Did Students Learn Less During the COVID-19 Pandemic? Reading and Mathematics Competencies Before and After the First Pandemic Wave

Johannes Schult[1], Nicole Mahler[1], Benjamin Fauth[1,2], and Marlit A. Lindner[3]

[1] Institute for Educational Analysis Baden-Württemberg (IBBW), Stuttgart, Germany

[2] Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Germany

[3] Leibniz Institute for Science and Mathematics Education (IPN), Kiel, Germany

Date: January 31, 2022

Author Note

Johannes Schult https://orcid.org/0000-0003-4576-4605 https://twitter.com/schult

Nicole Mahler https://orcid.org/0000-0003-1743-4748

Benjamin Fauth https://orcid.org/0000-0002-4316-4094

Marlit A. Lindner https://orcid.org/0000-0002-6277-1543

Correspondence concerning this article should be addressed to Johannes Schult, Institute for Educational Analysis Baden-Württemberg (IBBW), Heilbronner Straße 172, 70191 Stuttgart, Germany. Email: Johannes.Schult@ibbw.kv.bwl.de

Abstract

The first wave of the COVID-19 pandemic disrupted regular classes in spring 2020. Temporary school closures supposedly led to a considerable learning loss, particularly for low-achieving students. Schools in Baden-Württemberg, Germany, were closed for two months. Although distance learning was implemented, students spent less time learning. Additionally, teachers were faced with organizational and technological challenges of remote learning environments. The present study investigates the competencies of fifth-graders, using large-scale assessment results in reading and mathematics from annual mandatory tests in September (each $n > 80{,}000$). In line with studies from other countries, competence scores were slightly lower in 2020 compared with the three previous years (–0.07 standard deviations for reading comprehension, –0.09 for operations, and –0.03 for numbers). Low-achieving readers managed to attain pre-pandemic competence levels. Regarding mathematics competencies, low-achieving students seem to have a learning backlog that deserves attention in future education. School characteristic such as the average socio-cultural capital and the proportion of students with migration background played a minor role in mediating the schools' learning loss. Still, lower socio-cultural capital was positively associated with larger learning loss in mathematics.

Keywords: COVID-19, school closures, mathematics competencies, reading competencies, large-scale assessments

Introduction

Closing educational institutions is among the most effective non-pharmaceutical interventions to curb the spread of SARS-CoV-2, the virus responsible for the worldwide COVID-19 pandemic (Haug et al., 2020). Since the beginning of the pandemic in early 2020, almost every country has imposed countrywide school closures at some point (United Nations, 2020). In Germany, schools were closed in March 2020 and only gradually re-opened towards the end of the school year (for chronological details see Fickermann & Edelstein, 2020). Further school lockdowns followed in 2021. In spring 2020, elementary schools in the federal state of Baden-Württemberg, for example, were closed for two months. Classes shifted to remote learning during the lockdown of school buildings. There were many challenges for teachers, students and their families, including an unprepared digital infrastructure for virtual classrooms and frequently inadequate learning environments at home (Huber & Helm, 2020). Survey findings indicate that students were less engaged in school-related activities in general and learning activities in particular during the German lockdown in spring 2020 (Grewenig et al., 2021). Even though international studies suggest detrimental lockdown effects on student achievement (Engzell et al., 2021; Maldonado & De Witte, 2021), the city of Hamburg, Germany, reports no systematic long-term loss in learning achievements (Depping et al., 2021). Divergent results in these different reports might be due to differences between the studies in terms of assessment dates (e.g., in relation to school lockdowns during the pandemic), tested samples, implemented achievement tests, and other social and educational factors.

Educational management, policy making and educational sciences in times of the pandemic require a broad empirical base. The aim of the present study is to broaden the empirical base with long-term data from the German federal state of Baden-Württemberg and to provide insights from an educational monitoring perspective. Using data from a statewide mandatory large-scale educational assessment, we compare cognitive

competencies of students at the beginning of their first year in secondary school in grade 5 before and after the first wave of the COVID-19 pandemic. More specific, we compare students' test scores in reading and mathematics from the years 2017 to 2019 (pre-pandemic assessment) with the scores after the first school lockdown in Germany took place in the pandemic year 2020. Thus, our population data provide valuable insights into students' competence development following the pandemics' effects on formal education and beyond.

The Development of Student Achievement During the First Wave of the Pandemic

The unprecedented school lockdown changed the daily lives of students in many and profound ways. Expected consequences include less time spent on learning, increased stress, different and possibly fewer interactions with peers and teachers, reduced learning motivation, (more) distance learning, and less healthy nutrition (Di Pietro et al., 2020). According to a survey of 1,099 German parents of school-aged children (Grewenig et al., 2021), students spent considerably less time on learning activities during the lockdown (3.6 hours on average) than they did before (7.4 hours). This reduction was more pronounced for low-achieving students than for high-achieving students. There were no differences between students from academic or non-academic families. What made a difference was higher income, which can be associated with more time spent on learning activities (in a sample of English children; Andrew et al., 2020).

In some cases, however, there might be positive effects of school closures on learning. Additional cognitive, motivational, and financial support on the parents' side have the potential to be even more conducive to learning than regular classes (e.g., Immerfall, 2020). Still, such positive effects, if they exist, do not offset the detrimental consequences of closed schools, as indicated by an overall loss of learning in recent studies (Engzell et al., 2021; Kuhfeld et al., 2020a).

German schools offered remote learning which often had qualitative implications because they focused more on rehearsal of previously learned content. A lack of learning

spaces at home, poor internet connection and technical equipment were further issues which thwarted adequate (teaching and) learning quality (Grewenig et al., 2021). The variability of distance learning quality and content makes it difficult to pin down the causal effect of school lockdowns on student achievement. However, the analysis of educational large-scale assessments can provide an estimate of the overall pandemic effect and delivers a broad data base for an evaluation of the current status of students' achievement levels as compared to their respective peers in pre-pandemic years.

Currently, studies from the United States and from several European countries provide cross-sectional or longitudinal estimates of how the pandemic affected students' academic competencies. Moreover, they investigated whether the effects differ for students with different family backgrounds, focusing on migration background and socio-economic variables. These variables have repeatedly been shown to affect student learning. In general, socio-cultural capital is associated with resources that may help students to learn at home (Andrew et al., 2020). Migration background is associated with lower socio-economic status and more problems of understanding the language of instruction (Kempert et al., 2016).

A large-scale study of potential lockdown effects based on data from about 10 % of U.S.-American public schools indicates attenuated learning gains in 2020 compared to the previous year (Kuhfeld et al., 2020a). The intra-individual learning loss was small in reading (e.g., –2 percentile ranks for grade 4, $n > 325.000$) and more pronounced in mathematics (e.g., –10 percentile ranks for grade 4) across grades 4 to 8. The inter-individual comparison of the samples of fall 2020 and fall 2019 showed similar differences (with each $n > 344,000$ for grade 4). The differences in achievement associated with school closures were not related to socio-economic status.

Eight weeks of closed schools in the Netherlands in spring 2020 were associated with a loss of learning of 3 percentile ranks ($n > 250,000$, grades 4 through 7; Engzell et al., 2021). Dutch elementary school students had taken standardized exams just before the

schools were closed and soon after they were re-opened. Students' learning trajectories were then compared with data from the three previous years. Learning loss estimates were slightly larger for mathematics (0.14 standard deviations) than for reading (0.09 standard deviations; Engzell et al., 2021, p. 85). Fourth-graders showed a learning loss that was similar to the total sample (i.e., 3 percentile ranks). The effect was larger in schools with a high proportion of students with a migration background and in schools with a high proportion of students from non-academic families. In another study of Dutch students, learning trajectories in a formative assessment were compared for the school years 2018/19 and 2019/20. The assessment was administered online throughout the year in grade 1 to 6 and addressed mathematical skills as well as reading and spelling. In grade 4 (n = 25,336), there was a net increase of 0.2 standard deviations in 2019/20 for all domains combined with a particular spike during the lockdown phase (Meeter, 2021). During the lockdown and afterwards, students in this particular sample spent more time on the formative assessment tasks than the cohort from the previous year, presumably compensating the lack of classroom activity while schools were closed.

Schools in Belgium were closed for nine weeks in spring 2020. The comparison of test scores from June 2020 with test scores from the previous year showed a learning loss of 0.19 standard deviations in mathematics and of 0.23 standard deviations in Dutch (Maldonado & De Witte, 2021). The analysis was based on a sample of $n \geq$ 1,287 sixth-graders in Catholic schools in Flanders. For schools, participation in the assessment was voluntary. In 2020, the number of participating schools was reduced by more than 50 % compared to previous years. This sampling difference might compromise the validity of the comparison between the pandemic-affected cohort with the pre-pandemic cohorts to some extent. Results showed that disadvantaged neighbourhoods and foreign language spoken at home induced a higher estimated learning loss in the domain of reading but not in the domain of mathematics.

Currently, the only findings regarding the effects of the pandemic on German elementary schools pertain to the city of Hamburg (Depping et al., 2021). Test performances of incoming fifth-graders in 2020 were compared to test performances of incoming fifth-graders in 2019 (each year's n > 13,000). Surprisingly, students' reading scores were actually higher in the pandemic year 2020 compared to the previous year (0.05 standard deviations). Mathematics scores were slightly lower in 2020 compared to 2019 (0.02 standard deviations). The learning loss was most pronounced in schools with learners from a disadvantaged social background (e.g., a decrease of 0.21 standard deviations in mathematics). These findings suggest that remote teaching along with additional courses during summer break managed to compensate negative school lockdown effects. It should be mentioned, however, that test booklets with responses to less than 20 % of the test items were excluded from the analysis, which probably underrepresents weaker students in the reported sample. The positive pre-post-difference in reading might therefore be partly due to a larger number of excluded booklets in 2020 (5.5 %) than in 2019 (3.4 %).

Meta-analytic and systematic reviews, which included a preprint of the present study, estimate average learning losses between d = –0.09 (in reading; Hammerstein et al., 2021) and d = –0.17 (in mathematics; Zierer, 2021) during the first wave of the COVID-19 pandemic in the year 2020. Moreover, learning losses were larger, at least in reading tests, for schools with a large proportion of students with a migration background (Maldonado & De Witte, 2021) and for students from less-educated homes (Engzell et al., 2021; Maldonado & De Witte, 2021).

The Present Study

The state of Baden-Württemberg administers a statewide mandatory assessment of competencies in reading and mathematics at the beginning of secondary school in grade 5 ("Lernstand 5"). Students completed this assessment at the regularly scheduled time in September 2020, which allows unbiased comparisons with previous waves of the

assessment. By looking at students' achievements at the beginning of the new school year (and not directly after schools re-opened), one gets a clearer picture of potential long-term learning losses that includes compensation efforts after schools re-opened. After schools had closed, German teachers were initially instructed to repeat material that had already been learned and to postpone the introduction of new learning contents (Fickermann & Edelstein, 2020). There were also optional lessons during the last two weeks of summer break to remedy deficits of low-achieving students.

Thus, the present data provide further insights into the potential differences between the first (school) year of the COVID-19 pandemic and previous pre-pandemic years, extending the international findings discussed above and contributing additional insights into schools that were temporarily closed. Since the assessment took place at the very beginning of secondary school, the effects reported here mainly refer to what happened to students' learning achievements in their last year of elementary school (i.e. grade 4).

The present study investigates the competencies of all fifth-graders in one German federal state in two domains before and after the first pandemic wave, drawing on data that allow taking three important perspectives on changes of (1) population means, (2) high-, average-, and low-achieving groups, and (3) relations to school characteristics.

Research Questions (RQ)

RQ1: How do the competencies of incoming fifth-graders in reading and mathematics differ between the pandemic year 2020 and the three previous years?

RQ2: How do the competence levels of high-, average-, and low-achieving student groups differ between the pandemic year 2020 and the three previous years?

RQ3: How does the difference between student performance levels before and after the first school lockdown relate to school characteristics, specifically the proportion of students with migration background and the schools' socio-cultural capital?

Method

Study Design and Participating Student Cohorts

The present study is based on "Lernstand 5", an ongoing large-scale assessment that was developed by the former State Institute for School Development to provide teachers and schools with information regarding the competencies of the incoming fifth-graders in Baden-Württemberg (each year's n > 80,000). The aim of "Lernstand 5" is to assess to what degree students have achieved selected elementary competencies taught in elementary school and to offer (optional) tailored coaching materials (Fischer et al., 2017; Schulz et al., 2017). The assessment was introduced in 2015, featuring competence tests in German and mathematics at the very beginning of secondary school (i.e., grade 5). The tests are mandatory for all public secondary schools in the federal state. Each year's test consists exclusively of new items. Teachers administer the tests on a day of their choice during the second and third week of the school year, typically in the second half of September. Participation is voluntary for students with special needs. Students who are absent on the day of the test (e.g., due to illness) do not participate. Participation rates for each year and test domain are listed in Table 1. Teachers mark the tests using standardized coding instructions. The test results are not affecting grades in the classroom, making "Lernstand 5" a low-stakes assessment.

--- insert Table 1 here ---

Schools in Baden-Württemberg were closed on March 17, 2020 due to the COVID-19 pandemic. Classes in grade 4 resumed on-site teaching two months later on May 18, 2020. Class schedules rotated for the remaining time of the school year, resulting in a total of ten weeks of school absence for students in grade 4 (in addition to four weeks of regular holiday periods). For the duration of the lockdown, emergency childcare facilities were established for children whose parents work in essential services. The amount and frequency of distance learning opportunities varied; students from families with low socio-economic status tended to have less contact with their class (Immerfall, 2020).

Measures

"Lernstand 5" contains two short screening tests (reading speed and arithmetics) and three longer competence tests (reading comprehension, operations, and numbers). The following analysis is based on reading comprehension, operations, and numbers because these three tests were linked across the assessment years, allowing for a comparison of competencies on the same scale from 2015 to 2020. Table 1 contains reliability estimates and the number of items that were administered in each year.

A pilot study preceded each main data collection for "Lernstand 5". Each pilot study took place 14 months before the assessment, using a sample of around 8,000 elementary school children at the end of grade 4. The tests in the pilot studies contained new items along with a set of linking items that remained constant throughout the years.

Reading Comprehension.

The "Lernstand 5" reading comprehension test comprised four texts and between 32 and 38 items (50 minutes). It assessed different reading processes such as retrieving explicitly stated information and ideas, making inferences, interpreting ideas and information, and evaluating content and textual elements (Fischer et al., 2017).

Mathematics Competencies.

The "Lernstand 5" mathematics tests contained a test of the domain "operations" (20 minutes), and a test of the domain "numbers" (20 minutes). The operations test assessed competencies regarding the application and combination of arithmetic processes. The level of difficulty ranged from understanding simple operations in clear situations to applying multistep operations embedded in problem solving tasks. The numbers test assessed competencies regarding the understanding, interpretation, and application of different representations (numeric, verbal, figural) of natural numbers up to one million (Schult & Lindner, 2019; Schulz et al., 2017).

School Characteristics.

The proportion of students with migration background (at the grade level) was extracted from the official school statistics of the school year 2019/20[1]. The respective proportions for grade 5, 6, and 7 were obtained for each secondary school in the data, because they correspond to the three cohorts that were used as pre-pandemic comparison group. The three proportions (grade 5, 6, and 7) were averaged to obtain an estimate for each school (Cronbach's alpha = .95). Migration background was defined as having a foreign nationality, having a foreign place of birth, or speaking a different language than German at home.

Socio-cultural capital was estimated for each school based on the average number of books at the homes of its student body. This information was obtained at a school level from reports by Baden-Württemberg's eighth-graders in March 2019 and March 2020 during the VERA 8 assessment ($r_{2019,2020}$ = .88; for study details see Schult & Wagner, 2020). Responses were given on a six-point scale: 1 = "0–10 books", 2 = "11–25 books", 3 = "26–100 books", 4 = "101–200 books", 5 = "201–500 books", and 6 = "more than 500 books". We could link the background data to the competence scores at the school level.

Data Preparation and Data Analysis

Within each domain, competence test scores were placed on a common scale using a set of linking items administered in each pilot study (see section 2.2). Each domain's population mean in 2015 was set to 500 and the standard deviation to 100. The difficulty parameters of the linking items were estimated in 2014 after the first pilot study and retained for the scaling of all subsequent pilot studies (fixed parameter calibration; Kolen & Brennan, 2014, p. 182–183). The competencies in reading comprehension, operations, and numbers were modelled using the unidimensional one-parameter logistic model known as Rasch

[1] Data for the school year 2020/21 was not available yet.

model with the marginal maximum likelihood method (de Ayala, 2009). A set of ten plausible values was drawn to estimate each student's ability. Plausible values yield unbiased means and variances (Wu, 2005). We repeated the analysis for each plausible value and pooled the results. Because the standard deviations differ somewhat between the years, we computed Cohen's *d*, the effect size indicating the standardized difference (Lakens, 2013).

The dominant way to interpret these standardized mean differences in education is to compare the effect size with the average gain during a year of schooling for the respective age group. Based on large norm samples of various educational tests, Bloom (2008) estimated that the annual reading gain between grade 4 and grade 5 is $d = 0.40$ for U.S-American students. The corresponding average gain in mathematics tests is $d = 0.56$. Studies with samples of German fourth-graders suggest larger average gains within one year in the range between $d = 0.61$ (reading) and $d = 0.71$ (mathematics; Ditton & Krüsken, 2009; Fuchs & Brunner, 2017; Krüsken, 2007; Robitzsch et al., 2011; Rudolph et al., 2016; Wendt et al., 2017). With an annual learning of 0.65, the loss of learning of 0.08 (as reported by Engzell et al., 2021) would correspond to an eighth of a year of schooling (0.65 / 8 = 0.08).

All analyses were run in R (R Core Team, 2021) using the packages *eatRep* (Weirich et al., 2021), *ggplot2* (Wickham, 2016), *plyr* (Wickham, 2011), *psych* (Revelle, 2020), *tidyr* (Wickham, 2020), and *TAM* (Robitzsch et al., 2020). The analysis uses entire population cohorts. Even though there is no need to make inference from samples to these populations, we included complementary significance tests at the significance level of alpha = .01 for mean comparisons and correlations.

Descriptive statistics were computed for all years. Moreover, the results from 2017 to 2019 were combined to obtain a pre-pandemic comparison group. Choosing the three previous years as comparison group is in line with the approach taken by Engzell et al. (2021) and by Maldonado and De Witte (2021). Using more than one year for comparisons

reduces possible outlier effects. Using a higher number of pre-pandemic years for comparisons, however, could have introduced bias associated with more general long-term trends in student achievement (Stanat et al., 2017).

Comparing Population Means (RQ1).

Population means in reading comprehension, operations, and numbers were calculated for each year as well as for the pooled cohorts of 2017, 2018, and 2019. To answer RQ1, the standardized difference between 2020 and the three previous years combined was calculated for each of the three competence tests.

Comparing High-, Average-, and Low-Achieving Groups (RQ2).

Population percentiles at 5 %, 25 %, 50 %, 75 %, and 95 % were estimated in reading comprehension, operations, and numbers for each year as well as for the pooled cohorts of 2017, 2018, and 2019. To answer RQ2, the location of each percentile in 2020 and in the pooled group of the three previous years combined was compared for each of the three competence tests.

Exploring School Characteristics (RQ3).

The average test score $\overline{\theta_i}$was calculated for each school *i* for each of the three domains. The average test score for 2017, 2018, and 2019 was subtracted from the average test score for 2020 to indicate the change from before the COVID-19 pandemic to after the first pandemic wave: $\delta_i = \bar{\theta}_i(2020) - \bar{\theta}_i(2017, 2018, 2019)$. To answer RQ3, this change score was correlated with the proportion of students with migration background and with socio-cultural capital (average number of books at home). Furthermore, average school competence scores were estimated for each quartile of the two school variables.

Some schools comprised two or more types of schools (*Schulverbünde*). It was not possible to retrospectively assign their classes to the correct school type for the years 2015 to 2019. In addition, some schools changed their type of school, some were permanently

closed, and some were newly opened during those four years. These schools (k = 152) were excluded for the analysis pertaining to RQ3, leaving k = 1065 schools for the analysis.

Results

Comparison of Population Means (RQ1)

The descriptive statistics of the competence assessment for all years are listed in Table 2. The comparison of 2020 with the three previous years shows that the competencies of incoming fifth-graders were on average lower after the first COVID-19 wave than before (RQ1). The standardized difference was d = –0.07 for reading comprehension, d = –0.09 for operations, and d = –0.03 for numbers (all p < .001).[2]

--- insert Table 2 here ---

Comparisons for High-, Average-, and Low-Achieving Groups (RQ2)

Comparisons of selected percentiles revealed a mixed pattern (RQ2; see Figure 1 and Table 3). In reading comprehension, low-achieving students (i.e., the bottom 5 %) scored beneath a similar cut level (2017–2019: 299 vs. 2020: 300). The other cut scores at the 25 %, 50 %, 75 %, and 95 % percentile were 9 to 17 points lower in 2020 than in the three previous years, indicating lower achievement scores for medium- to high-achieving students. In operations, high-achieving students (i.e., those above the 95 % percentile) scored above a higher cut level in 2020 (2017–2019: 656 vs. 2020: 667). Students in the groups with lower levels of competence at the 5 %, 25 %, 50 %, and 75 % percentile had scored 1 to 28 points less in 2020 than in the three previous pre-pandemic years. In numbers, cut levels were similar for the 50 %, 75 %, and 95 % percentiles and 5 to 7 points lower for the 25 % and 5 % percentiles. Taken together, the differences in reading comprehension were larger for

[2] The year-to-year comparison of 2020 with 2019 shows standardized differences of d = +0.07 for reading comprehension, d = –0.07 for operations, and d = –0.02 for numbers.

high-achieving students, whereas in mathematics they were larger for low-achieving students.

--- insert Figure 1 and Table 3 here ---

School Change Scores and School Characteristics (RQ3)

The correlation between schools' change scores in the two mathematics competence domains were large ($r = .50$, $p < .001$). The correlations between the reading change score and the mathematics change scores were medium ($r = .26$ and $r = .29$, both $p < .001$). This suggests that learning losses (or gains) at a school level could differ considerably between reading and mathematics.

The bivariate correlations of the change scores with migration background and socio-cultural capital are listed in Table 4. Schools with higher socio-cultural capital tended to have more positive school change scores in operations ($r = .36$, $p < .001$). The remaining correlations between the school characteristics and the change scores were weak and heterogeneous ($|r| < .3$; $p < .001$, except for numbers both $p > .01$; RQ3). However, the correlations between the two background variables with each other ($r = –.71$, $p < .001$) as well as with the average school competence scores ($|r| > .65$, $p < .001$) were very large. Consequently, the average school change scores at different quartiles of the background variables, which are shown in Figure 2 (and listed in Table 5), were similar to the differences shown in Figure 1.

--- insert Table 4 and Figure 2 and Table 5 here ---

Discussion

During the first COVID-19 wave in spring 2020, the roughly two months with closed schools disrupted regular education in Germany. The purpose of this study was to gain a better understanding of student competencies at the beginning of secondary school in 2020 compared to the pre-pandemic school years by investigating the score differences across cohorts in a mandatory annual large-scale assessment in Baden-Württemberg, Germany. The

reading and mathematics competencies of incoming fifth-graders were slightly lower after schools closed for two months in 2020 than they were compared to the average of the three previous years (2017–2019). These cohort differences are in line with the expectations that closed schools and further pandemic-related issues (e.g., increased absentism; Sosu & Klein, 2021) would prevent students from learning as much as they would have learned in an uninterrupted tuition system, presumably due to less time being spent on learning activities, fewer interactions with peers, and less direct instruction and feedback from teachers (Andrew et al., 2020; Grewenig et al., 2021). The results provide further evidence that student competencies were attenuated after schools had been closed for about a fifth of the school year (RQ1). The extrapolation of learning loss estimates from summer breaks and (partial) absentism yields expected learning gains for the school year 2019/20 that correspond to 37–80 % of the learning gains expected for a typical school year in reading and 67–82 % in mathematics (for grade 4; see Kuhfeld et al., 2020b, Table C5). The annual learning gains in the present study were larger than these projections, ranging from 83–89 % in reading to 84–87 % in numbers and 95–96 % in numbers.[3] In summary, the cohort differences translate to roughly one month of learning backlog.

Studies that compared student performance declines directly after the re-opening of schools tend to show larger deficit estimates (Engzell et al., 2021; Kuhfeld et al., 2020a; Maldonado & De Witte, 2021). The “Lernstand 5” competence assessments took place at the beginning of the new school year in September 2020. Effects of distance learning as well as additional remedial teaching after schools re-opened are therefore also reflected in the assessed test performances. Data from Hamburg, Germany, from August/September 2019 and September 2020 show similar results (Depping et al., 2021). Like the findings from Hamburg, the present analysis shows an unexpected positive difference for reading

[3] The percentages were calculated using the effect sizes from Table 2 along with the annual gain estimates from Germany (reading: $d = 0.61$; mathematic: $d = 0.71$; see above) and Bloom (2008; reading: $d = 0.40$; mathematics: $d = 0.56$) as baseline, respectively.

competencies between 2019 and 2020. A possible explanation might be that students increased their reading skills by spending more time on reading activities during the school lockdown. However, the positive difference in Hamburg might be partially related to an increased number of excluded booklets with very few responses, whereas the positive difference in Baden-Württemberg could be somewhat related to particular test characteristics and other fluctuations in the year 2019. Including the years 2017, 2018 and 2019 in the pre-pandemic comparison group in the present study shows more plausible negative difference, though.

The separate analysis of three competence domains across different competence percentiles provides further insights. The lockdown findings from Baden-Württemberg show slightly diverging patterns for reading and mathematics, depending on the level of student achievement. This may indicate that school closures are not followed by a uniform loss of learning across all groups of students but rather by a heterogeneous learning backlog (RQ2). Overall, the present findings suggest that teachers' (as well as parents' and students') efforts to adapt to the disruption were often beneficial, because the cohort differences turned out to be smaller than the estimated short-term loss of learning during the actual school lockdown (Engzell et al., 2021; Maldonado & De Witte, 2021).

Differences in mathematics assessment scores before and after the first wave of the pandemic were on average closer to zero in the numbers domain than in the operations domain. This may reflect that teachers put a stronger emphasis on arithmetic tasks that students could practice at home on their own during the pandemic year 2020. In contrast, there may have been less time than usual to focus on the application of operations in tasks that require multistep problem solving and a translation of mathematical relations to real world settings. The group of high-achieving students managed to attain similar competence levels compared to previous cohorts regarding operations. On the other hand, average- and low-achieving students showed less understanding of operations than previous cohorts. A

different pattern can be found in the domain of reading comprehension, where low-achieving students attained similar competence levels in 2020 compared to the previous years. A possible explanation could be that the lockdown forced teachers and students to spend more time with basic reading tasks, including distance learning via email, chat, and work sheets.

School characteristic such as the average socio-cultural capital and the proportion of students with migration background did not show substantial relationships with schools' competence change scores (RQ3). This result is in line with findings from Hamburg, where social school characteristics played only a minor role in the explanation of school differences (Depping et al., 2021). As expected, socio-cultural capital was positively associated with slightly larger school change scores in mathematical operations and, to a lesser degree, numbers. Contrary to previous findings (e.g., Maldonado & De Witte, 2021), students' migration background at the school level showed a small positive relation to the reading change score whereas socio-cultural capital had a small negative relation to the reading change score. Overall, the school background variables were highly correlated with the average school competence scores. It was therefore not possible to statistically isolate specific effects of school characteristics over and above competence scores.

The present study is based on mandatory standardized tests that had been administered shortly before the second COVID-19 wave in Germany caused further disruptions in regular education. The data provide a unique perspective on the effects of the pandemic on student competencies after two months of closed schools and distance education. Specifically, differences could be shown for lower levels of student achievement. Judging from the comparably low 5 % percentile in 2020, the operations domain in mathematics deserves more attention during (and after) potential future school lockdowns. The variability in test scores between students has already been large before the COVID-19 pandemic (see 2015 to 2019 in Figure 1). Competencies within one year – and even within many individual schools – are spread across more than three standard deviations. This

heterogeneity within schools may well pose a greater challenge to teachers than a slight pandemic-related attenuation of the schools' mean competence level. In order to help teachers cope with large student competence variances in their classes, the "Lernstand 5" assessment is directly linked to remedial teaching modules. These modules help teachers support students to improve their competence levels with material that is specifically matched to serve the individual's educational needs.

Limitations

The present study design does not allow a causal interpretation. Several potentially confounding aspects must be taken into account. The analysis is based on between-student comparisons. Thus, the results do not reflect intra-individual learning gains. The background characteristics were only available at the school level. The analysis of school characteristics (RQ3) cannot be translated directly to individual students without risking an ecological fallacy. Moreover, the average number of books at home was an estimate from previous student cohorts. However, the measure showed little fluctuation across cohorts, and can thus be used to approximate the average socio-cultural capital of the assessed schools.

As shown in Figure 1, there was a considerable year-to-year fluctuation in the pre-pandemic years. The 2020 scores may reflect not just effects of school closures but also specific variance introduced by the particular tests that were used in 2020 despite the standardized development and scaling. Still, the present study presents important results regarding students' achievement that are necessary to conduct meta-analytic investigations. Combining the present findings with other large-scale studies yields learning loss estimates that are slightly larger than our effect sizes (d = –0.10 in native language domains such as reading; d = –0.17 in mathematics; Zierer, 2021).

Schools are places of learning. In Germany (and most other countries), the curriculum includes factual knowledge, cognitive competencies (such as performing tasks in reading and mathematics) as well as social skills and critical thinking. Closing schools

disrupted not only learning and instruction, but also peer interactions and everyday family life. With its focus on specific competencies, the present study only provides insights regarding changes of reading and mathematics competencies. However, school closures do not only affect students' acquisition of domain specific competencies. Thus, additional research should also focus on potential changes in well-being and social relatedness of both students and teachers.

In Germany, students with low achievements often have to repeat a grade (n = 262 after the school year 2018/19 in Baden-Württemberg). In the wake of the first COVID-19 wave, such mandatory grade repetitions were suspended. This might also have led to a larger than usual number of low-achieving students in the 2020 cohort. Assuming that the additional students are on the lower end of the empirical ability distribution, this effect may theoretically account for up to 0.01 standard deviations.[4]

Implications for Teaching

When global emergencies such as the COVID-19 pandemic trigger the political decision to close schools, educational psychology offers clear recommendations how classes can continue educating their students in beneficial ways. Some recommendations are simple but may require substantial resources. For example, Andrew et al. (2020) suggest that adding one extra hour of instruction per week during the next school years might offset the pandemic effects by the time today's fifth-graders graduate from high school. The digital infrastructure at many schools in Germany might need to be improved in order to facilitate distance learning and digitally augmented education in general (Huber & Helm, 2020).

Online learning benefits students during a school lockdown. In a study of 1,835 Chinese ninth-graders, learning loss during a seven-week school lockdown was 0.22 standard deviations smaller for students with online lessons than for students with no formal instruction (Clark et al., 2021). Formative online assessment seems to be a promising tool to

[4] The estimate is based on a simulation of 100 data sets with M = 500, SD = 100, and n = 80,000, in which the bottom 262 cases were duplicated.

engage students while schools are closed. Empirical studies of formative assessments during the first COVID-19 wave showed a performance increase compared to the regular classes in previous years (Meeter, 2021; for older students see Spitzer & Musslick, 2020). Still, the exact causal links remain unclear, because the samples were self-selected and the interventions were at the same time the competence assessment used as outcome (which facilitates method effects).

More generally, learning and instruction need to be adapted to distance learning with a focus on cognitive activation, student support, and classroom management (Voss & Wittwer, 2020). Previous research provides a rich database for factors that make teaching effective, such as high-quality learning tasks, effective feedback, and students' actual time on task. We suggest that similar factors are also critical when it comes to distance learning (Muijs, 2021). Challenges include motivating students and providing feedback while schools are closed. Advanced digitalization of schools seems to be a necessary precondition for effective distance teaching and learning. School management may facilitate good teaching by providing clear communication, reliable digital infrastructure, and resources that support students' learning regardless of their situation at home (Andrew et al., 2020).

Conclusions

The present cross-sectional study indicates a minor drop in the mean competence scores of incoming fifth-graders after two months of closed schools in early 2020. These results complement previous findings for intra-individual competence development, which showed a stronger attenuation of competence levels during the first wave of the pandemic. The present population data suggest that the overall efforts of teachers, students, and parents throughout the pandemic-stricken school year might have mitigated the expectable detrimental effects of school closing on student achievement to some degree. The findings also provide a unique benchmark for future studies of educational disruption and school lockdowns during further waves of the COVID-19 pandemic.

References

Andrew, A., Cattan, S., Costa-Dias, M., Farquharson, C., Kraftman, L., Krutikova, S., Phimister, A., & Sevilla, A. (2020). Learning during the lockdown: Real-time data on children's experiences during home learning. *IFS Briefing Note BN288*. Institute for Fiscal Studies. Retrieved from https://www.ifs.org.uk/publications/14848

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness, 1*, 289–328. https://doi.org/10.1080/19345740802400072

Clark, A. E., Nong, H., Zhu, H., & Zhu, R. (2021). Compensating for academic loss: Online learning and student performance during the COVID-19 pandemic. *China Economic Review, 68*, 101629. https://doi.org/10.1016/j.chieco.2021.101629

de Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Press.

Depping, D., Lücken, M., Musekamp, F., & Thonke, F. (2021). Kompetenzstände Hamburger Schüler*innen vor und während der Corona-Pandemie [Alternative pupils' competence measurement in Hamburg during the Corona pandemic]. *DDS – Die Deutsche Schule*, Beiheft 17, 51–79. https://doi.org/10.31244/9783830993315.03

Di Pietro, G., Biagi, F., Costa P., Karpiński Z., & Mazza, J. (2020). *The likely impact of COVID-19 on education: Reflections based on the existing literature and recent international datasets*. Publications Office of the European Union. https://doi.org/10.2760/126686

Ditton, H., & Krüsken, J. (2009). Denn wer hat, dem wird gegeben werden? Eine Längsschnittstudie zur Entwicklung schulischer Leistungen und den Effekten der sozialen Herkunft in der Grundschulzeit [To those who have, will more be given? A longitudinal study concerning the development of school achievement and the effects of social background during primary school]. *Journal for Educational Research Online, 1*(1), 33–61.

Engzell, P., Frey, A. & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences, 118*(17) e2022376118 https://doi.org/10.1073/pnas.2022376118

Fickermann, D., & Edelstein, B. (2020) "Langsam vermisse ich die Schule ...": Schule während und nach der Corona-Pandemie ["I'm starting to miss school ...": Schooling during and after the Corona pandemic.] *Die Deutsche Schule, Beiheft 16*, 9–33. https://doi.org/10.31244/9783830992318.01

Fischer, U., Merz, G., & Wagner, S. (2017). Kompetenzorientierung im Leseverstehensunterricht: Verknüpfung von Diagnose und Förderung in Lernstand 5. *leseforum.ch* (3), 1–22.

Fuchs, G., & Brunner, M. (2017). Wie gut können bildungsstandardbasierte Tests den schulischen Erfolg von Grundschulkindern vorhersagen? [How well can standard-based tests predict the success of primary-school students?] *Zeitschrift für Pädagogische Psychologie, 31*, 27–39. https://doi.org/10.1024/1010-0652/a000195

Grewenig, E., Lergetporer, P., Werner, K., Woessmann, L., & Zierow, L. (2021). COVID-19 and educational inequality: How school closures affect low-and high-achieving students. *European Economic Review, 140*, 103920. https://doi.org/10.1016/j.euroecorev.2021.103920

Hammerstein, S., König, C., Dreisoerner, T., & Frey, A. (2021). Effects of COVID-19-related school closures on student achievement—A systematic review. *Frontiers in Psychology, 12*, 4020. https://doi.org/10.3389/fpsyg.2021.746289

Haug, N., Geyrhofer, L., Londei, A., Dervic, E., Desvars-Larrive, A., Loreto, V., Pinior, B., Thurner, S., & Klimek, P. (2020). Ranking the effectiveness of worldwide COVID-19 government interventions. *Nature Human Behaviour, 4*, 1303–1312. https://doi.org/10.1038/s41562-020-01009-0

Huber, S., G. & Helm, C. (2020). COVID-19 and schooling: Evaluation, assessment and accountability in times of crises—reacting quickly to explore key issues for policy,

practice and research with the school barometer. *Educational Assessment, Evaluation and Accountability, 32*, 237–270. https://doi.org/10.1007/s11092-020-09322-y

Immerfall, S. (2020). Schule in der Pandemie: Erfahrungen aus Ostwürttemberg. *Aus Politik und Zeitgeschichte*, (51). Retrieved from https://www.bpb.de/apuz/schule-2020/322686/schule-in-der-pandemie-erfahrungen-aus-ostwuerttemberg

Kempert, S., Edele, A., Rauch, D., Wolf, K. M., Paetsch, J., Darsow, A., Maluch, J., & Stanat, P. (2016). Die Rolle der Sprache für zuwanderungsbezogene Ungleichheiten im Bildungserfolg. In: C. Diehl, D. Hunkler, & C. Kristen (Eds.): *Ethnische Ungleichheiten im Bildungsverlauf* (pp. 157–241). Springer VS. https://doi.org/10.1007/978-3-658-04322-3_5

Kolen M. J., & Brennan R. L. (2014). *Test equating, scaling, and linking: Methods and practices (3rd ed.).* Springer. https://doi.org/10.1007/978-1-4939-0317-7

Krüsken, J. (2007). Entwicklung von Schülerleistungen und Zensuren in der Grundschule. In: H. Ditton (Ed.), *Kompetenzaufbau und Laufbahnen im Schulsystem* (pp. 41–62). Waxmann.

Kuhfeld, M., Ruzek, E., Johnson, A., Tarasawa, B., & Lewis, K. (2020a). *Technical appendix for: Learning during COVID-19: Initial findings on students' reading and math achievement and growth.* NWEA. https://www.nwea.org/research/publication/technical-appendix-for-learning-during-covid-19-initial-findings-on-students-reading-and-math-achievement-and-growth/

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020b). Projecting the potential impacts of COVID-19 school closures on academic achievement. *Educational Researcher, 49*(8), 549–565. https://doi.org/10.3102/0013189X20965918

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology, 4*(863), 1–12. https://doi.org/10.3389/fpsyg.2013.00863

Maldonado, J. E., & De Witte, K. (2021). The effect of school closures on standardised student test outcomes. *British Educational Research Journal*. https://doi.org/10.1002/berj.3754

Meeter, M. (2021). Primary school mathematics during the COVID-19 pandemic: No evidence of learning gaps in adaptive practicing results. *Trends in Neuroscience and Education, 25*, 100163. https://doi.org/10.1016/j.tine.2021.100163

Muijs, D. (2021). *What's working well in remote education*. Office for Standards in Education, Children's Services and Skills. Retrieved from https://www.gov.uk/government/publications/whats-working-well-in-remote-education

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from https://www.r-project.org

Revelle, W. (2020). *psych: Procedures for psychological, psychometric, and personality research*. R package version 2.0.12, https://CRAN.R-project.org/package=psych

Robitzsch, A., Dörfler, T., Pfost, M., & Artelt, C. (2011). Die Bedeutung der Itemauswahl und der Modellwahl für die längsschnittliche Erfassung von Kompetenzen: Lesekompetenzentwicklung in der Primarstufe [Relevance of item selection and model selection for assessing the development of competencies: The development in reading competence in primary school students]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 43*, 213–227. https://doi.org/10.1026/0049-8637/a000052

Robitzsch, A., Kiefer, T., & Wu, M. (2020). *TAM: Test Analysis Modules*. R package version 3.5-19, https://CRAN.R-project.org/package=TAM.

Rudolph, J., Schoreit, E., & Lipowsky, F. (2016). Interessen- und Leistungsentwicklung im Mathematikunterricht des vierten Schuljahres. *Psychologie in Erziehung und Unterricht, 63*, 151–166. https://doi.org/10.2378/peu2016.art13d

Schult, J., & Lindner, M. A. (2019). Zur Messgüte von geschlossenen und offenen Antwortformaten in Lernstandserhebungen [Psychometric properties of multiple-choice and constructed response formats in proficiency tests]. *Psychologie in Erziehung und Unterricht, 66*(4), 260–272. https://doi.org/10.2378/peu2018.art31d

Schult, J., & Wagner, S. (2020). *VERA 8 in Baden-Württemberg 2020* (Beiträge zur Bildungsberichterstattung). Institut für Bildungsanalysen Baden-Württemberg. Retrieved from https://ibbw.kultus-bw.de/site/pbs-bw-km-root/get/documents_E1653873552/KULTUS.Dachmandant/KULTUS/Dienststellen/ibbw/Systemanalysen/Bildungsberichterstattung/Ergebnisberichte/VERA_8/Ergebnisse_VERA8_2020.pdf

Schulz, A., Leuders, T., & Rangel, U. (2017). Arithmetische Basiskompetenzen am Übergang zu Klasse 5 – eine empirie- und modellgestützte Diagnostik als Grundlage für spezifische Förderentscheidungen. In A. Fritz, S. Schmidt, & G. Ricken (Eds.), *Handbuch Rechenschwäche* (3rd ed., pp. 396–417). Beltz.

Sosu, E., & Klein, M. (2021, January 14). *Socioeconomic disparities in school absenteeism after the first wave of COVID-19 school closures in Scotland*. https://doi.org/10.31235/osf.io/f4jsy

Spitzer, M. W. H., & Musslick, S. (2021). Academic performance of K-12 students in an online-learning environment for mathematics increased during the shutdown of schools in wake of the COVID-19 pandemic. *PloS ONE, 16*(8), e0255629. https://doi.org/10.1371/journal.pone.0255629

Stanat, P., Schipolowski, S., Rjosk, C., Weirich, S., & Haag, N. (Eds.) (2017*). IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich*. Waxmann.

United Nations (2020). Policy brief: The impact of COVID-19 on children. Retrieved from https://unsdg.un.org/sites/default/files/2020-04/160420_Covid_Children_Policy_Brief.pdf

Voss, T., & Wittwer, J. (2020). Unterricht in Zeiten von Corona: Ein Blick auf die Herausforderungen aus der Sicht von Unterrichts- und Instruktionsforschung [Teaching in times of corona: A look at the challenges from the perspective of research on learning and instruction]. *Unterrichtswissenschaft, 48*, 601–627. https://doi.org/10.1007/s42010-020-00088-2

Weirich, S., Hecht, M., & Becker, B. (2021). *eatRep: Educational assessment tools for replication methods. R package version 0.13.4*, https://CRAN.R-project.org/package=eatRep.

Wendt, H., Kasper, D., Bos, W., Vennemann, M., & Goy, M. (2017) Wie viele Punkte auf der TIMSS-Metrik entsprechen einem Lernjahr? In: T. Eckert & B. Gniewosz (Eds.), *Bildungsgerechtigkeit* (pp. 121–152). Springer VS. https://doi.org/10.1007/978-3-658-15003-7_8

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software, 40*(1), 1–29. Retrieved from http://www.jstatsoft.org/v40/i01/

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer. https://doi.org/10.1007/978-3-319-24277-4

Wickham, H. (2020). *tidyr: Tidy messy data. R package version 1.1.2*, https://CRAN.R-project.org/package=tidyr.

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation, 31*, 114–128. https://doi.org/10.1016/j.stueduc.2005.05.005

Zierer, K. (2021). Effects of pandemic-related school closures on pupils' performance and learning in selected countries: A rapid review. *Education Sciences, 11*(6), 252. https://doi.org/10.3390/educsci11060252

Table 1.

Test Characteristics and Participation Rates for the "Lernstand 5" Assessment Across Years.

Year	2015	2016	2017	2018	2019	2020
Number of items in the reading test	32	34	38	37	35	35
Number of items in the operations test	14	14	14	14	14	14
Number of items in the numbers test	14	14	14	14	14	14
Cronbach's alpha for the reading test	.86	.89	.91	.89	.92	.90
Cronbach's alpha for the operations test	.78	.80	.79	.75	.74	.78
Cronbach's alpha for the numbers test	.78	.77	.73	.72	.75	.72
Participation rate in reading (in %)	96.7	96.4	96.1	96.6	96.3	94.9
Participation rate in mathematics (in %)	96.5	96.5	96.1	96.6	96.2	94.6

Note: Participation rates denote the ratio of participating fifth-graders to the total number of fifth-graders reported by schools each year.

Table 2.

Descriptive Statistics for the Competence Assessment Across Years.

		Year							
Subject / Domain		2015	2016	2017	2018	2019	2020	2017–2019	*d*
German	*n*	84894	83158	82958	84385	82519	81810	249862	
Reading	*M*	500	492	516	498	481	490	498	–0.07
	SD	100	111	118	102	134	115	119	
Mathematics	*n*	84746	83290	83085	84404	82620	81550	250109	
Operations	*M*	500	487	494	498	493	485	495	–0.09
	SD	100	116	112	95	91	112	100	
Numbers	*M*	500	511	515	511	511	509	512	–0.03
	SD	100	101	88	86	95	92	90	

Note: d = standardized difference between competence scores in the year 2020 and in the three previous years (Cohen's *d*).

Table 3.

Selected Percentiles for the Competence Distributions Across Years in Reading and Mathematics (see Figure 1).

		Year							
Domain	Percentile	2015	2016	2017	2018	2019	2020	2017–2019	Δ
Reading	5 %	334	300	333	328	256	300	299	+1.3
	25 %	430	419	432	434	387	412	421	–9.1
	50 %	510	493	519	500	494	491	504	–12.4
	75 %	560	571	600	567	577	570	579	–8.9
	95 %	660	665	709	667	690	672	688	–16.6
Operations	5 %	328	303	310	341	342	301	329	–28.2
	25 %	436	408	419	441	432	408	431	–23.2
	50 %	500	476	503	495	493	476	496	–20.0
	75 %	576	567	566	566	556	562	563	–1.3
	95 %	661	676	675	655	639	667	656	+11.4
Numbers	5 %	336	338	367	367	351	357	361	–4.7
	25 %	435	441	458	452	442	445	452	–7.1
	50 %	493	513	519	510	510	509	512	–3.4
	75 %	574	581	574	571	572	572	572	–0.2
	95 %	664	672	658	655	662	659	659	–0.1

Note: Δ denotes the difference between the percentile in 2020 and the corresponding percentile in 2017 to 2019.

Table 4.

Descriptive Statistics and Correlations for School Competence Scores, School Achievement Change Scores δ_i (Pandemic Year 2020 Compared to the Three Previous Years) and School Characteristics.

			r									
Variable	*M*	*SD*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Reading Comprehension (2017–2019)	502	68										
(2) Operations (2017–2019)	498	52	.98*									
(3) Numbers (2017–2019)	515	44	.98*	.99*								
(4) Reading Comprehension (2020)	494	67	.96*	.95*	.95*							
(5) Operations (2020)	490	62	.95*	.95*	.95*	.95*						
(6) Numbers (2020)	513	47	.94*	.94*	.94*	.94*	.96*					
(7) δ_i Reading Comprehension	–7	19	–.20*	–.15*	–.15*	.09*	–.04	–.04				
(8) δ_i Operations	–9	19	.36*	.31*	.35*	.44*	.59*	.50*	.26*			
(9) δ_i Numbers	–2	14	.01	–.01	–.04	.10*	.15*	.30*	.29*	.50*		
(10) Migration background (in %)	27	19	–.73*	–.71*	–.69*	–.70*	–.68*	–.65*	.14*	–.22*	.04	
(11) Socio-Cultural Capital (1–6)	3.4	0.7	.93*	.91*	.91*	.90*	.89*	.88*	–.13*	.36*	.04	–.71*

Note: * = $p < .01$; k = 1065 schools. Means and standard deviations were weighted by the number of students participating in the 2020 assessment.

Table 5.

Average School Scores Across Domains for Each Quartile of the Proportion of Students with Migration Background and for Each Quartile of the Schools' Mean Socio-Cultural Capital (see Figure 2)

			Migration Background		Socio-Cultural Capital	
Domain	Quartile	*k*	2017–2019	2020	2017–2019	2020
Reading	1	267	549	539	401	398
	2	266	518	512	451	444
	3	266	484	478	507	501
	4	266	416	411	571	561
Operations	1	267	533	530	422	405
	2	266	510	505	459	441
	3	266	485	473	503	494
	4	266	434	417	550	552
Numbers	1	267	544	543	451	450
	2	266	524	522	481	475
	3	266	504	501	518	516
	4	266	461	459	559	559

Note: $k = 1065$ schools. Means were weighted by the number of students participating in the 2020 assessment.

Figure 1.

Five Selected Percentiles for Each Tested Competence Domain Across Cohorts in the Years before the Pandemic (2015–2019) and After the First Wave of the Pandemic (2020)

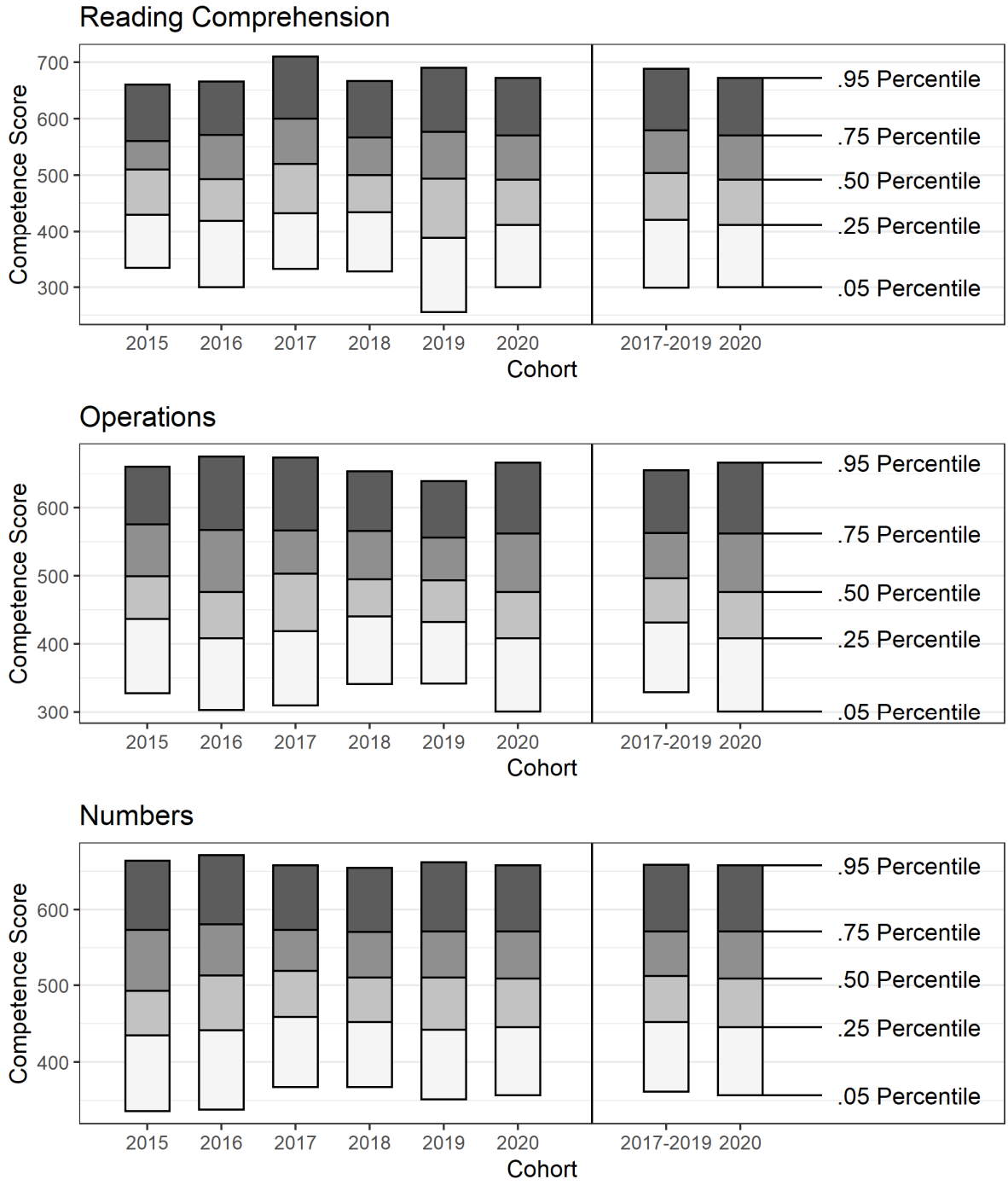
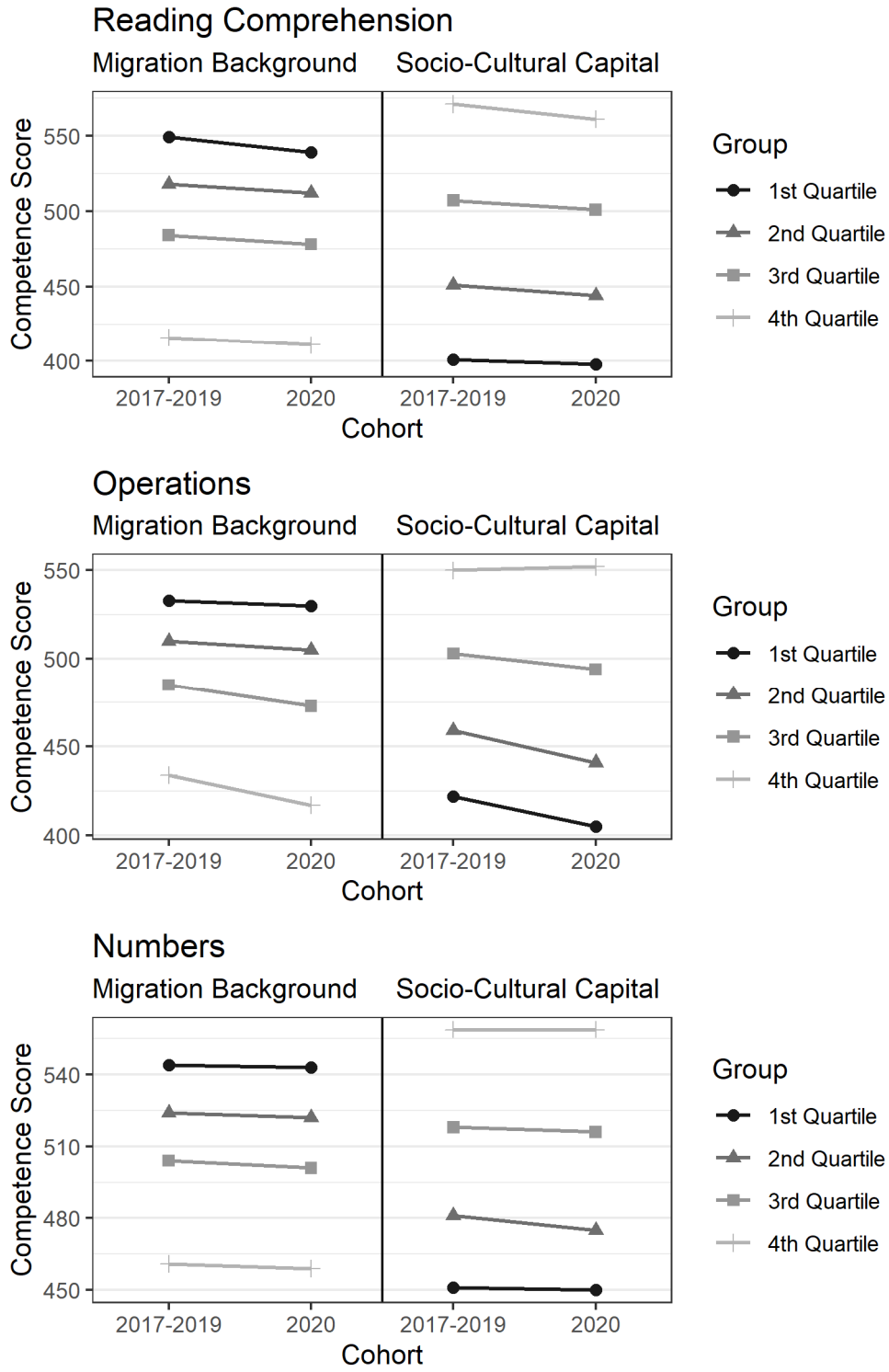


Figure 2.

Average School Scores Across Domains for Each Quartile of the Proportion of Students With Migration Background (Left) and for Each Quartile of the Schools' Average Socio-Cultural Capital (i.e., Number of Books at Home; Right).



Note. The average competence scores were weighted by the number of students participating in the 2020 assessment.