Effects of pandemic-related school closures on pupils' performance and learning in selected countries: a rapid review

Klaus Zierer

Review

Effects of Pandemic-Related School Closures on Pupils' Performance and Learning in Selected Countries: A Rapid Review

Klaus Zierer

Department of School Education, University of Augsburg, 86159 Augsburg, Germany; klaus.zierer@phil.uni-augsburg.de

Abstract: The coronavirus pandemic has led to drastic measures around the world, which have also affected the education system. Schools were closed in the spring of 2020 in almost every country in the world, and many children and young people are still involved in distance learning to this day. What effect these measures have on children's and young people's learning performance is important in view of the time pressure under which educational policy decisions have to be taken. The rapid review presented in this paper delivers evidence on the effects of school closures to contain the coronavirus pandemic in the spring of 2020 on children's and young people's learning performance in five countries (USA, Belgium, the Netherlands, Switzerland and Germany).

Keywords: rapid review; corona pandemic; school closures

1. Introduction

The coronavirus pandemic has led to drastic measures around the world, which have also affected the education system. Schools were closed in the spring of 2020 in almost every country in the world, and many children and young people are still involved in distance learning to this day. What effect these measures have on children's and young people's learning performance is interesting from an educational perspective as there has never been such a social experiment in human history. The question of the effects of school closures on children's and young people's learning performance is also important in terms of educational policy, current decisions must be based on these effects to be able to meet the state-anchored educational mandate comprehensibly.

As of March 2020, more than 160 states have closed schools to contain the spread of the coronavirus, according to UNESCO. In this respect, one can speak of a global school closure. Since then, there have been differences in the ways schools were reopened. For example, there are countries where the duration of school closures exceeds 40 weeks (e.g., Brazil, Argentina and India), while other countries, such as Germany and Italy, have had to close schools only for around 30 weeks. In a few countries, it was only around 10 weeks to date, as in France and Switzerland [1]. However, the situation is dynamic, and differentiation must be made with regard to the age of the learners and the concept of teaching (e.g., alternating lessons in Germany where half of the class learns at school and the other half of the class learns at home). There is no doubt, however, that school closures represent an important political tool for containing the coronavirus pandemic and one that requires educational investigation.

In view of the time pressure under which these educational policy decisions have to be taken, a particularly efficient method for generating evidence is necessary. The aspiration is as follows: to be quick in collecting and evaluating data and to meet all standards of scientific work. For this task, a procedure was established in many disciplines during the coronavirus pandemic, which will also be used in this article. The rapid review is a "form of knowledge synthesis that accelerates the process of conducting a traditional

systematic review through streamlining or omitting various methods to produce evidence for stakeholders in a resource-efficient manner." [2] (p. 15) Compared to a systematic review, several decisions are made to reduce the time it takes to identify, summarize and answer the questions. In order to show trends, meta-analytical procedures are taken up and adapted. Although the concept of a rapid review is not new, from a methodological point of view, it remains a procedure that must be further differentiated [3]. The speed at which educational policy decisions are to be made in certain situations underlines the necessity, the characteristics and the use of a rapid review.

Thus, the aim of this paper is to use a rapid review to obtain empirical data on the effect of school closures on children's and young people's learning performance. Since the school closures occurred globally in March 2020, this period is of particular interest.

2. Literature Search

Studies should be identified which examined the consequences of the school closures in the spring of 2020 to contain the coronavirus pandemic in terms of children's and young people's learning performance. To this end, a literature search was conducted on ResearchGate and Google Scholar until 10 March 2021. Even though both platforms do not have the reputation of, for example, Web of Science and Scopus, they are suitable for a rapid review mainly because they allow a quick overview of worldwide research, provided that it has already been published in journals or else as gray literature. Given the short time span between the school closures in spring 2020 and the present rapid review, this is certainly an advantage in the choice of platforms for the literature research, but it also entails the disadvantage of a publication bias and a possible lack of research quality, which imposes limitations on the interpretation of the data and will be explicitly addressed below. The search was based on the terms "learning performance", "COVID" and "school closure". To obtain an initial narrowing, only reports since 2020 were considered. In total, this yielded 75 literature hits. The vast majority of these articles had not collected their own empirical data or relied on data not collected during the coronavirus pandemic. They were therefore excluded. Thus, according to the screening of titles/abstracts and, if necessary, full-text analysis, only studies that pursued the above-mentioned question and were published in English or German were included. In addition, the bibliographies of the papers selected by screening of titles/abstracts were viewed. After a subsequent full-text analysis, seven studies were included. They come from the USA, Belgium, the Netherlands, Switzerland and Germany, focus on learners in primary school and secondary school and examine learning performance in mathematics, native language or several subjects combined. Following this system, the studies are described in further detail below.

3. Effects on Mathematics Performance in the Primary School

Kuhfeld et al. [4], in their study, found considerable effects of school closures on mathematics performance, among other things. Grades three to five were particularly affected [4] (p. 8). All samples from grades three to eight included more than 300,000 test persons from different school districts in the USA [4] (p. 9). Since primary schools in the USA have a different number of grades depending on the school district, only the results for third and fourth graders are presented here. Across all ethnic groups tested, lower percentile values were found in the test in the autumn of 2020 compared to a group of pupils of the same age who were tested in the autumn of 2019. In the third grade, mathematics performance decreased by 9 percentile positions; in the fourth grade, it decreased by 9 percentile positions [4] (p. 13). The authors point out that pupils from ethnic minorities were largely not tested in the autumn of 2020, which could lead to the extent of the differences reported here being underestimated [4] (p. 5).

Moldonado and De Witte [5] investigated the effect of the seven-week-long school closures on primary school pupils of Flemish schools in Belgium and determined a decrease in the test values achieved in mathematics by 0.19 standard deviations [5] (p. 1). The heterogeneity of performance increased both within the schools examined and between

the institutions. It is noticeable that the performance deficits were greater for pupils who were already disadvantaged [5] (p. 1). As part of the study, data that were obtained using standardized tests that were carried out in the sixth grade, the last level of primary school, were evaluated [5] (p. 6). The comparison between the years 2019 and 2020 appears particularly relevant, especially since the exact same test was used in both years (p. 10). When comparing the mathematics test values at the aforementioned measurement times without taking control variables into account, a significantly negative effect of −0.190 ($p < 0.05$) was found with a random sample size of $n = 1287$ schools [5] (p. 12). The effect found was robust to the introduction of control variables into the model. Even if the sample was restricted to schools that participated in each year of the test (from 2015 to 2020 in total), the result remained similar [5] (p. 24). The proportion of schools that took part in the test in 2020 was around half the participation rate in previous years [5] (p. 32). It should be pointed out that this fact could have an influence on the results presented here.

Engzell et al. [6] dealt with the effects of the eight-week school closures in the Netherlands in their study, using a sample of around $n = 350,000$ primary school pupils (15% of Dutch primary schools). The data were based on the results of six-monthly tests of the core subjects [6] (p. 1). Overall, a learning loss of 3 percentile points or 0.08 standard deviations was determined, whereby the losses in the case of pupils from less educated parents turned out to be 60% greater [6] (p. 1). Standardized mathematics tests developed by the Dutch National Institute for Educational Measurement were carried out in grades four to seven [6] (pp. 5, 50). With a sample size of $n = 286,515$, a significant difference in mean values was found when comparing the control group and the test group. The group of pupils that was tested after the school closures showed significantly worse test values than a comparison group that did not experience any school closures [6] (p. 85). The reported mean difference corresponds to an effect size of Cohen's d = 0.142, with a 95% confidence interval of −0.15−−0.135 (own calculation according to Lenhard, W. & Lenhard, A., 2016). In order to be able to classify the results, Engzell et al. [6] found that the time-measured learning loss in relation to the progress that would be achieved in a normal school year corresponds to a period of 7.9 to 10.5 weeks of non-learning [6] (p. 10). The authors also noted that access to technology alone is not enough to guarantee quality distance learning [6] (p. 11).

Meeter [7] presented the effects of school closures on school learning based on data from the Netherlands. These were collected during the school closures with the help of adaptive exercise software (Snappet). The sample consisted of 53,656 primary school pupils from grades two to six and included 810 schools. Only schools that used the learning software Snappet were included in the analyses. For the school year 2018/19 the data from 48,815 pupils were available; for the school year 2019/20, there were 53,656 pupils [7] (p. 4). As part of the study, Meeter differentiated between effects on learning, attainment and practice; in addition, the pupils were structured into three groups of equal size according to their performance in the first half of the year [7] (p. 5). The reported effects of the school closures were all positive and ranged between Cohen's d ~0.045 and Cohen's d ~0.85. More precise information was not possible because the author only presented effect sizes and confidence intervals graphically [7] (p. 7). The greatest positive effects can be seen in the lower grades, especially in the area of learning [7] (p. 6). It is precisely in these grades that the reopening of schools shows the greatest negative effects in the area of learning. Positive effects between Cohen's d ~0.045 and Cohen's d ~0.18 are to be noted for all groups of pupils also with regard to attainment, in the time of closed schools [7] (p.6). After the reopening, these turned out to be smaller, but still positive. In the area of practice, greater effects were demonstrated during school closures than before this measure was taken. After the institutes were reopened, these effects were even greater [7] (p. 7). In the case of weaker pupils from lower grades, positive effects of the school closures up to the end of the 2020 school year were demonstrated [7] (p. 8). Although some of the positive effects found were equalized after the reopening of the schools, an increase of 0.2 standard deviations for the school year 2019/20 can be noted for all domains recorded with a sample with

n = 25,336 observations in the fourth grade. Meeter [7] pointed out that for reasons of data protection, multi-level analyses were not possible, which in turn increases the probability of false-positive results [7] (p. 10). The software Snappet is also both a digital learning environment and the measurement tool used here [7] (p. 3ff.). It seems misguided to assume that the use of this fully developed learning program can be equated with average distance learning. This fact should be kept in mind when interpreting the study results.

Depping et al. [8] dealt with the skill levels of Hamburg pupils before and during the coronavirus pandemic in their study. In doing so, they compared the skill levels of comparable pupil cohorts in the areas of competence of German reading comprehension and mathematics at times before and during the coronavirus pandemic [8] (p. 51). Data collected using the LERNSTAND and KERMIT instruments in grades four and five in 2020 and grades three and five in 2019 were evaluated [8] (2021, p. 51). Since the regularly planned survey with KERMIT was not possible in the spring of 2020, the LERNSTAND instruments for the fourth and fifth grades and FLIP for the third and seventh grades were used as alternatives [8] (p. 51 et seqq.). At this point, only the results of the analyses in the primary school area are of interest. In the context of the study, the pupils were divided into three exposure groups, depending on the school's social index [8] (p. 59). When comparing the scores achieved with KERMIT 3 (*n* = 13,238) and LERNSTAND 4 (*n* = 13,165), the greatest observable effect was shown, with Cohen's d = −0.061, in exposure group 2 [8] (p. 62). Neither a *p*-value nor a confidence interval was given for this value. A comparison of the percentage distribution across skill levels KS I to KS III showed that both cohorts, 2019 and 2020, were roughly the same in mathematics [8] (p. 63). As a result, the differences "must not be interpreted as meaningful in practice and can probably also be traced back to normal fluctuations in performance between cohorts of different age groups" [8] (p. 63). The authors pointed out the fact that in the 2020 survey "significantly more pupils [...] processed too few tasks for the calculation of a competence value" [8] (p. 74) than was the case in 2019 (p. 79). Due to this, the results of the study could be biased [8] (p. 74).

4. Effects on Mathematics Performance in Secondary Education

Schult et al. [9] evaluated the results of large-scale competence tests (LERNSTAND 5) in the areas of reading and mathematics in their study. A sample consisting of over 80,000 fifth grade pupils in the German federal state of Baden Württemberg was used for this purpose. Primary schools there were closed for eight weeks [9] (p. 2). Test values from the years 2017 to 2019 (before the school closures, *n* = 250,109) and of 2020 (after the school closures, *n* = 81,550) were used [9] (pp. 4, 29). In the area of competence in mathematics, a distinction was made between operations and numbers. When comparing the averaged test values from 2017 to 2019 with the averaged test values from 2020, a negative effect of Cohen's d = −0.09 with a 95% confidence interval of −0.107−0.092 was seen in the "Operations" area [9] (p. 29). In the "numbers" area there was also a negative effect of Cohen's d = −0.03, with a 95% confidence interval of −0.041 to −0.025 [9] (p. 29). The authors came to the conclusion that, based on an augmented competence of 0.65 standard deviations in a normal school year, the pupils lagged behind in their competence development by one month due to the school closures [9] (p. 16). The results of the investigations also suggest that average and weaker pupils in the 2020 test showed greater deficits in terms of mathematical operations than the cohorts from previous years. This difference could not be demonstrated in the high-performance pupils [9] (p. 17). The effect of school characteristics, such as socio-economic capital and the proportion of pupils with an immigrant background, carried hardly any weight in relation to the changes of the competence values at the schools [9] (p. 17). Schult et al. [9] stated that the dispersion of the competence levels of pupils within the individual schools could pose greater challenges for the teachers than a slight pandemic-related weakening of the average competence level at the school [9] (p. 18).

Also, with regard to mathematics competence, Depping et al. [8] provided results with their study on the skill levels of Hamburg pupils. With the instruments KERMIT 5

(2019) and LERNSTAND 5 (2020), 13,703 pupils were tested in 2019 and 13,938 pupils in 2020 [8] (p. 66). Both samples contained observations from grammar schools and from district schools [8] (p. 65). A comparison of the distribution of pupils across the three competence areas KS I, KS II and KS III in 2019 and 2020 revealed no major deviations [8] (p. 71) This finding is supported by the fact that a comparison of the test values achieved on average did not reveal any statistically significant difference [8] (p. 71) This finding is underpinned by the fact that a comparison of the mean test values does not reveal any statistically significant difference [8] (p. 70), Table 1.

Table 1. Studies on mathematics performance in primary school and secondary education—overview.

School Form	Authors	Title [a]	Year	Country	Measuring Instrument	*n*	DV	Result(s)	Form of Presentation of the Results
					Primary School				
	Kuhfeld et al. [4]	Technical appendix for: Learning during COVID-19	2020	USA	NWEA MAP Growth	>650,000 P.	Math RIT Scores	−9 percentile positions (third grade) −11 percentile positions (fourth grade)	Change in relation to performance quintile, percentage-wise
	Moldonado and De Witte [5]	The effect of school closures	2020	Belgium	Standardized school performance tests	1287 Schools	Mathematics Score	Effect size = −0.19 (SD = 0.087)	Effect size
	Engzell et al. [6]	Learning loss due to school closures	2021	Netherlands	Half-yearly tests in the core subjects	286,515 P.	Mathematics Test Result	Cohen's d = −0.142 [−0.15−−0.135] [b]	Mean values, *p*-values
	Meeter [7]	Primary school mathematics during Covid-19	2021	Netherlands	Snappet (Practice software)	25,336 P.	Learning Attainment Practice	Effect size = 0.2 [c]	Effect size, graphic. Exact values (d, confidence level) not exactly readable
	Depping et al. [8]	Competence levels of Hamburg pupils before and during the coronavirus pandemic	2021	Germany	LERNSTAND 4 and KERMIT 3	13,238 P. (2019) 13,165 P. (2020)	Competence levels mathematics	No significant differences	Comparison of percentage distribution on the skill levels KS I, KS II and KS III+
					Secondary education				
	Schult et al. [9]	Did Students Learn Less During the COVID-19 Pandemic?	2021	Germany	LERNSTAND 5	250,109 P. (2017–2019) 81,550 P. (2020)	Operations Figures	Cohen's d = −0.09 [−0.107−−0.092] [d] Cohen's d = −0.03 [−0.041−−0.025] [e]	Effect size, Cohen's d, mean values, standard deviations
	Kuhfeld et al. [4]	Technical appendix for: Learning during COVID-19	2020	USA	NWEA MAP Growth	>975,000 P.	Math RIT Scores	−11 percentile positions (fifth grade) −4 percentile positions (sixth grade) −4 percentile positions (seventh grade)	Change in relation to Performance quintile, percentage-wise
	Depping et al. [8]	s.a.	2021	Germany	LERNSTAND 5 and KERMIT 5	13,703 P. (2019) 13,938 P. (2020)	Competence levels mathematics	No significant differences	Comparison percentage distribution on the skill levels KS I, KS II and KS III+

Notes. P. = pupils, s.a. = see above, DV = Dependent variable or relationship. [a] The study titles are only given in abbreviated form here, for the complete title, see Bibliography; [b] own calculation based on reported mean values, sample sizes and standard deviations with 95% confidence intervals in square brackets; [c] the effect size reported here was based on the study by Schult et al. (2021); [d] own calculation of the 95% confidence interval in square brackets based on the values given in the study; [e] own calculation of the 95% confidence interval in square brackets based on values given in the study.

5. Effects on School Performances in the Area of Language in Primary School

Kuhfeld et al. [4] also dealt with the effects of school closures on American pupils' reading ability as part of their study. By analogy to the mathematics performance, the reading ability was assessed with the "NWEA MAP Growth" [4] (p. 2). A comparison of percentile values showed that pupils who were tested in the autumn of 2019 achieved similarly high test values in the area of reading than pupils who were tested in the autumn of 2020 [4] (p. 5). The percentage of pupils who showed an increase in performance, no change in performance or a decrease in performance with regard to reading were roughly the same in 2019 and 2020 [4] (p. 7). Black and Hispanic pupils in the upper elementary school grades showed a slightly worse reading performance in the autumn of 2020 than in the autumn of the previous year [4] (p. 5). The increase in reading performance in the period from the winter of 2020 (start of the COVID 19 pandemic) to the autumn of 2020, corresponds to the increase that is to be expected in a "typical" year [4] (p. 6).

Kogan and Lavertu [10,11] analyzed the results of the "Third Grade English Language Arts (ELA)" test, which is performed several times a year in the US state of Ohio [10] (p. 2). While the test scores of 124,816 in the third year were available for the year 2019, data from 124,710 pupils were used for the autumn of 2020 [11] (p. 5). The results showed that the language performance provided by third graders in the period from the autumn of 2019 to

the autumn of 2020 decreased significantly, by about 0.23 standard deviations [10] (p. 2). This is an adjusted value that was calculated by estimating missing scaled values based on data from the second grade [11] (p. 5). The unadjusted effect was −0.198 [11] (p. 5). The decrease in performance corresponds to about a third of the increase in learning that is normally recorded in a year and affected economically disadvantaged and particularly pupils of color [10] (p. 2). The authors pointed out that the school closures were not the only reason for the decline in performance in the language test. Negative effects were also demonstrated for pupils who had hybrid or face-to-face lessons during the study period [10] (p. 3).

In addition to the mathematics performance, Moldonado and De Witte [5] also examined changes in primary school pupil performance in the language area, in this case, Dutch and French, at Flemish schools in Belgium [5] (p. 1). Standardized school performance tests were used to assess performance for both subjects [5] (pp. 6, 13). The comparison of the test values from the year 2019 with those from 2020 showed a significant, negative effect of the school closures, which was estimated at −0.237 (p 0.01), with n = 1480 schools [5] (p. 13). The heterogeneity of performance in the subject "Dutch" increased by 20% within the schools examined and by 18% between the institutions [5] (p. 1). The data also showed that the performance deficits were greater in disadvantaged groups of pupils [5] (p.1). In the course of the comparison of the standardized test values in the "French" department from 2019 and 2020, a significant, negative effect of the dummy variable "COVID 19", which was given as −0.191 (p 0.01), with n = 1325 schools [5] (p. 15), was reported. The effects reported here are those that were calculated without taking control variables into account.

As part of the study by Engzell et al. [6] effects of eight weeks of school closures in the Netherlands relating to reading and spelling are also reported. The available sample was particularly rich at around 350,000 primary school pupils [6] (p. 1). Standardized tests were used to assess pupils' performance in the areas of reading and spelling, which were developed by the Dutch National Institute for Educational Measurement [6] (pp. 5, 50). For the reading area, with a sample size of n = 217,875 (control group) and n = 54,487 (test group), a significant difference in mean values emerged in the course of the comparison between the control group and the test group. The group of pupils that was tested after the school closures showed significantly worse test values (M = −1.01, SD = 18.94) than the comparison group (M = 0.66, SD = 18.96) that experienced no school closures [6] (p. 85). The strength of the effects amounts to Cohen's d = −0.088, at a 95% confidence interval of −0.097–−0.079 (own calculation according to [12]). Also, with regard to the area of spelling, a significantly lower test mean value was reported for the test group (M = 0.71, SD = 17.07, n = 58,627) compared to the control group (M = 0.05, SD = 17.27, n = 284,499) (p. 85). The effect size is Cohen's d = −0.044, at a 95% confidence interval of −0.059–−0.03 (own calculation according to [12]).

Depping et al. [8] ascertained the skill level of primary school pupils in Hamburg and also took German reading comprehension into account in their study. Based on data collected with the instruments LERNSTAND 4 and KERMIT 3, the skill levels of pupils from the school year 2019/20 were compared with those from 2020/21 in order to make possible changes visible [8] (p. 51). By analogy to the evaluation of the change in competence in the subject of mathematics, the results with regard to German reading comprehension were divided into the three skill levels KS I, KS II and KSIII+ (p. 63). Next, the percentage distributions of the skill levels from the year 2019 (KERMIT 3) were compared with those from 2020 (LERNSTAND 4), whereby no significant differences were found [8] (p. 63 et seq.). Just like with the investigation of competence in mathematics, it should also be noted that in some cases no competence value could be calculated due to insufficiently completed tasks [8] (p. 74). This can lead to a distortion of the results.

6. Effects on School Performances in the Area of Language in Secondary Education

Schult et al. [9] used data that were collected with the LERNSTAND 5 competence test in their study. The reading comprehension part of this test tests various reading processes,

such as recalling explicitly stated information, interpreting ideas, evaluating content and text elements, and the ability to draw conclusions [9] (p. 10). The result is an overall test value, the development of which will be examined over the years 2017 to 2020. When comparing the standardized test values from the years 2017 to 2019 (before school closure, *n* = 249,862) with those that were collected in 2020 (after school closure, *n* = 81,810), there is a negative effect of Cohen's d = −0.07 [9] (p. 29). The differences that were found with regard to reading competence were more pronounced in the group of high-performance pupils than in the group of low-performance pupils [9] (p. 14). According to the authors of the study, this difference could be due to the fact that during school closures, more time is spent on basic reading tasks, such as reading e-mails and text in the chat [9] (p. 17).

Depping et al. [8], among other things, collected the skill levels of fifth graders at Hamburg schools in the field of German reading comprehension. The instruments KERMIT 5 (2019) and LERNSTAND 5 (2020) were used. The sample from the year 2019 comprised 13,716 pupils, and from the year 2020, 13,936 observations were available [8] (p. 65). In the course of analyzing the differences between the mean scores obtained with said test, no statistically significant differences were found [8] (p. 67), Table 2.

Table 2. Studies on the performance in the area of language in primary school and secondary education—overview.

School Form	Authors	Title [a]	Year	Country	Measuring Instrument	*n*	DV	Result(s)	Form of Presentation of the Results
					Primary School				
	Kuhfeld et al. [4]	Technical appendix for: Learning during COVID-19	2020	USA	NWEA MAP Growth	>650,000 P.	Reading Scores	−1 percentile positions (third grade) −3 percentile positions (fourth grade)	Percentage changes in percentiles
	Kogan and Lavertu [10,11]	Covid and Student Achievement	2021	USA (Ohio)	Third-Grade English Language Arts (ELA)	249,526 P.	Language performance	Effect size = −0.198 [−0.206–−0.19] [b]	Change in standard deviations
	Moldonado and De Witte [5]	The effect of school closures	2020	Belgium	Standardized School performance s-tests	1480 Schools 1325 Schools	Dutch Score French Score	Effect size = −0.237 (SD = 0.063) Effect size = −0.191 (SD = 0.068)	Effect size *n*
	Engzell et al. [6]	Learning loss due to school closures	2021	Netherlands	Half early tests in the core subjects	217,875 P. 284,499 P.	Reading Spelling	Cohen's d = −0.088, [−0.097–−0.079] [c] Cohen's d = −0.044, [−0.059–−0.03] [d]	Mean values, *p*-values
	Depping et al. [8]	Competence levels of Hamburg pupils before and during the coronavirus pandemic	2021	Germany	LERNSTAND 4 and KERMIT 3	13,238 P.	Deutsch-reading comprehension	No significant differences	Comparison percentage distribution to the skill levels KS I, KSII and KS III+
					Secondary education				
	Schult et al. [9]	Did Students Learn Less During the COVID-19 Pandemic?	2021	Germany	LERNSTAND 5	249,862 P. (2017–2019) 81,810 P. (2020)	German reading comprehension	Cohen's d = −0.07 [−0.077–−0.061] [e]	Effect size, Cohen's d, mean values, standard deviations
	Kuhfeld et al. [4]	Technical appendix for: Learning during COVID-19	2020	USA	NWEA MAP Growth	>975,000 000 P.	Reading Scores	−2 percentile positions (fifth grade) −1 percentile positions (6th grade) −2 percentile positions (seventh grade)	Percentage changes in percentiles
	Depping et al. [8]	s.a.	2021	Germany	LERNSTAND 5 and KERMIT 5	13,716 P. (2019) 13,936 P. (2020)	German-reading comprehension	No significant differences	Comparison of percentage distribution to the skill levels KS I, KS II and KS III+

Notes. P. = pupils, s.a. = see above, DV = Dependent variable or relationship. [a] The study titles are only given in abbreviated form here, for the complete, title see Bibliography; [b] own calculation based on reported mean values, sample sizes and standard deviations with 95% confidence intervals in square brackets; [c] the effect size reported here was based on the study by Schult et al. (2021); [d] own calculation based on values given in the study with 95% confidence interval in square brackets; [e] own calculation of the 95% confidence interval in square brackets based on values given in the study.

7. Effects on the School Performance in Primary School—Several Subjects Combined

Tomasik et al. [13] compared pupils' augmented learning in the eight weeks before and during school closures using a latent growth model based on a sample of a total of *n* = 28,685 Swiss pupils [13] (p. 1). At this point, the results relating to primary education were of interest, where effects could be found on the basis of the data in contrast to secondary education [13] (p. 1). The online computer system MINDSTEPS, which contains exam questions generated by teachers, was used to collect data. It provides values for the areas of mathematics, reading and grammar. The sample included observations from *n* = 13,134 primary school pupils [13] (p. 4). In the course of the analyses, no distinction was made between the areas of mathematics, reading and grammar. The results of the

study show that the augmented learning of the primary school pupils who were taught face-to-face was more than twice as high as the augmented learning of primary school pupils in distance teaching. This difference of $\Delta\chi^2$ (1) = 8.86 is significant ($p < 0.001$) [13] (p. 6). Tomasik et al. [13] point out that particularly inter-individual differences in learning increased sharply during the time of distance learning and that a decrease in the augmented learning can have negative long-term consequences [13] (p. 7).

Engzell et al. [6] reported a composite score as part of their study, in which the test values from the areas of mathematics, reading and writing converge. As previously reported, standardized tests developed by the Dutch National Institute for Educational Measurement were used to measure performance [6] (pp. 5, 50). The underlying sample indicated n = 289,189 primary school pupils in the case of the control group; in the test group, there were n = 69,190 observations [6] (p. 85). The test values of the group of pupils that were tested after the school closures (test group, M = −1.28, SD = 11.83), were on average significantly lower than those of the comparison group (control group, M = 0.28, SD = 10.99) that had not experienced any school closures [6] (p. 85). The effect size calculated from this is Cohen's d = −0.14, at a 95% confidence interval of −0.148 to −0.132 (own calculation according to [12]). Engzell and colleagues quantified the learning loss, taking into account all the differences measured by them, with an overall negative effect of −0.08 standard deviations [6] (p. 1).

Moldonado and De Witte [5] calculated the effect of school closures on the grade point average in mathematics, Dutch, French, natural sciences and social sciences in their study on the effects of the COVID 19 pandemic. When comparing the values from 2019 with those from 2020 after the school closures, a negative effect of −0.228 was demonstrated in the fully saturated model. The underlying sample included 513 Belgian primary schools [5] (p. 16). For the model in which no control variables were taken into account, an effect of −0.225 was reported for a sample of n = 719 schools [5] (p. 16).

8. Effects on the School Performance in Secondary Education—Several Subjects Combined

By analogy to the procedure for evaluating primary schools, Tomasik et al. [13] developed a latent growth model in order to be able to compare the learning progress of Swiss secondary school pupils before and during school closures [13] (p. 1). The data required for this were collected using a system that was originally developed to provide teachers and pupils with formative feedback. However, it can also be used to enhance the capabilities of pupils in the course of time. The augmented learning was collected into the areas of mathematics, reading and grammar; the dataset used comprised observations of 15,551 secondary school pupils [13] (p. 5). In contrast to primary education, in the case of secondary education, no significant differences could be found between the learning progress of pupils in face-to-face and those in distance learning ($\Delta\chi^2$(1) = 1.01, p = 0.31). The increase during face-to-face teaching was ν_{inp} = 0.012 (SEν = 0.005, $p < 0.05$) and during distance teaching it was ν_{dis} = 0.008 (SEν = 0.004, p = 0.05) [13] (p. 6). The authors of the study point out that their analyses did not compare the best possible face-to-face teaching with the best possible distance teaching. This must be taken into account when interpreting the results [13] (p. 7), Table 3.

Table 3. Studies on several subjects combined in primary school and secondary education—overview.

School Form	Authors	Title [a]	Year	Country	Measuring Instrument	n	DV	Result(s)	Form of Presentation of the Results
					Primary School				
	Tomasik et al. [13]	Educational gains of in-person vs. distance learning in primary and secondary schools	2020	Switzerland	MINDSTEPS	13,134 (P.)	Learning progress in mathematics, reading and grammar combined	ν = 0.042 (SE = 0.007) augmented learning in face-to-face teaching twice as high as in distance teaching	Change of the slopes
	Engzell et al. [6]	Learning loss due to school closures during the COVID-19 pandemic	2021	Netherlands	Half-yearly tests in the core subjects	289,189 P. (Ctrl.) 69,190 P. (Tg.)	Composite score from mathematics, spelling and reading	Cohen's d = −0.14 [−0.148−−0.132] [b]	mean values, *p*-values

Table 3. *Cont.*

School Form	Authors	Title [a]	Year	Country	Measuring Instrument	*n*	DV	Result(s)	Form of Presentation of the Results
	Moldonado and De Witte [5]	The effect of school closures	2020	Belgium	Standardized school performance tests	719 Schools	Grade Point Average (GPA) from mathematics, Dutch, French, natural sciences and social sciences	Effect size non-saturated model = −0.225 (SD = 0.130)	Effect size *n*
					Secondary education				
	Tomasik et al. [13]	s.a.	2020	Switzerland	MINDSTEPS	13,134 (P.)	Learning progress in mathematics, reading and grammar combined	ν = 0.008 (SE = 0.004) No significant difference between distance and face-to-face classes	Change of the slopes

Notes. P. = pupils, Ctrl. = control group, Tg. = test group, s.a. = see above, DV = Dependent Variable or relationship. [a] The study titles indicated are abbreviated, for the complete title, see Bibliography; [b] own calculation based on the values indicated in the study at 95% confidence interval in square brackets.

9. Interpretation of the Results

In order to simplify the evaluation of the study results, an overview is first given of how the changes found are directed. This is intended to give an impression of the fundamental tendency that can be read from the studies presented above (Table 4).

Table 4. Nature and number of the differences found regardless of the relationship.

Direction of the Change	School Form	Authors	Title [a]	DV
		Positive change		
	Primary School			
		Meeter [7]	Primary school mathematics during Covid-19	Mathematics (learning, performance, practice)
	Secondary education	-	-	
		No differences		
	Primary School			
		Kuhfeld et al. [4]	Technical appendix for: Learning during COVID-19	Mathematics performance and reading performance
		Depping et al. [8]	Competence levels of Hamburg pupils before and during the coronavirus pandemic	Competence level mathematics and German reading comprehension
	Secondary education			
		Depping et al. [8]	s.a.	s.a.
		Tomasik et al. [13]	Educational gains of in-person vs. distance learning in primary and secondary schools	Learning progress in mathematics, reading and grammar combined
		Negative change		
	Primary School			
		Kuhfeld et al. [4]	s.a.	s.a.
		Moldonado & De Witte [5]	The effect of school closures	Performance in Dutch, French, social sciences, mathematics, natural sciences and Overall Grade Average
		Engzell et al. [6]	Learning loss due to school closures	Performance in mathematics, spelling, reading individually and overall value
		Kogan & Lavertu [10,11]	Covid and Student Achievement	Language performance
	Secondary education			
		Schult et al. [9]	Did Students Learn Less During the COVID-19 Pandemic?	Competence level in mathematics, divided into operations and numbers and German reading comprehension

Notes. DV = Dependent Variable or relationship. [a] The study titles indicated are abbreviated here, for the complete title, see Bibliography.

Positive effects were found during the school closures in only one of the available studies. These were in the area of mathematics [7]. In contrast, no significant differences were reported in three studies [4,8,13]. Tomasik et al. [13] found insignificant effects in secondary education. One of the studies reported insignificant differences for both primary and secondary schools [8]. A total of five studies on the subject show negative effects of school closures [4–6,9–11].

Schult et al. [9], referring to the studies by Depping et al. [8], Moldonado and De Witte [5], Meeter [7], Engzell et al. [6] and Kuhfeld et al. [4] stated that the data across all these studies and domains indicate an average learning loss of around 0.1 standard deviations during the first wave of the COVID 19 pandemic in 2020 [9] (p. 7). If, in addition to this, one also takes into account the results that Schult et al. [9] and Kogan and Lavertu [10,11] come to, there is not expected to be any change in this estimated value.

10. General Effect of School Closures in the Analyzed Countries

In order to estimate how the pandemic-related school closures in the spring of 2020 affected pupils' learning performance, an attempt was made to make a stable estimate of the relationship with the help of the quantitative data available to date [14,15].

A general statement about the effects of pandemic-related school closures in the five countries (USA, Belgium, the Netherlands, Switzerland and Germany) can be made by calculating the measure of the effect size *d* [16]. Cohen's *d* was initially determined for each of the primary studies for which this was possible on the basis of the data given. The primary study effects were then combined into an averaged effect size, with study effects with higher precision being weighted higher. The precision was determined on the basis of the respective sample size of the primary studies, under the assumption that the unsystematic error variance of a primary study can be determined by the variance of the effect size [17]. Differential effects were examined on the basis of two moderator variables—the learner's age group and the school subject.

The significance of the averaged effect size depends on the quality of the dataset. For example, it must be decided up to which order of magnitude the heterogeneity between the studies is negligible and then the calculation of an averaged effect size makes sense. If the individual study effects show great heterogeneity, an averaged effect size can only be meaningfully interpreted to a limited extent. Furthermore, the averaged effect size is usually calculated from aggregated data. Subgroup analyses and the consideration of potential confounding variables can often only take place to a limited extent or not at all. When integrating published data, there is always the problem of publication bias, which manifests itself in the fact that significant effects are published more often than non-significant ones [14,15]. The informative value of averaged effects also depends on their homogeneity. The Q test can be used to determine whether there is significant heterogeneity between primary study effects [18]. To quantify the extent of the heterogeneity, the test variable I^2 is used [19].

If an attempt is made to calculate a mean effect size hereinafter, then this is with the intention of being able to make an initial general statement on the effects of school closures on learning performance against the background of international studies and thus beyond country-specific peculiarities. More precise and differentiated analyses in the form of a comprehensive meta-analysis are necessary following this rapid review.

The mean effect size, which was calculated on the basis of 12 effects out of seven described primary studies that contained information on the calculation of *d*, is –0.14 standard deviations (95% CI: –0.19, –0.10). As a result of the pandemic-related school closure in spring 2020, pupils lost between 23% (with 0.60 augmented learning per school year) and 35% (with 0.40 augmented learning per school year) in learning time, depending on the ratio. The high heterogeneity of the primary study effects (Q = 594.798; $p < 0.001$) suggests that the impairment of learning performance greatly depends on moderating variables.

If the learner's age and the subject are used as moderator variables (Table 5), it can be seen that primary school pupils were more affected than secondary school pupils and that the effects on the subject of mathematics were more serious than those on the teaching of the native language in secondary education.

Table 5. Results of a provisional moderator analysis.

Moderator	k	N	Q	P	I^2	d	CI
Learners' age							
Primary school	8	>2,435,000	347.52	0.001	97.99%	–0.17	–0.25, –0.10
Secondary education	4	>2,450,000	205.13	0.001	98.54%	–0.10	–0.16, –0.03
Subject							
Mathematics	5	>2,290,000	118.38	0.001	96.62%	–0.17	–0.22, –0.11
Native language	6	>2,530,000	77.90	0.001	93.58%	–0.11	–0.15, –0.07
Subject (only primary school)							

Table 5. *Cont.*

Moderator	k	N	Q	P	I²	d	CI
Mathematics	3	>1,065,000	35.61	0.001	94.38%	–0.19	–0.28, –0.11
Native language	4	>1,298,000	69.28	0.001	95.67%	–0.14	–0.22, –0.06
Several subjects	1	>72,000	-	-	-	–0.23	-
Subject (only secondary education)							
Mathematics	2	>1,225,000	11.53	0.001	91.33%	–0.13	–0.20, –0.06
Native language	2	>1,225,000	0.26	0.611	0.0%	–0.06	–0.07, –0.05

NB. k: Number of primary studies; N: Number of the test persons; Q: Testing size for the heterogeneity; I^2: Extent of the heterogeneity; d: Effect size; CI: 95% Confidence interval.

11. Strengths and Limitations

One of the strengths of the present work is the rapid synopsis of evidence on the effects of school closures to contain the coronavirus pandemic in the spring of 2020 on children's and young people's learning performance. This creates an initial evidence basis for educational policy decisions in the analyzed countries. It should be noted as a limitation that in the course of the method of a rapid review only two databases are used as a source. Therefore, a publication bias cannot be ruled out, and the conclusions should be evaluated with caution. It is likely that further relevant work was not taken into account. Further literature search, which then also takes advantage of platforms such as Web of Science and Scopus, is indispensable. In addition, the methodological quality of the studies included was only assessed conditionally, so possible biases may arise in the presentation of results. Further studies and, if submitted, a more in-depth meta-analytical procedure, are therefore desirable. Another limitation of the present work is that the data are from five countries (USA, Belgium, the Netherlands, Switzerland and Germany). While these countries seem to be similar in many terms of political and social structure, comparisons should always be made with caution due to cultural specificities. A transfer to other countries, especially those with different political, cultural and social conditions (e.g., Brazil, Argentina and India), is not possible.

12. Conclusions

It can be seen that pupils in the analyzed countries (USA, Belgium, the Netherlands, Switzerland and Germany) of all ages and across disciplines had already fallen behind in learning due to the first school closure in the spring of 2020. The high heterogeneity of the effects, which could not be substantially restricted by the moderator variables used here, does show, however, that how school closures affect learning success greatly depends on individual schools and individual teachers. An essential next step would, therefore, be to investigate in these countries which school organizational, pedagogical and didactic concepts help minimize the loss of learning time due to school closures. In addition, it seems worthwhile to include social and cultural factors in further research, as these may also have effects on the impact of school closures on children's and young people's learning performance.

Educational policy administrators should take this evidence into account when making further decisions. In the countries analyzed, it cannot be ruled out that learning performance will decline even more as a result of further school closures. In addition, the results suggest that children and adolescents from educationally disadvantaged backgrounds are more affected. This raises the question of social justice. Therefore, in terms of education policy, measures need to be taken that can help compensate for these learning deficits. Additional courses for individual support, such as summer schools, are just as conceivable as the further professionalization of distance learning with the help of digital media. No statement can be made about the effects of school closures in other countries, especially since it can be assumed that political, social and cultural factors also have an influence.

References

1. UNESO. Education: From Disruption to Recovery. 2021. Available online: https://en.unesco.org/covid19/educationresponse (accessed on 20 May 2021).
2. Garritty, C.; Gartlehner, G.; Kamel, C.; King, V.J.; Nussbaumer-Streit, B.; Stevens, A.; Hamel, C.; Affengruber, L.; Cochrane Rapid Reviews. Interim Guidance from the Cochrane Rapid Reviews Methods Group. 2020. Available online: https://methods.cochrane.org/rapidreviews/sites/methods.cochrane.org.rapidreviews/files/public/uploads/cochrane_rr_-_guidance-23mar2020-final.pdf (accessed on 20 May 2021).
3. Tricco, A.C.; Antony, J.; Zarin, W.; Strifler, L.; Ghassemi, M.; Ivory, J.; Perrier, L.; Hutton, B.; Moher, D.; Straus, S.E. A scoping review of rapid review methods. *BMC Med.* **2015**, *13*, 224. [CrossRef] [PubMed]
4. Kuhfeld, M.; Ruzek, E.; Johnson, A.; Tarasawa, B.; Lewis, K. Technical appendix for: Learning during COVID-19. Initial Findings on Students' Reading and Math Achievement and Growth. NWEA. 2020. Available online: https://www.nwea.org/content/uploads/2020/04/COVID-technical-appendix_2020.04.08.pdf (accessed on 20 May 2021).
5. Maldonado, J.E.; De Witte, K. The Effect of School Closures on Standardised Student Test Outcomes. Discussion Paper Series, DPS20.17. KU Leuven Department of Economics. 2020. Available online: https://feb.kuleuven.be/research/economics/ces/documents/DPS/2020/dps2017.pdf (accessed on 20 May 2021).
6. Engzell, P.; Frey, A.; Verhagen, M.D. Learning inequality during the COVID-19 pandemic. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, 17. [CrossRef] [PubMed]
7. Meeter, M. Primary school mathematics during Covid-19: No evidence of learning gaps in adaptive practicing results. Available online: https://psyarxiv.com/8un6x/ (accessed on 20 May 2021).
8. Depping, D.; Lücken, M.; Musekamp, F.; Thonke, F. Kompetenzstände Hamburger Schüler*innen vor und während der Corona-Pandemie [Alternative pupils' competence measurement in Hamburg during the Corona pandemic]. *DDS Die Dtsch. Sch. Beiheft* **2021**, *17*, 51–79.
9. Schult, J.; Mahler, N.; Fauth, B.; Lindner, M.A. Did Students Learn Less During the COVID-19 Pandemic? Reading and Mathematics Competencies Before and After the First Pandemic Wave. Available online: https://psyarxiv.com/pqtgf/ (accessed on 20 May 2021).
10. Kogan, V.; Lavertu, S. The COVID-19 Pandemic and Student Achievement on Ohio's Third-Grade English Language Arts Assessment. 2021. Available online: http://glenn.osu.edu/educational-governance/reports/reports-attributes/ODE_ThirdGradeELA_KL_1-27-2021.pdf (accessed on 20 May 2021).
11. Kogan, V.; Lavertu, S. Covid and Student Achievement. Early Evidence from Ohio's Fall Tests. 2021. Available online: https://www.educationnext.org/covid-student-achievement-early-evidence-from-ohios-fall-tests/ (accessed on 20 May 2021).
12. Lenhard, W.; Lenhard, A. Berechnung von Effektstärken. 2016. Available online: https://www.psychometrica.de/effektstaerke.html (accessed on 20 May 2021).
13. Tomasik, M.J.; Helbling, L.A.; Moser, U. Educational gains of in-person vs. distance learning in primary and secondary schools: A natural experiment during the COVID-19 pandemic school closures in Switzerland. *Int. J. Psychol.* **2020**. [CrossRef] [PubMed]
14. Lipsey, M.W.; Wilson, D.B. The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *Am. Psychol.* **1993**, *48*, 1181–1209. [CrossRef] [PubMed]
15. Schmidt, F.L.; Hunter, J.E. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*, 3rd ed.; Sage: Thousand Oaks, CA, USA, 2014.
16. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Erlbaum: Hillsdale, NJ, USA, 1988.
17. Hedges, L.V.; Olkin, I. *Statistical Methods for Meta-Analysis*; Academic Press: New York, NY, USA, 1985.
18. Shadish, W.R.; Haddock, C.K. Combining estimates of effect size. In *The Handbook of Research Synthesis*; Cooper, H., Hedges, L.V., Eds.; Russell Sage Foundation: New York, NY, USA, 1994; pp. 261–281.
19. Deeks, J.J.; Higgins, J.P.; Altman, D.G. Analysing data and undertaking meta-analyses. In *Cochrane Handbook for Systematic Reviews of Interventions*; Higgins, J.P., Green, S., Eds.; John Wiley & Sons Ltd.: London, UK, 2008; pp. 243–296.