? Intensity of involvement in private tuition and examination performance. *Educational Research and Evaluation*, 14(5), 465–476.
- Thomas, J., van Kessel, M., Lohrmann, K., & Haag, L. (2006). Wirkfaktoren im Wissen und Handeln der Nachhilfelehrer – Einzelfallbetrachtungen. [Effecting mechanisms in the thought processes and actions of tutoring teachers—A study based on single cases]. *Psychologie in Erziehung Und Unterricht*, 53(1), 35–43.
- Trautwein, U. (Ed.). (2003). *Schule und Selbstwert*. [School and self-esteem]. Waxmann.
- Valtin, R., Wagner, C., Ostrop, G., & Darge, K. (2000). *Das Forschungsprojekt SABA Plus: Schulische Adaptation und Bildungsaspiration (Fortsetzung im 7. Schuljahr): Bericht für die an der Untersuchung beteiligten Schulen*. Retrieved from Humboldt-Universität zu Berlin website: https://amor.cms.hu-berlin.de/~h0319kfm/download/bericht_schulen_7_klassen.pdf.
- Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67. <https://doi.org/10.18637/jss.v045.i03>.
- West, S. G., Duan, N., Pequegnat, W., Gaist, P., Des Jarlais, D. C., Holtgrave, D., José Szapocznik, J., Fishbein, M., Rapkin, B., Clatts, M., & Dolan Mullen, P. (2008). Alternatives to the randomized controlled trial. *American Journal of Public Health*, 98, 1359–1366. <https://doi.org/10.2105/AJPH.2007.124446>.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>.
- Zhang, W., & Bray, M. (2019). Mapping and analyzing private supplementary tutoring: Approaches and themes. *ECNU Review of Education*, 2(1), 3–6. <https://doi.org/10.1177/2096531119840865>.

Karin Guill. IPN - Leibniz Institute for Science and Mathematics Education, Olshausenstraße 62, D-24118 Kiel, Germany. E-mail: guill@leibniz-ipn.de

Current themes of research:

Predictors and effects of attending private tutoring (shadow education). Low-achieving students.

Transition points in the educational systems. Effects of tracking and student composition.

Most relevant publications in the field of Psychology of Education:

Guill, K., Lüdtke, O., & Schwanenber, J. (2020). A two-level study of predictors of private tutoring attendance at the beginning of secondary schooling in Germany: The role of individual learning support in the classroom. *British Educational Research Journal*, 46(2), 437–457. doi: 10.1002/BERJ.3586.

Guill, K., Lüdtke, O., & Köller, O. (2020). Assessing the instructional quality of private tutoring and effects on student outcomes: Analyses from the German National Educational Panel Study. *British Journal of Educational Psychology*, 90(2), 282–300. doi: 10.1111/bjep.12281.

Guill, K. & Lintorf, K. (2019). Private tutoring when stakes are high: insights from the transition from primary to secondary school in Germany. *International Journal of Educational Development*, 65, 172–182. doi: 10.1016/j.ijedudev.2018.08.001.

Melike Ömeroğulları. IPN - Leibniz Institute for Science and Mathematics Education, Olshausenstraße 62, D-24118 Kiel, Germany. FAU – Friedrich-Alexander-Universität Erlangen-Nürnberg, Regensburgerstraße 160, D-90478 Nürnberg, Germany. E-mail: melike.oemerogullari@fau.de

Current themes of research:

Transition points in the educational systems. Development of emotions and well-being in school. Expectancy and emotions at school enrolment. Affective characteristics of privileged and disadvantaged children

Most relevant publications in the field of Psychology of Education:

Ömeroğulları, M. & Gläser-Zikuda, M. (2021). Entwicklung affektiv-motivationaler Merkmale am Übergang in die Sekundarstufe – Wie ergeht es bildungsbenachteiligten Kindern? [Development of affective-motivational characteristics at the transition to secondary school – A special focus on disadvantaged students]. *Zeitschrift für Bildungsforschung*. <https://doi.org/10.1007/s35834-021-00314-6>.

Ömeroğulları, M., Guill, K., & Köller, O. (2020). Effectiveness of private tutoring during secondary schooling in Germany: Do the duration of private tutoring and tutor qualification affect school achievement?. *Learning and Instruction*, 66. <https://dx.doi.org/10.1016/j.learninstruc.2020.101306>.

Guill, K., Bahr, I., & Ömeroğulları, M. (2020). Qualität der Hausaufgabenhilfe in Elternhaus und Nachhilfeunterricht als Prädiktoren des hausaufgabenbezogenen Arbeitsverhaltens [Quality of homework support at home and during private tutoring]. *Psychologie in Erziehung und Unterricht*, 67. <https://dx.doi.org/10.2378/peu2020.art08d>.

Olaf Köller. IPN - Leibniz Institute for Science and Mathematics Education, Olshausenstraße 62, D-24118 Kiel, Germany. E-mail: koeller@leibniz-ipn.de

Current themes of research:

The role of school environments for individual development. Large-scale assessments. Educational assessment and educational measurements. Implementation of innovations in schools.

Most relevant publications in the field of Psychology of Education:

Lindner, M. A., Eitel, A., Barenthien, J. & Köller, O. (2021). An integrative study on learning and testing with multimedia: Effects on students' performance and metacognition. *Learning and Instruction*, 71(2021), 101100, 1-9. DOI: 10.1016/j.learninstruc.2018.01.002.

Meyer, J., Fleckenstein, J. & Köller, O. (2019). Expectancy value interactions and academic achievement: Differential relationships with achievement measures. *Contemporary Educational Psychology*, 58, 58-74. DOI: j.cedpsych.2019.01.006.

Saß, S., Schütte, K., Kampa, N. & Köller, O. (2021). Continuous time models support the reciprocal relations between academic achievement and fluid intelligence over the course of a school year. *Intelligence*, 87, 101560. DOI: 10.1016/j.intell.2021.101560.

Teerling, A., Bernholdt, A., Igler, J., Schlitter, T., Ohle-Peters, A., McElvany, N. & Köller, O. (2020). The attitude does matter: The role of principals' and teachers' concerns in an implementation process. *International Journal of Educational Research*, 100, 101528. DOI: 10.1016/j.ijer.2019.101528.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.