A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic

Bastian A. Betthäuser*, [1–3], Anders M. Bach-Mortensen[2], Per Engzell[3–5]

January 18, 2023

*To whom correspondence should be addressed. Email: bastian.betthaeuser@sciencespo.fr [1]Centre for Research on Social Inequalities (CRIS), Sciences Po, Paris, France. [2]Department of Social Policy and Intervention, University of Oxford, Oxford, United Kingdom. [3]Nuffield College, University of Oxford, Oxford, United Kingdom. [4]Social Research Institute, University College London, London, United Kingdom. [5]Swedish Institute for Social Research, Stockholm University, Stockholm, Sweden

Abstract

To what extent has the learning progress of school-aged children slowed down during the COVID-19 pandemic? A growing number of studies address this question, but findings vary depending on context. We conduct a pre-registered systematic review, quality appraisal, and meta-analysis of 42 studies across 15 countries to assess the magnitude of learning deficits during the pandemic. We find a substantial overall learning deficit (Cohen's $d = -0.14$, 95% c.i. $-0.17, -0.10$), which arose early in the pandemic and persists over time. Learning deficits are particularly large among children from low socio-economic backgrounds. They are also larger in math than in reading, and in middle-income countries, relative to high-income countries. There is a lack of evidence on learning progress during the pandemic in low-income countries. Future research should address this evidence gap and avoid the common risks of bias that we identify.

The COVID-19 pandemic has led to one of the largest disruptions to learning in history. To a large extent this is due to school closures, which are estimated to have affected 95 percent of the world's student population.[1] But even when face-to-face teaching resumed, instruction has often been compromised by hybrid teaching, and by children or teachers having to quarantine and miss classes. The effect of limited face-to-face instruction is likely compounded by the pandemic's consequences for children's out-of-school learning environment, as well as their mental and physical health. Lockdowns have restricted children's movement and their ability to play, meet other children, and engage in extra-curricular activities. Children's well-being and family relationships have also suffered due to economic uncertainties and conflicting demands of work, care and learning. These negative consequences can be expected to be most pronounced for children from low socio-economic family backgrounds, exacerbating pre-existing educational inequalities.

It is critical to understand the extent to which learning progress has changed since the onset of the COVID-19 pandemic. We use the term 'learning deficit' to encompass both a delay in expected learning progress, as well as a loss of skills and knowledge already gained. The COVID-19 learning deficit is likely to affect children's life chances through their education and labor market prospects. At the societal level, it can have important implications for growth, prosperity, and social cohesion. As policy-makers across the world are seeking to limit further learning deficits and to devise policies to recover learning deficits that have already been incurred, assessing the current state of learning is crucial. A careful assessment of the COVID-19 learning deficit is also necessary to weigh the true costs and benefits of school closures.

A number of narrative reviews have sought to summarize the emerging research on COVID-19 and learning, mostly focusing on learning progress relatively early in the pandemic.[2–6] Moreover, two reviews harmonized and synthesized existing estimates of learning deficits during the pandemic.[7,8] In line with the narrative reviews, these two reviews find a statistically significant reduction in learning progress during the pandemic. However, this finding is based on a relatively small number of studies (18 and 10 studies respectively). The limited evidence that was available at the time these reviews were conducted also precluded them from meta-analyzing variation in the magnitude of learning deficits over-time and across subjects, different groups of students, or country contexts.

In this paper, we conduct a systematic review and meta-analysis of the evidence on COVID-19 learning deficits two and a half years into the pandemic. Our primary preregistered research question was 'What is the effect of the Covid-19 Pandemic on learning progress amongst school-age children?' and we address this using evidence from studies examining changes in learning outcomes during the pandemic. Our second preregistered research aim was 'to examine whether the effect of the Covid-19 Pandemic on learning differs across different social background groups, age groups, boys and girls, learning areas or subjects, national contexts'.

We contribute to the existing research in two ways. First, we describe and appraise the up-to-date body of evidence and its geographic reach and quality. More specifically, we ask (a) What is the state of the evidence, in terms of the available peer-reviewed research and gray-literature, on learning progress of school-aged children during the COVID-19 pandemic?, (b) Which countries are represented in the available evidence?, and (c) What is the quality of the existing evidence?

Our second contribution is to harmonize, synthesize and meta-analyze the existing evidence, with special attention to variation across different sub-populations and country contexts. Based on the identified studies, we ask (d) To what extent has the learning progress of school-aged children changed since the onset of the pandemic?, (e) How has the magnitude of the learning deficit evolved since the beginning of the pandemic?, (f) To what extent has the pandemic reinforced inequalities between children from different socio-economic backgrounds? (g) Are there differences in the magnitude of learning deficits between subject domains (math and reading) and between age groups (primary and secondary school)?, and (h) To what extent does the does the magnitude of learning deficits vary across national contexts?

Below, we report our answers to each of these questions in turn. The questions correspond to the analysis plan set out in our pre-registered protocol (see `https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021249944`), but we have adjusted the order and wording to aid readability. We had planned to examine gender differences in learning progress during the pandemic, but found there to be insufficient evidence to conduct this sub-group analysis, as the large majority of the identified studies do not provide evidence on learning deficits separately by gender. We also planned to examine how the magnitude of learning deficits differs across groups of students with varying exposures to school closures. This was not possible as the available data on school closures lacks sufficient depth with respect to variation of school closures within countries, across grade levels, and with respect to different modes of instruction, to meaningfully examine this association.

Results

The state of the evidence

Our systematic review identified 42 studies on learning progress during the COVID-19 pandemic that met our inclusion criteria. To be included in our systematic review and meta-analysis, studies had to use a measure of learning that can be standardized (using Cohen's d) and base their estimates on empirical data collected since the onset of the COVID-19 pandemic (rather than making projections based on pre-COVID-19 data). As shown in Fig. 1, the initial literature search resulted in 5,153 hits after removal of duplicates. All studies were double-screened by the first two authors. The formal database search process identified 15 eligible studies. We also hand-searched relevant preprint repositories and policy databases. Further, to ensure that our study selection was as up-to-date as possible, we conducted two full forward and backward citation searches of all included studies on February 15, 2022, and on August 8, 2022. The citation and preprint hand-searches allowed us to identify 27 additional eligible studies, resulting in a total of 42 studies. Most of these studies were published after the initial database search, which illustrates that the body of evidence continues to expand. Most studies provide multiple estimates of COVID-19 learning deficits, separately for math and reading and for different school grades. The number of estimates ($n = 291$) is therefore larger than the number of included studies ($n = 42$).

The geographic reach of evidence is limited

Table 1 shows all included studies and estimates of COVID-19 learning deficits (in brackets), grouped by the 15 countries represented: Australia, Belgium, Brazil, Colombia, Denmark, Germany, Italy, Mexico, the Netherlands, South Africa, Spain, Sweden, Switzerland, the United Kingdom and the United States. About half of the estimates ($n = 149$) are from the United States, 58 are from the United Kingdom, a further 70 are from other European countries, and the remaining 14 estimates are from Australia, Brazil, Colombia, Mexico, and South Africa. As this list shows, there is a strong over-representation of studies from high-income countries, a dearth of studies from middle-income countries, and no studies from low-income countries. This skewed representation should be kept in mind when interpreting our synthesis of the existing evidence on COVID-19 learning deficits.

The quality of evidence is mixed

We assessed the quality of the evidence using an adapted version of the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool.[50] More specifically, we analyzed the risk of bias of each estimate from confounding, sample selection, classification of treatments, missing data, the measurement of outcomes, and the selection of reported results. A.M.B.-M. and B.A.B. performed the risk of bias assessments, which were independently checked by the respective other author. We then assigned each study an overall risk of bias rating (low, moderate, serious, or critical) based on the estimate and domain with the highest risk of bias.

Fig. 2a shows the distribution of all studies of COVID-19 learning deficits according to their risk of bias rating separately for each domain (top six rows), as well as the distribution of studies according to their overall risk of bias rating (bottom row). The overall risk of bias was considered 'low' for 15% of studies, 'moderate' for 30% of studies, 'serious' for 25% of studies, and 'critical' for 30% of studies.

In line with ROBINS-I guidance, we exclude studies rated to be at critical risk of bias ($n = 19$) from all of our analyses and figures, except for Figure 2a, which visualizes the distribution of studies according to their risk of bias.[50] These are thus not part of the 42 studies included in our meta-analysis. Supplementary Table 2 provides an overview of these studies as well as the main potential sources of risk of bias. Moreover, in Supplementary Fig. 3–6, we replicate all our results excluding studies deemed to be at serious risk of bias.

As shown in Fig. 2a, common sources of potential bias were confounding, sample selection, and missing data. Studies rated at risk of confounding typically compared only two time points, without accounting for longer time trends in learning progress. The main causes of selection bias were the use of convenience samples and insufficient consideration of self-selection by schools or students. Several studies found evidence of selection bias, often with students from a low socio-economic background or schools in deprived areas being underrepresented after (as compared to before) the pandemic, but this was not always adjusted for. Some studies also reported a higher amount of missing data post-pandemic, again generally without adjustment, and several studies did not report any information on missing data. For an overview of the risk of bias ratings for each domain of each study see Supplementary Fig. 1 and Supplementary Tables 1–2.

No evidence of publication bias

Publication bias can occur if authors self-censor to conform to theoretical expectations, or if journals favor statistically significant results. To mitigate this concern, we include not only published papers, but also preprints, working papers and policy reports.

Moreover, Fig. 2b tests for publication bias by showing the distribution of z-statistics for the effect size estimates of all identified studies. The dotted line indicates $z = 1.96$ ($p = 0.050$), the conventional threshold for statistical significance. The overlaid curve shows a normal distribution. If there was publication bias, we would expect a spike just above the threshold, and a slump just below it. There is no indication of this. Moreover, we do not find a left-skewed distribution of p-values (see p-curve in Supplementary Fig. 2a), or an association between estimates of learning deficits and their standard errors (see funnel plot in Supplementary Fig. 2b) that would suggest publication bias. Publication bias does thus not appear to be a major concern.

Having assessed the quality of the existing evidence, we now present the substantive results of our meta-analysis, focusing on the magnitude of COVID-19 learning deficits and on the variation in learning deficits over time, across different groups of students, and across different country contexts.

Learning progress slowed substantially during the pandemic

Fig. 3 shows the effect sizes that we extracted from each study (averaged across grades and learning subject) as well as the pooled effect size (red diamond). Effects are expressed in standard deviations, using Cohen's d. Estimates are pooled using inverse variance weights. The pooled effect size across all studies is $d = -0.14$, $t(41) = -7.30$, p two-tailed $= 0.000$, 95% c.i. $-0.17, -0.10$. Under normal circumstances, students generally improve their performance by around 0.4 standard deviations per school year.[51–53] Thus, the overall effect of $d = -0.14$ suggests that students lost out on 0.14/0.4, or about 35%, of a school year's worth of learning. On average, learning progress of school-aged children has slowed substantially during the pandemic.

Learning deficits arose early in the pandemic and persist

One may expect that children were able to recover learning that was lost early in the pandemic, after teachers and families had time to adjust to the new learning conditions and structures for online learning and for recovering early learning deficits were set up. However, existing research on teacher strikes in Belgium[54] and Argentina,[55] shortened school years in Germany,[56] and disruptions to education during World War II[57] suggests that learning deficits are difficult to compensate and tend to persist in the long run.

Fig. 4 plots the magnitude of estimated learning deficits (on the vertical axis) by the date of measurement (on the horizontal axis). The color of the circles reflects the relevant country, the size of the circles indicates the sample size for a given estimate, and the line displays a linear trend. The figure suggests that learning deficits opened up early in the pandemic and have neither closed nor substantially widened since then. We find no evidence that the slope coefficient is different from zero (β months $= -0.00,\ t(41) = -7.30,\ p$ two-tailed $= 0.097$, 95% c.i. $-0.01, 0.00$). This implies that efforts by children, parents, teachers, and policy-makers to adjust to the changed circumstance have been successful in preventing further learning deficits, but so far have been unable to reverse them. As shown in Supplementary Fig. 8, the pattern of persistent learning deficits also emerges within each of the three countries for which we have a relatively large number of estimates at different time points: the United States, the United Kingdom and the Netherlands. However, it is important to note that estimates of learning deficits are based on distinct samples of students. Future research should continue to follow the learning progress of cohorts of students in different countries to reveal how learning deficits of these cohorts have developed and continue to develop since the onset of the pandemic.

Socio-economic inequality in education increased

Existing research on the development of learning gaps during summer vacations,[58,59] disruptions to schooling during the Ebola outbreak in Sierra Leone and Guinea,[60] and the 2005 earthquake in Pakistan,[61] shows that the suspension of face-to-face teaching can increase educational inequality between children from different socio-economic backgrounds. Learning deficits during the COVID-19 pandemic are likely to have been particularly pronounced for children from low socio-economic backgrounds. These children have been more affected by school closures than children from more advantaged backgrounds.[62] Moreover, they are likely to be disadvantaged with respect to their access and ability to use digital learning technology, the quality of their home learning environment, the learning support they receive from teachers and parents, and their ability to study autonomously.[63–65]

Most studies we identify examine changes in socio-economic inequality during the pandemic, attesting to the importance of the issue. Because studies use different measures of socio-economic background (e.g., parental income, parental education, free school meal eligibility, or neighborhood disadvantage), pooling the estimates is not possible. Instead, we code all estimates according to whether they indicate a reduction, no change, or an increase in learning inequality during the pandemic. Fig. 5 displays this information. Estimates that indicate an increase in inequality are shown on the right, those that indicate a decrease on the left, and those that suggest no change in the middle. Squares represent estimates of changes in inequality during the pandemic in reading performance, and circles represent estimates of changes in inequality in math performance. The shading represents when in the pandemic educational inequality was measured, differentiating between the first, second and third year of the pandemic. Estimates are also arranged horizontally by grade level. A large majority of estimates indicate an increase in educational inequality between children from different socio-economic backgrounds. This holds for both math and reading, across primary and secondary education, at each stage of the pandemic, and independently of how socio-economic background is measured.

Learning deficits are larger in math than in reading

Available research on summer learning deficits,[58,66] student absenteeism,[67,68] and extreme weather events,[69] suggests that learning progress in mathematics is more dependent on formal instruction than in reading. This might be due to parents being better equipped to help their children with reading, and children advancing their reading skills (but not their math skills) when reading for enjoyment outside of school. Fig. 6a shows that similarly to earlier disruptions to learning, the estimated learning deficits during the COVID-19 pandemic are larger for math than for reading (mean difference $\delta = -0.07$, $t(41) = -4.02$, p two-tailed = 0.000, 95% c.i. $-0.11, -0.04$). This difference is statistically significant and robust to dropping estimates from individual countries (see Supplementary Fig. 9).

No evidence of variation across grade levels

One may expect learning deficits to be smaller for older than for younger children, as older children may be more autonomous in their learning and better able to cope with a sudden change in their learning environment. However, older students were subject to longer school closures in some countries, such as Denmark,[14] based partly on the assumption that they would be better able to learn from home. This may have offset any advantage that older children would otherwise have had in learning remotely.

Fig. 6b shows the distribution of estimates of learning deficits for students at the primary and secondary level, respectively. Our analysis yields no evidence of variation in learning deficits across grade levels (mean difference $\delta = -0.01$, $t(41) = -0.59$, p two-tailed = 0.556, 95% c.i. $-0.06, 0.03$). Due to the limited number of available estimates of learning deficits, we cannot be certain about whether learning deficits differ between primary and secondary students or not.

Learning deficits are larger in poorer countries

Low and middle-income countries were already struggling with a learning crisis before the pandemic. Despite large expansions of the proportion of children in school, children in low and middle-income countries still perform poorly by international standards, and inequality in learning remains high.[70–72] The pandemic is likely to deepen this learning crisis and to undo past progress. Schools in low- and middle-income countries have not only been closed for longer, but have also had fewer resources to facilitate remote learning.[73,74] Moreover, the economic resources, ICT equipment and ability of children, parents, teachers, and governments to support learning from home are likely to be lower in low- and middle-income countries.[75]

As discussed above, most evidence on COVID-19 learning deficits comes from high-income countries. We found no studies on low-income countries that met our inclusion criteria, and evidence from middle-income countries is limited to Brazil, Colombia, Mexico, and South Africa. Fig. 6c groups the estimates of COVID-19 learning deficits in these four middle-income countries together (on the right) and compares them to estimates from high-income countries (on the left). The learning deficit is appreciably larger in middle-income countries than in high-income countries (mean difference $\delta = -0.29$, $t(41) = -2.78$, p two-tailed = 0.008, 95% c.i. $-0.50, -0.08$). In fact, the three largest estimates of learning deficits in our sample are from middle-income countries (see Fig. 3).[12,22,28]

Discussion

Two years since the COVID-19 pandemic, there is a growing number of studies examining the learning progress of school-age children during the pandemic. This paper first systematically reviews the existing literature on learning progress of school-age children during the pandemic and appraises its geographic reach and quality. Second, it harmonizes, synthesizes and meta-analyzes the existing evidence in order to examine the extent to which learning progress has changed since the onset of the pandemic, and how this varies across different groups of students and across country contexts.

Our meta-analysis suggests that learning progress has slowed substantially during the COVID-19 pandemic. The pooled effect size of $d = -0.14$, implies that students lost out on about 35%, of a normal school year's worth of learning. This confirms initial concerns that substantial learning deficits would arise during the pandemic.[51,78,79] But our results also suggest that fears of an accumulation of learning deficits as the pandemic continues have not materialized.[80,81] On average, learning deficits emerged early in the pandemic and have neither closed nor widened substantially. Future research should continue to follow the learning progress of cohorts of students in different countries to reveal how learning deficits of these cohorts have developed and continue to develop since the onset of the pandemic.

Most studies that we identify find that learning deficits have been largest for children from disadvantaged socio-economic backgrounds. This holds across different time points during the pandemic, countries, grade levels, and learning subjects, and independently of how socio-economic background is measured. This suggests that the pandemic has exacerbated educational inequalities between children from different socio-economic backgrounds, which were already large before the pandemic.[84,85] Policy initiatives to compensate learning deficits need to prioritize support for children from low socio-economic backgrounds in order to allow them to recover the learning they lost during the pandemic. There is a need for future research to assess how the COVID-19 pandemic has affected gender inequality in education. To date, there is very little evidence on this issue. The large majority of the studies that we identify do not empirically examine learning deficits separately by gender.

Comparing estimates of learning deficits across subjects, we find that learning deficits tend to be larger in math than in reading. As noted above, this may be due to the fact that parents and children have been in a better position to compensate school-based learning in reading by reading at home. Accordingly, there are grounds for policy initiatives to prioritize the compensation of learning deficits in math and other science subjects.

A limitation of this study and the existing body of evidence on learning progress during the COVID-19 pandemic is that the existing studies primarily focus on high-income countries, while there is a dearth of evidence from low- and middle-income countries. This is particularly concerning because the small number of existing studies from middle-income countries suggest that learning deficits have been particularly severe in these countries. Learning deficits are likely to be even larger in low-income countries, considering that they already faced a learning crisis before the pandemic, generally implemented longer school closures, and were under-resourced and ill-equipped to facilitate remote learning.[72–76] It is critical that this evidence gap on low- and middle-income countries is addressed swiftly, and that the infrastructure to collect and share data on educational performance in middle- and low-income countries is strengthened. Collecting and making available this data is a key prerequisite for fully understanding how learning progress and related outcomes have changed since the onset of the pandemic.[77]

A further limitation is that about half of the studies that we identify are rated as having a serious or critical risk of bias. We seek to limit the risk of bias in our results by excluding all studies rated to be at critical risk of bias from all of our analyses. Moreover, in Supplementary Fig. 3–6, we show that our results are robust to further excluding studies deemed to be at serious risk of bias. Future studies should minimize risk of bias in estimating learning deficits by employing research designs that appropriately account for common sources of bias. These include a lack of accounting for secular time trends, non-representative samples, and imbalances between treatment and comparison groups.

The persistence of learning deficits two and a half years into the pandemic highlights the need for well-designed, well-resourced and decisive policy initiatives to recover learning deficits. Policy-makers, schools, and families will need to identify and realize opportunities to complement and expand on regular school-based learning. Experimental evidence from low- and middle-income countries suggests that even relatively low-tech and low-cost learning interventions can have substantial, positive effects on students' learning progress in the context of remote learning. For example, sending SMS messages with numeracy problems accompanied by a short phone call was found to lead to substantial learning gains in numeracy in Botswana.[82] Sending motivational text messages successfully limited learning losses in math and Portuguese in Brazil.[83]

More evidence is needed to assess the effectiveness of other interventions for limiting or recovering learning deficits. Potential avenues include the use of the often extensive summer holidays to offer summer schools and learning camps, extending school days and school weeks, and organizing and scaling up tutoring programs. Further potential lies in developing, advertising and providing access to learning apps, online learning platforms, or educational TV programs that are free at the point of use. Many countries have already begun investing significant resources to capitalize on some of these opportunities. If these implemented interventions prove effective, and if the momentum of existing policy efforts is maintained and expanded, the disruptions to learning during the pandemic may be a window of opportunity to improve the education afforded to children.

Methods

Eligibility criteria

We consider all types of primary research, including peer-reviewed publications, preprints, working papers, and reports for inclusion. To be eligible for inclusion, studies have to measure learning progress using test scores that can be standardized across studies using Cohen's *d*. Moreover, studies have to be in English, Danish, Dutch, French, German, Norwegian, Spanish or Swedish.

Search strategy and study identification

We identify relevant studies using the following steps. First, we developed a Boolean search string defining our population (school-aged children), exposure (the COVID-19 pandemic), and outcomes of interest (learning progress). The full search string can be found in Section 1.1 of the Supplementary Information. Second, we used this string to search the following academic databases: Coronavirus Research Database, the Education Resources Information Centre (ERIC), International Bibliography of the Social Sciences (IBSS), Politics Collection (PAIS index, policy file index, political science database, and worldwide political science abstracts), Social Science Database, Sociology Collection (applied social science index [ASSIA] and abstracts, sociological abstracts, and sociology database), CINAHL, and Web of Science. Second, we hand-searched multiple preprint and working paper repositories (SSRN, MPRA, IZA, NBER, OSF Preprints, PsyArXiv, SocArXiv, and EdArXiv) and relevant policy websites, including the websites of the Organization for Economic Co-operation and Development (OECD), the United Nations (UN), the World Bank, and the Education Endowment Foundation (EEF). Third, we periodically posted our protocol via Twitter in order to crowdsource additional relevant studies not identified through the search. All titles and abstracts identified in our search were double-screened using the Rayyan online application.[89] Our initial search was conducted on April 27, 2021, and we conducted two forward and backward citation searches of all eligible studies identified in the above steps, on February 14, 2022, and on August 8, 2022, to ensure that our analysis includes recent relevant research.

Data extraction

From the studies that meet our inclusion criteria we extract all estimates of learning deficits during the pandemic, separately for math and reading and for different school grades. We also extract the corresponding sample size, standard error, date(s) of measurement, author name(s), and country. Last, we record whether studies differentiate between children's socio-economic background, which measure is used to this end, and whether studies find an increase, decrease or no change in learning inequality. We contacted study authors if any of the above information was missing in the study. Data extraction was performed by B.A.B. and validated independently by A.M.B.-M., with discrepancies resolved through discussion and by conferring with P.E.

Measurement and standardization

We standardize all estimates of learning deficits during the pandemic using Cohen's d, which expresses effect sizes in terms of standard deviations. Cohen's d is calculated as the difference in the mean learning gain in a given subject (math or reading) over two comparable periods before and after the onset of the pandemic, divided by the pooled standard deviation of learning progress in this subject:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s},$$

where

$$s = \sqrt{\frac{(s_1^2 + s_2^2)}{2}}.$$

Effect sizes expressed as β coefficients are converted to Cohen's d:

$$d = \frac{\beta}{se} \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

Subject We use a binary indicator for whether the study outcome is math or reading. One study does not differentiate the outcome but includes a composite of math and reading scores.[31]

Level of education We distinguish between primary and secondary education. We first consulted the original studies for this information. Where this was not stated in a given study, students' age was used in conjunction with information about education systems from external sources to determine the level of education.[86]

Country income level We follow the World Bank's classification of countries into four income groups: low, lower-middle, upper-middle, and high-income. Four countries in our sample are in the upper-middle group: Brazil, Colombia, Mexico and South Africa. All other countries are in the high-income group.

Data synthesis

We synthesize our data using three synthesis techniques. First, we generate a forest plot, based on all available estimates of learning progress during the pandemic. We pool estimates using a random-effects REML model and inverse variance weights to calculate an overall effect size (see Fig. 3).[87] Second, we code all estimates of changes in educational inequality between children from different socio-economic backgrounds during the pandemic, according to whether they indicate an increase, decrease, or no change in educational inequality. We visualize the resulting distribution using a harvest plot (see Fig. 5).[88] Third, given that the limited amount of available evidence precludes multivariate or causal analyses, we examine the bivariate association between COVID-19 learning deficits and the months in which learning was measured using a scatter plot (see Fig. 4), and the bivariate association between COVID-19 and learning subject, grade, and countries' income level, using a series of violin plots (see Fig. 6). The reported estimates, confidence intervals and statistical significance tests of these bivariate associations are based on common-effects models with standard errors clustered by study, and two-sided tests. With respect to statistical tests reported, the data distribution was assumed to be normal, but this was not formally tested. The distribution of estimates of learning deficits is shown separately for the different moderator categories in Fig. 6.

Pre-registration

We prospectively registered a protocol of our systematic review and meta-analysis in the International Prospective Register of Systematic Reviews (CRD42021249944) on 19 April 2021 (https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021249944).

Data availability

The data used in the analyses for this manuscript were compiled by the authors based on the studies identified in the systematic review. The data are available on the Open Science Framework repository (https://doi.org/10.17605/osf.io/u8gaz).

For our systematic review, we searched the following databases: Coronavirus Research Database (https://proquest.libguides.com/covid19), Education Resources Information Centre (ERIC) database (https://eric.ed.gov), International Bibliography of the Social Sciences (IBSS) (https://about.proquest.com/en/products-services/ibss-set-c/), Politics Collection (https://about.proquest.com/en/products-services/ProQuest-Politics-Collection/), Social Science Database (https://about.proquest.com/en/products-services/pq_social_science/), Sociology Collection (https://about.proquest.com/en/products-services/ProQuest-Sociology-Collection/), CINAHL (https://www.ebsco.com/products/research-databases/cinahl-database), and Web of Science (https://clarivate.com/webofsciencegroup/solutions/web-of-science/). We also searched the following preprint and working paper repositories: SSRN (https://papers.ssrn.com/sol3/DisplayJournalBrowse.cfm), MPRA (https://mpra.ub.uni-muenchen.de), IZA (https://www.iza.org/content/publications), NBER (https://www.nber.org/papers?page=1&perPage=50&sortBy=public_date), OSF Preprints (https://osf.io/preprints/), PsyArXiv (https://psyarxiv.com), SocArXiv (https://osf.io/preprints/socarxiv), and EdArXiv (https://edarxiv.org).

Code availability

All code needed to replicate our findings is available on the Open Science Framework repository (https://doi.org/10.17605/osf.io/u8gaz).

Acknowledgments.

Carlsberg Foundation grant CF19-0102 (A.M.B.-M.); Leverhulme Trust Large Centre Grant (P.E), the Swedish Research Council for Health, Working Life and Welfare (FORTE) grant 2016-07099 (P.E); the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program LIEPP (ANR-11-LABX-0091, ANR-11-IDEX-0005-02) and the Université Paris Cité IdEx (ANR-18-IDEX-0001) (P.E.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author Contributions.

B.A.B., A.M.B.-M., and P.E. designed the study; B.A.B., A.M.B.-M., and P.E. planned and implemented the search and screened studies; B.A.B., A.M.B.-M., and P.E. extracted relevant data from studies; B.A.B., A.M.B.-M., and P.E. conducted the quality appraisal; B.A.B., A.M.B.-M., and P.E. conducted the data analysis and visualization; B.A.B., A.M.B.-M., and P.E. wrote the manuscript.

Competing interests.

The authors declare no competing interests.

Tables

Table 1: Studies and estimates by country

Country	Studies
Australia [4]	Gore et al. 2021 [4][9]
Belgium [4]	Gambi and De Witte 2021 [2],[10] Maldonado and De Witte 2021 [2][11]
Brazil [2]	Lichand et al. 2022 [2][12]
Colombia [2]	Vegas 2022 [2][13]
Denmark [7]	Birkelund et al. 2021 [7][14]
Germany [9]	Depping et al. 2021 [4],[15] Ludewig et al. 2022 [1],[16] Schult et al. 2022a [2],[17] Schult et al. 2022b [2][18]
Italy [11]	Bazoli et al. 2022 [6],[19] Borgonovi and Ferrara 2022 [4],[20] Contini et al. 2022 [1][21]
Mexico [2]	Hevia et al. 2022 [2][22]
Netherlands [27]	Engzell et al. 2021 [8],[23] Haelermans 2021 [2],[24] Haelermans et al. 2021 [2],[25] Haelermans et al. 2022 [9],[26] Schuurman et al. 2021 [6][27]
South Africa [2]	Ardington et al. 2021 [2][28]
Spain [3]	Arenas and Gortazar 2022 [3][29]
Sweden [9]	Hallin et al. 2022 [9][30]
Switzerland [2]	Tomasik et al. 2020 [2][31]
United Kingdom [58]	Blainey and Hannay 2021a [12],[32] Blainey and Hannay 2021b [12],[33] Blainey and Hannay 2021c [12],[34] Department for Education 2021a [6],[35] Department for Education 2021b [2],[36] GL Assessment 2021 [4],[37] Rose et al. 2021a [2],[38] Rose et al. 2021b [4][39] Weidman et al. 2021 [4][40]
United States [149]	Domingue et al. 2021a [8],[41] Domingue et al. 2021b [4],[42] Kogan and Lavertu 2021a [1],[43] Kogan and Lavertu 2021b [9],[44] Kozakowski et al. 2021 [12],[45] Kuhfeld and Lewis 2022 [48],[46] Lewis et al. 2021 [12],[47] Locke et al. 2021 [14],[48] Pier et al. 2021 [25],[49]

Note: Countries and corresponding studies on COVID-19 learning deficits. The number of estimates are shown in brackets, by country (left) and study (right). Full references are indicated by superscript and listed in the bibliography.

Figures

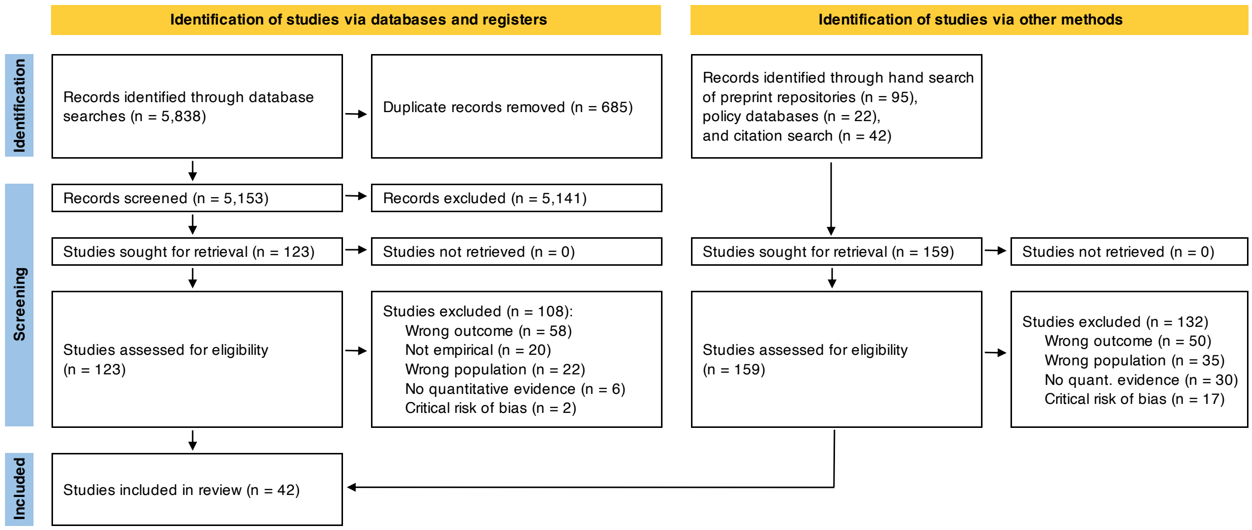

Figure 1: Study identification and selection process (PRISMA flow diagram)

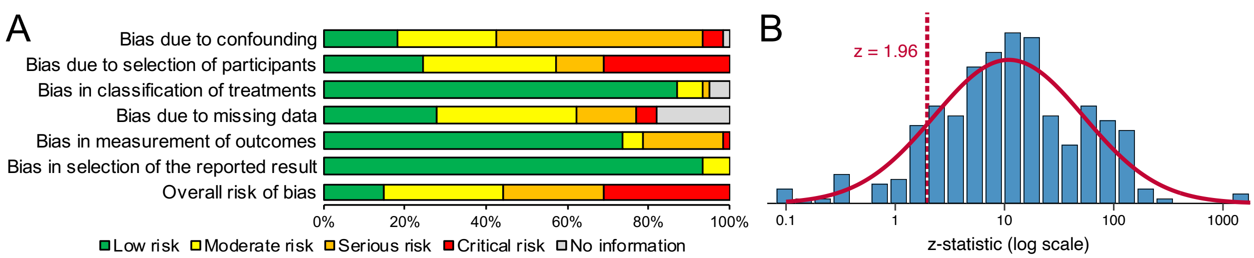

Figure 2: Risk of bias and publication bias. (a) Domain-specific and overall distribution of studies of COVID-19 learning deficits by risk of bias rating using ROBINS-I. Figure 2a includes studies rated to be at critical risk of bias ($n = 19$ out of a total of $n = 61$ studies shown in this figure). In line with ROBINS-I guidance, studies rated to be at critical risk of bias were excluded from all analyses and other figures in this article and in the Supplementary Information (including Figure 2b); (b) z-curve: Distribution of the z-scores of all estimates included in the meta-analysis ($n = 291$) to test for publication bias. The dotted line indicates $z = 1.96$ ($p = 0.050$), the conventional threshold for statistical significance. The overlaid curve shows a normal distribution. The absence of a spike in the distribution of the z-scores just above the threshold for statistical significance and the absence of a slump just below it indicate the absence of evidence for publication bias.

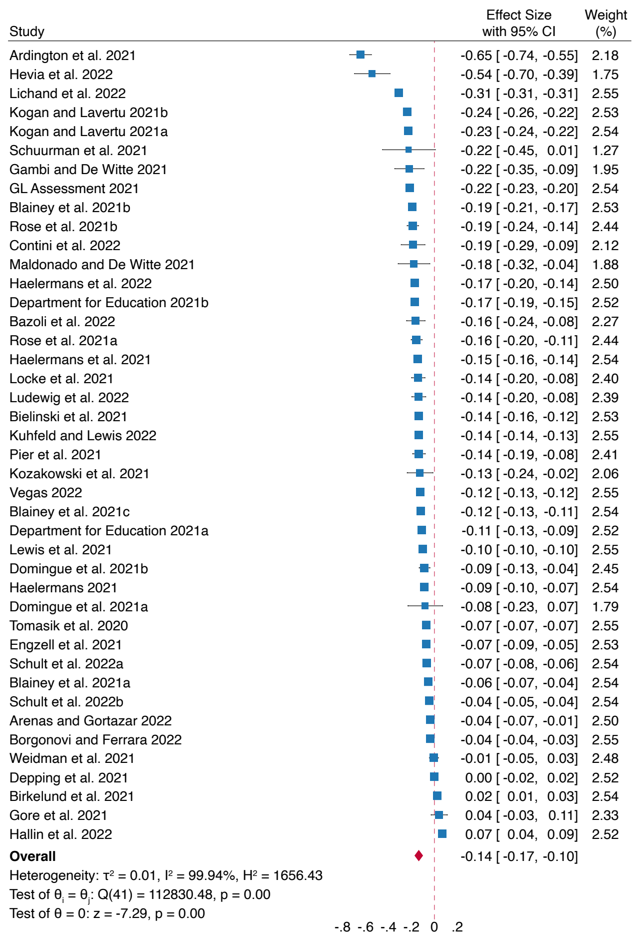

Figure 3: Forest plot showing individual estimates by study ($n = 42$), averaged across subjects and grade levels), and the overall effect size estimate, pooled using inverse variance weights and a random-effects model. Effect sizes are expressed in standard deviations, using Cohen's d, with 95% confidence intervals, and are sorted by magnitude.

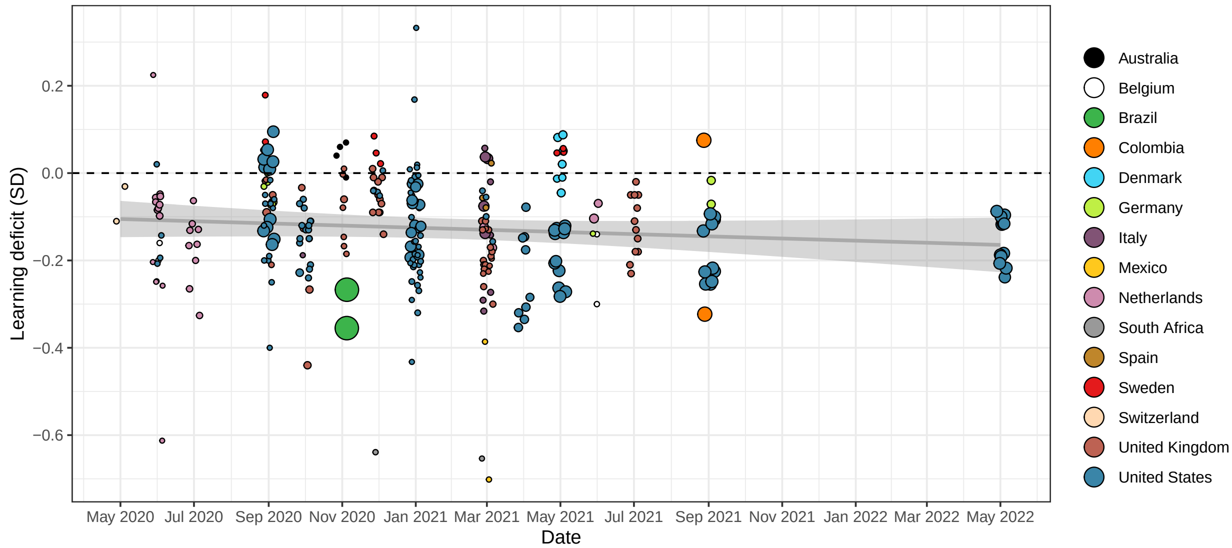

Figure 4: Estimates of COVID-19 learning deficits ($n = 291$), by date of measurement. The horizontal axis displays the date on which learning progress was measured. The vertical axis displays estimated learning deficits, expressed in standard deviations using Cohen's d. The color of the circles reflects the respective country, the size of the circles indicates the sample size for a given estimate, and the line displays a linear trend with a 95% confidence interval. The trend line is estimated as a linear regression using ordinary least squares, with standard errors clustered at the study level ($n = 42$ clusters). β months $= -0.00$, $t(41) = -7.30$, p two-tailed $= 0.097$, 95% c.i. $-0.01, 0.00$.

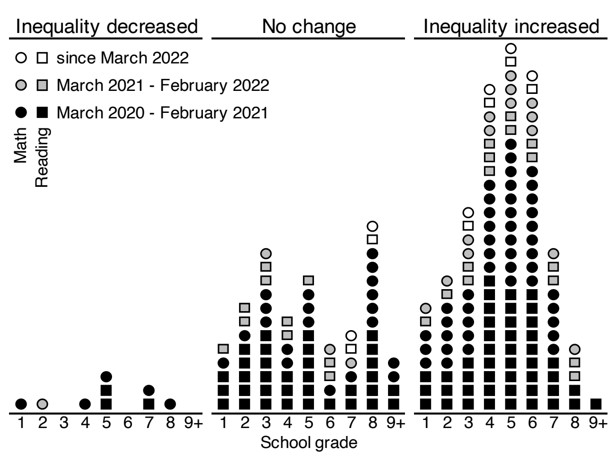

Figure 5: Harvest plot summarizing the evidence on changes in educational inequality between students from different socio-economic backgrounds during the pandemic. Each circle/square refers to one estimate of over-time change in inequality in math/reading performance ($n = 211$). Estimates that find a decrease/no change/increase in inequality are grouped on the left/middle/right. Within these categories, estimates are ordered horizontally by school grade. The shading indicates when in the pandemic a given measure was taken.

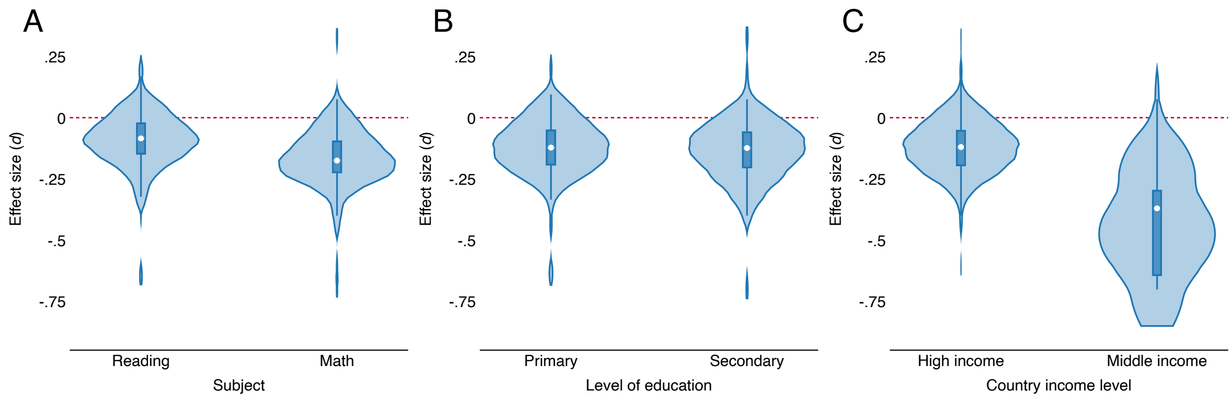

Figure 6: Variation in estimates of COVID-19 learning deficits ($n = 291$) across different characteristics. Each plot shows the distribution of COVID-19 learning deficit estimates for the respective subgroup, with the box marking the interquartile range and the white circle denoting the median. Whiskers mark upper and lower adjacent values: the furthest observation within 1.5 interquartile range of either side of the box. (a) Learning subject (reading vs. math). Median: reading $= -0.09$, math $= -0.18$. Interquartile range: reading $-0.15, -0.02$, math $-0.23, -0.09$. (b) Level of education (primary vs. secondary). Median: primary $= -0.12$, secondary $= -0.12$. Interquartile range: primary $-0.19, -0.05$, secondary $-0.21, -0.06$. (c) Country income level (high vs. middle). Median: high $= -0.12$, middle $= -0.37$. Interquartile range: high $-0.20, -0.05$, middle $-0.65, -0.30$.

References

[1] United Nations. The impact of COVID-19 on children. UN Policy Briefs (2020).

[2] Donnelly, R. & Patrinos, H. A. Learning loss during COVID-19: An early systematic review. *Prospects* 1–9 (2021).

[3] Hammerstein, S., König, C., Dreisörner, T. & Frey, A. Effects of COVID-19-related school closures on student achievement: A systematic review. *Frontiers in Psychology* 4020 (2021).

[4] Panagouli, E. *et al.* School performance among children and adolescents during COVID-19 pandemic: A systematic review. *Children* **8**, 1134 (2021).

[5] Patrinos, H. A., Vegas, E. & Carter-Rau, R. An analysis of COVID-19 student learning loss. World Bank, Washington, DC (2022).

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

[7] König, C. & Frey, A. The impact of COVID-19-related school closures on student achievement: A meta-analysis. *Educational Measurement: Issues and Practice* **41**, 16–22 (2022).

[8] Storey, N. & Zhang, Q. A meta-analysis of COVID learning loss. EdArXiv (2021).

[9] Gore, J., Fray, L., Miller, A., Harris, J. & Taggart, W. The impact of COVID-19 on student learning in New South Wales primary schools: An empirical study. *The Australian Educational Researcher* **48**, 605–637 (2021).

[10] Gambi, L. & De Witte, K. The resiliency of school outcomes after the COVID-19 pandemic: Standardised test scores and inequality one year after long term school closures. *FEB Research Report Department of Economics* (2021).

[11] Maldonado, J. E. & De Witte, K. The effect of school closures on standardised student test outcomes. *British Educational Research Journal* **48**, 49–94 (2021).

[12] Lichand, G., Doria, C. A., Leal-Neto, O. & Fernandes, J. P. C. The impacts of remote learning in secondary education during the pandemic in Brazil. *Nature Human Behaviour* **6**, 1079–1086 (2022).

[13] Vegas, E. COVID-19's impact on learning losses and learning inequality in Colombia. Center for Universal Education at Brookings (2022).

[14] Birkelund, J. F. & Karlson, K. B. No evidence of a major learning slide 14 months into the COVID-19 pandemic in Denmark. SocArXiv (2021).

[15] Depping, D., Lücken, M., Musekamp, F. & Thonke, F. Kompetenzstände Hamburger Schüler*innen vor und während der Corona-Pandemie. In Fickermann, D. & Edelstein, B. (eds.) *Schule während der Corona-Pandemie. Neue Ergebnisse und Überblick über ein dynamisches Forschungsfeld*, 51–79 (Münster & New York: Waxmann, 2021).

[16] Ludewig, U. *et al.* Die COVID-19 Pandemie und Lesekompetenz von Viertklässler*innen: Ergebnisse der IFS-Schulpanelstudie 2016-2021. Institut für Schulentwicklungsforschung, Universität Dortmund (2022).

[17] 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. *School Effectiveness and School Improvement* 1–20 (2022).

[18] Schult, J., Mahler, N., Fauth, B. & Lindner, M. A. Long-term consequences of repeated school closures during the COVID-19 pandemic for reading and mathematics competencies. *Frontiers in Education* **13** (2022). URL `https://doi.org/10.3389/feduc.2022.867316`.

[19] Bazoli, N., Marzadro, S., Schizzerotto, A. & Vergolini, L. Learning loss and students' social origins during the COVID-19 pandemic in Italy. *FBK-IRVAPP Working Papers* **3** (2022).

[20] Borgonovi, F. & Ferrara, A. The effects of COVID-19 on inequalities in educational achievement in Italy. Available at SSRN 4171968 (2022).

[21] Contini, D., Di Tommaso, M. L., Muratori, C., Piazzalunga, D. & Schiavon, L. Who lost the most? Mathematics achievement during the COVID-19 pandemic. *The BE Journal of Economic Analysis & Policy* **22**, 399–408 (2022).

[22] Hevia, F. J., Vergara-Lope, S., Velásquez-Durán, A. & Calderón, D. Estimation of the fundamental learning loss and learning poverty related to COVID-19 pandemic in Mexico. *International Journal of Educational Development* **88**, 102515 (2022).

[23] Engzell, P., Frey, A. & Verhagen, M. D. Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences* **118** (2021).

[24] Haelermans, C. Learning growth and inequality in primary education: Policy lessons from the COVID-19 crisis. The European Liberal Forum (ELF)-FORES (2021).

[25] Haelermans, C. *et al.* A full year COVID-19 crisis with interrupted learning and two school closures: The effects on learning growth and inequality in primary education. Maastricht University, Research Centre for Education and the Labour Market (ROA) (2021).

[26] Haelermans, C. *et al.* Sharp increase in inequality in education in times of the COVID-19-pandemic. *PLoS One* **17**, e0261114 (2022).

[27] Schuurman, T. M., Henrichs, L. F., Schuurman, N. K., Polderdijk, S. & Hornstra, L. Learning loss in vulnerable student populations after the first COVID-19 school closure in the Netherlands. *Scandinavian Journal of Educational Research* 1–18 (2021).

[28] Ardington, C., Wills, G. & Kotze, J. COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development* **86**, 102480 (2021).

[29] Arenas, A. & Gortazar, L. Learning loss one year after school closures. *Esade Working Paper* (2022).

[30] Hallin, A. E., Danielsson, H., Nordström, T. & Fälth, L. No learning loss in Sweden during the pandemic evidence from primary school reading assessments. *International Journal of Educational Research* **114**, 102011 (2022).

[31] 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. *International Journal of Psychology* **56**, 566–576 (2021).

[32] Blainey, K. & Hannay, T. The impact of school closures on autumn 2020 attainment. RS Assessment from Hodder Education (2021).

[33] Blainey, K. & Hannay, T. The impact of school closures on spring 2021 attainment. RS Assessment from Hodder Education (2021).

[34] Blainey, K. & Hannay, T. The effects of educational disruption on primary school attainment in summer 2021. RS Assessment from Hodder Education (2021).

[35] Department for Education. Understanding progress in the 2020/21 academic year: Complete findings from the autumn term. London: Department for Education (2021).

[36] Department for Education. Understanding progress in the 2020/21 academic year: Initial findings from the spring term. London: Department for Education (2021).

[37] GL Assessment. Impact of COVID-19 on attainment: Initial analysis. Brentford: GL Assessment (2021).

[38] Rose, S. *et al.* Impact of school closures and subsequent support strategies on attainment and socio-emotional wellbeing in key stage 1: Interim paper 1. London: National Foundation for Educational Research (NFER), Education Endowment Foundation (EEF).

[39] Rose, S. *et al.* Impact of school closures and subsequent support strategies on attainment and socio-emotional wellbeing in key stage 1: Interim paper 2. London: National Foundation for Educational Research (NFER), Education Endowment Foundation (EEF).

[40] Weidmann, B. *et al.* COVID-19 disruptions: Attainment gaps and primary school responses. *Education Endowment Foundation: London, UK* (2021).

[41] Domingue, B. W., Hough, H. J., Lang, D. & Yeatman, J. Changing patterns of growth in oral reading fluency during the COVID-19 pandemic. PACE Working Paper, Policy Analysis for California Education (2021).

[42] Domingue, B. *et al.* The effect of COVID on oral reading fluency during the 2020-2021 academic year. EdArXiv (2021).

[43] Kogan, V. & Lavertu, S. The COVID-19 pandemic and student achievement on Ohio's third-grade English language arts assessment. Ohio State University (2021).

[44] Kogan, V. & Lavertu, S. How the COVID-19 pandemic affected student learning in Ohio: Analysis of spring 2021 Ohio state tests. Ohio State University (2021).

[45] Kozakowski, W., Gill, B., Lavallee, P., Burnett, A. & Ladinsky, J. Changes in academic achievement in Pittsburgh public schools during remote instruction in the COVID-19 pandemic. Institute of Education Sciences (IES), U.S. Department of Education (2020).

[46] Kuhfeld, M. & Lewis, K. Student achievement in 2021–2022: Cause for hope and continued urgency. NWEA (2022).

[47] Lewis, K., Kuhfeld, M., Ruzek, E. & McEachin, A. Learning during COVID-19: Reading and math achievement in the 2020-21 school year. NWEA (2021).

[48] Locke, V. N., Patarapichayatham, C. & Lewis, S. Learning loss in reading and math in US schools due to the COVID-19 pandemic. Istation (2021).

[49] Pier, L., Christian, M., Tymeson, H. & Meyer, R. H. COVID-19 impacts on student learning: Evidence from interim assessments in California. PACE Working Paper, Policy Analysis for California Education (2021).

[50] Sterne, J. A. *et al.* ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* **355** (2016).

[51] Azevedo, J. P., Hasan, A., Goldemberg, D., Iqbal, S. A. & Geven, K. Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates. *World Bank Policy Research Working Paper* (2020).

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

[53] Hill, C. J., Bloom, H. S., Black, A. R. & Lipsey, M. W. Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives* **2**, 172–177 (2008).

[54] Belot, M. & Webbink, D. Do teacher strikes harm educational attainment of students? *Labour* **24**, 391–406 (2010).

[55] Jaume, D. & Willén, A. The long-run effects of teacher strikes: Evidence from Argentina. *Journal of Labor Economics* **37**, 1097–1139 (2019).

[56] Cygan-Rehm, K. Are there no wage returns to compulsory schooling in Germany? A reassessment. *Journal of Applied Econometrics* **37**, 218–223 (2022).

[57] Ichino, A. & Winter-Ebmer, R. The long-run educational cost of World War II. *Journal of Labor Economics* **22**, 57–87 (2004).

[58] Cooper, H., Nye, B., Charlton, K., Lindsay, J. & Greathouse, S. The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research* **66**, 227–268 (1996).

[59] Allington, R. L. *et al.* Addressing summer reading setback among economically disadvantaged elementary students. *Reading Psychology* **31**, 411–427 (2010).

[60] Smith, W. C. Consequences of school closure on access to education: Lessons from the 2013–2016 Ebola pandemic. *International Review of Education* **67**, 53–78 (2021).

[61] Andrabi, T., Daniels, B. & Das, J. Human capital accumulation and disasters: Evidence from the Pakistan earthquake of 2005. *Journal of Human Resources* 0520–10887R1 (2021).

[62] Parolin, Z. & Lee, E. K. Large socio-economic, geographic and demographic disparities exist in exposure to school closures. *Nature Human Behaviour* **5**, 522–528 (2021).

[63] Goudeau, S., Sanrey, C., Stanczak, A., Manstead, A. & Darnon, C. Why lockdown and distance learning during the COVID-19 pandemic are likely to increase the social class achievement gap. *Nature Human Behaviour* **5**, 1273–1281 (2021).

[64] Bailey, D. H., Duncan, G. J., Murnane, R. J. & Au Yeung, N. Achievement gaps in the wake of COVID-19. *Educational Researcher* **50**, 266–275 (2021).

[65] van de Werfhorst, H. G. Inequality in learning is a major concern after school closures. *Proceedings of the National Academy of Sciences* **118** (2021).

[66] Alexander, K. L., Entwisle, D. R. & Olson, L. S. Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis* **23**, 171–191 (2001).

[67] Aucejo, E. M. & Romano, T. F. Assessing the effect of school days and absences on test score performance. *Economics of Education Review* **55**, 70–87 (2016).

[68] Gottfried, M. A. The detrimental effects of missing school: Evidence from urban siblings. *American Journal of Education* **117**, 147–182 (2011).

[69] Goodman, J. Flaking out: Student absences and snow days as disruptions of instructional time. National Bureau of Economic Research (2014).

[70] Angrist, N., Djankov, S., Goldberg, P. K. & Patrinos, H. A. Measuring human capital using global learning data. *Nature* **592**, 403–408 (2021).

[71] Torche, F. Educational mobility in the developing world. In Iversen, V., Krishna, A. & Sen, K. (eds.) *Social Mobility in Developing Countries: Concepts, Methods, and Determinants*, 139–171 (Oxford University Press, 2021).

[72] World Bank. *World development report 2018: Learning to realize education's promise* (The World Bank, 2018).

[73] United Nations. Policy brief: Education during COVID-19 and beyond. United Nations (2020).

[74] UNESCO. One year into COVID-19 education disruption: Where do we stand? (2021). Accessed: 2022-09-01.

[75] Azevedo, J. P., Hasan, A., Goldemberg, D., Geven, K. & Iqbal, S. A. Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates. *The World Bank Research Observer* **36**, 1–40 (2021).

[76] Angrist, N. *et al.* Building back better to avert a learning catastrophe: Estimating learning loss from covid-19 school shutdowns in africa and facilitating short-term and long-term learning recovery. *International Journal of Educational Development* **84**, 102397 (2021).

[77] Conley, D. & Johnson, T. Opinion: Past is future for the era of COVID-19 research in the social sciences. *Proceedings of the National Academy of Sciences* **118** (2021).

[78] Major, L. E., Eyles, A., Machin, S. *et al.* Learning loss since lockdown: Variation across the home nations. Centre for Economic Performance, London School of Economics and Political Science (2021).

[79] Di Pietro, G., Biagi, F., Costa, P., Karpinski, Z. & Mazza, J. The likely impact of COVID-19 on education: Reflections based on the existing literature and recent international datasets. Publications Office of the European Union (2020).

[80] Fuchs-Schündeln, N., Krueger, D., Ludwig, A. & Popova, I. The long-term distributional and welfare effects of COVID-19 school closures. National Bureau of Economic Research (2020).

[81] Kaffenberger, M. Modelling the long-run learning impact of the COVID-19 learning shock: Actions to (more than) mitigate loss. *International Journal of Educational Development* **81**, 102326 (2021).

[82] Angrist, N., Bergman, P. & Matsheng, M. Experimental evidence on learning using low-tech when school is out. *Nature Human Behaviour* **6**, 941–950 (2022).

[83] Lichand, G., Christen, J. & van Egeraat, E. Do behavioral nudges work under remote learning? Evidence from Brazil during the pandemic. University of Zurich, Department of Economics, Working Paper (2022).

[84] Attewell, P. & Newman, K. S. *Growing gaps: Educational inequality around the world* (Oxford University Press, 2010).

[85] Betthäuser, B. A., Kaiser, C. & Trinh, N. A. Regional variation in inequality of educational opportunity across europe. *Socius* **7**, 23780231211019890 (2021).

[86] Eurydice. Eurybase: the information database on education systems in Europe (2021).

[87] Borenstein, M., Hedges, L. V., Higgins, J. P. & Rothstein, H. R. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods* **1**, 97–111 (2010).

[88] Ogilvie, D. *et al.* The harvest plot: A method for synthesising evidence about the differential effects of interventions. *BMC Medical Research Methodology* **8**, 1–7 (2008).

[89] Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan—a web and mobile app for systematic reviews. *Systematic reviews* **5**, 1–10 (2016).