Education
Endowment
Foundation

Impact of school closures on the attainment gap:

Rapid Evidence Assessment

May 2020

This report should be cited as: Education Endowment Foundation (2020) Impact of school closures on the attainment gap: Rapid Evidence Assessment, London: Education Endowment Foundation.

The report was authored by a team drawn from the EEF and beyond. The core team comprised: Rob Coe, Ben Weidmann, Robbie Coleman and Jon Kay. Additional support was provided by: Amy Clark, Amy Ellis-Thompson, Alexander Gnanapragasam, Trent Grassian, Steve Higgins, Christine Kelly, Harry Madgwick, Sue Morgan, Loren Parry, Shelby Roberts, Kirsten Sadler, Julie Sonnemann, Jennifer Stevenson, Sarah Tillotson, Matthew van Poortvliet and Mohammad Zaman.

The Education Endowment Foundation (EEF) is an independent charity dedicated to breaking the link between family income and educational achievement, ensuring that children and young people from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged pupils in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus (formerly Impetus Trust) and received a founding £125m grant from the Department for Education. Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

For more information about the EEF or this report please contact:

Rob Coe
Education Endowment Foundation
5th Floor, Millbank Tower
21-24 Millbank
SW1P 4QP

0207 802 1653

rob.coe@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

Executive Summary

This rapid evidence assessment aimed to examine the potential impact of school closures on the attainment gap, based on a systematic search of existing literature. Eleven studies were identified that provided quantitative evidence about the impact of school closures on attainment gaps. A subset of nine studies provided comparable estimates that could be synthesised.

Although the search included school closures with a range of causes (including due to adverse weather, hurricanes, teacher industrial action and pandemics) the final group of robust estimates all came from studies of summer holidays for primary-aged children.

Key findings and implications

1. **School closures are likely to reverse progress made to narrow the gap in the last decade**

The projections suggest that school closures will widen the attainment gap between disadvantaged children and their peers, likely reversing progress made to narrow the gap since 2011. The median estimate indicates that the gap would widen by 36%.

However, the estimated rate of gap widening varied substantially between studies, meaning that there is a high level of uncertainty around this average. Plausible "good" and "bad" estimates range from the gap widening from 11% to 75%.

2. **Supporting effective remote learning will mitigate the extent to which the gap widens**

Pupils can learn through remote teaching. However, ensuring the elements of effective teaching are present – for example through clear explanations, scaffolding and feedback – is more important than how or when lessons or support are provided.

To increase access to teaching, it would also be valuable to test the feasibility of online tuition as a way to supplement the support provided by teachers to disadvantaged children.

3. **Sustained support will be needed to help disadvantaged pupils catch up**

It is highly likely that the gap will have widened when pupils return to school, even if the strongest possible mitigatory steps are put in place. It is highly unlikely that a single or short-term catch-up strategy will be sufficient to compensate for lost learning due to school closures.

There is a risk that high levels of absence after schools formally reopen poses a particular risk for disadvantaged pupils.

Limitations

School closures due to coronavirus are different to the closures included in our rapid evidence assessment, meaning that the estimates above should be viewed as an imperfect guide.

For example, the search provided no examples of unplanned closures of the length already experienced by schools in England and the existing evidence on school closures almost exclusively focuses on summer holidays and younger children.

The projections do not incorporate information about activity during school closures due to coronavirus. It is possible that factors such as the provision of support for remote learning, or different rates of engagement with learning while at home mean that the projections are over- or under-estimates.

This briefing focuses on learning and does not aim to inform decisions about when pupils return to school, which should be based on pupil and teacher safety.

Introduction

Background and rationale for the review

The Covid-19 pandemic has led to school closures across the UK and many countries across the world, with the majority of pupils in these systems out of school, though supported and taught in various ways. There has been great concern that school closures will lead to slower rates of learning or learning loss, and there is a risk that the negative impact will be worse for pupils who are economically disadvantaged.

In this context a number of researchers and policy organisations have produced quick analyses of the potential impact of the school closures (e.g. Sims, 2020; Burgess and Sievertsen, 2020; Kuhfeld, & Tarasawa, 2020). These are impressive in their speed and relevance for policy thinking, but they highlight the diversity and potentially contested nature of the evidence that may be relevant. Some of the earlier reviews of the impact of school closure, although still widely cited, have been subjected to considerable criticism (e.g. von Hippel, 2019).

A rapid evidence assessment seeks to address this heterogeneity by ensuring that, as far as possible, all relevant evidence has been captured and considered. We believe the most recent systematic review of the evidence on summer learning loss is Cooper et al.'s (1996) study, and have not found any systematic review that covers the impact of other causes of closure (e.g. due to epidemics and adverse weather), or that focuses specifically on the differential impact of closure on disadvantaged pupils.

Previous research on the impact of school closures

The Cooper et al. (1996) meta-analysis has been a key source of evidence about the impact of summer closures and has been widely cited. Our inclusion date of 1995 onwards was set partly in order to capture any studies not included in that review. Cooper et al. reviewed a history of more than a hundred years of research on summer learning loss but focused their meta-analysis on studies published since 1975, of which they found 13 (two from Canada, all the others from the US; median publication date 1981). The headline estimate for summer learning loss was 10% of a standard deviation, or about one month of learning, slightly higher in maths and lower in reading, and increasing with age, at least in reading. They estimated that in reading and language, "on average, summer vacations created a gap of about 3 months between middle- and lower-class students" (p261). However, "the meta-analysis revealed no differential effect of summer on the mathematics skills of middle- and lower-class students" (p261). We should note that summer vacations in the US are typically around three months, about twice as long as those in England.

Despite its dominance in the field of summer learning loss, we believe the Cooper et al. (1996) meta-analysis suffers from a number of limitations that reduce the relevance of its claims to our questions. Some of these limitations derive from the technical methodological issues we discuss below, for example, problems of the scaling and standardisation of test scores. We consider two additional concerns here that are more specific to the Cooper et al study: the SES comparison and weightings.

*SES comparison***.** Our primary focus for this review is the impact of school closure on the disadvantage gap: the interaction between the amount of summer learning loss and students' socioeconomic status (SES). Part of the concern here is that the operationalisation of SES in the studies reviewed by Cooper et al. (1996) is not very clear. Different study populations were described as 'middle-income' or 'low-income' without much more detail: students in 28 samples were described as coming from low-income families, and students in 20 samples were described as coming from middle-income families. Generally, this assessment was based on the community served by a participating school or on the percentage of students in a sample who were eligible for free or reduced-price lunch (Cooper et al., p.252).

A bigger concern is that the comparison of effect sizes for income groups was largely a between-studies comparison. This allows considerable scope for confounds: any differences between study populations, measures of learning, other variables collected and/or controlled for, or analysis methods could affect their estimates. Comparing these estimates across studies mixes any genuine differences in rates of learning loss for different SES groups with these different artefacts, many of which are known to be capable of affecting the results substantially (von Hippel and Hamrock, 2019). A better approach, and the one we have adopted in our analysis, is to draw estimates of the impact on the gap from studies that compared the impact for both groups. That way, some of the main study-level artefacts are better controlled, since they are likely to affect both groups equally.

Weightings. Cooper et al. (1996) were faced with something of a no-win dilemma when their systematic search process included a single study that was four times as big as all the others combined. The Sustaining Effects Study from 1976 (confusingly abbreviated to SES in their paper) had already been the subject of a good deal of controversy as different researchers used different analytical approaches, different subsets of the data and different interpretations of the same results to argue different positions (Cooper et al. devote pages 247-250 to discussion of this history).

Most problematically, this large study appeared to find no evidence of overall summer learning loss. The standard approach in a meta-analysis is to weight the different study estimates, so that those with more precision (usually because they are larger) count more. However, in this case, a weighted average would simply represent the result from this one study – a positive (i.e. summer learning gain) effect of 2% of a standard deviation. Instead, Cooper et al. seem to prioritise either the unweighted mean effect size, or an estimate with the Sustaining Effects Study removed – with negative effects of 9% and 13% of a standard deviation, respectively. Their justification for this is largely that the Sustaining Effects Study included a longer interval between tests (140 days) that included about 8 weeks of instructional time. However, the average for all studies was 131 days, so it is not clear that this study was an outlier in that respect.

Methodological challenges in evaluating the impact of closure on the gap

Some of the more recent analyses of summer learning loss draw attention to a range of methodological issues and demonstrate that they can make a considerable difference to the estimate of the gap-widening effect of closure. We consider three methodological issues here: interval scales, standardisation, and analytical choices.

Interval scales. If we want to compare the gap between two groups, either on two different tests, or even on the same test at different points on the scale, we need to know that the intervals between scores on the test are equal across that range. Defining what is meant by ‘equal’ is not simple – we can easily get drawn into a complex technical argument about the nature of measurement (e.g. Perline et al., 1979) – but an equivalent change in learning must correspond to the same difference in scores. An extreme example where this fails would be a test with a ceiling effect, where candidates with quite different amounts of learning could be awarded the same (maximum) score.

Even where there are not clear ceiling (or floor) effects, most tests have different numbers of questions targeted at different levels of difficulty and hence differences in the number of marks associated with an equivalent change in performance at different points on the scale. Partly for this reason, modern tests generally use Item Response Theory (IRT) models to create equal-interval scales instead of just scoring as ‘number correct’. Von Hippel and Hamrock (2019) provide a detailed and clear explanation of how this problem can lead to the appearance of gap-widening that is a pure artefact of non-interval scales.

Standardisation and reliability. A related problem arises when different tests are used at the two time points. Most of the earlier studies, and some of the best known, estimate summer learning gaps by testing students at the end of one school year with a test designed for that grade, and then testing again at the start of the next with a different test for the next grade. Although this may seem like a necessary and perhaps obvious way to proceed, it generates problems for comparing gaps on two quite different tests. Without proper vertical scaling using IRT, the usual approach is to standardise the tests; in other words, to subtract each group’s mean and divide by its standard deviation. Unfortunately, under this procedure, if two tests have different reliabilities, and we split the group into two sub-groups with different means, the expected means of the standardised scores for each sub-group will not be equal.

For example, if the first test is less reliable (which is often the case, for example, as children move from Kindergarten to Grade 1), its standard deviation is inflated by random error and hence standardisation leads to shrinkage: the mean of a high-SES subgroup is depressed, while the mean of a low-SES subgroup is raised. The result is that the gap appears to have widened on the second test, even if nothing actually changed.

According to von Hippel and Hamrock (2019) these two measurement artefacts of interval scales and standardisation account for much of what has been claimed as a gap widening effect:

> *There are well-known findings suggesting that substantial test score gaps accumulate over summer vacation, but those findings were obtained using test scales that spread with age and fixed-form tests that change at the end of the summer. Patterns of summer gap growth do not necessarily replicate when using modern adaptive tests that are scored on IRT ability scales. If summer learning gaps are present, most of them are small and hard to discern through the fog of potential measurement artifacts (von Hippel and Hamrock, 2019, p.75).*

Different analytical choices. In any analysis there are choices to be made, some of which affect the results. Sometimes these are arbitrary choices where there is not a clear best option, but results will nevertheless differ. Sometimes there is a best way, but researchers do not choose it. Sometimes the choice reflects a different framing of the question: if you ask a different question you get a different answer. An example of this last kind is provided by Quinn et al. (2016) who show that we could think about a change in the scores of two subgroups either in terms of their absolute difference on an interval scale, or in terms of their relative overlap. Each approach is defensible as answering an important question about the gap-widening effects of the school year and summer vacation, but the answers they give are not the same. Similarly, Dumont and Ready (2020) frame the choice as an example of Lord's paradox, but also introduce a further dimension of choice: whether the disadvantage gap is defined between individual students who differ in their socioeconomic status, or between the students who attend schools with differing socioeconomic composition. Again, each of the resulting four choices leads to a different conclusion and, according to Dumont and Ready (2020), these differences largely account for the different perspectives and conflicting claims among different groups of scholars.

A further twist is that both Quinn et al. (2016) and Dumont and Ready (2020) analyse the same dataset, from the ECLS-K:2011 survey. Our analysis suggests that when they ask the same question, their results are pretty close, although not identical.

Understanding 'learning loss'

The studies we reviewed do not consider the question of what is meant by 'learning' and 'learning loss' in this context. Although a number of studies do compare the effects of school closure on different tests measuring different kinds of learning, and some even offer theoretical explanations for these differences, there is little consideration of the nature of the learning entailed and whether it is lost or has merely become rusty with disuse.

The distinction has implications for the remedy. If learning has been truly lost, it must be regained, which may be slow and painful. On the other hand, if it is merely rusty, it may be quickly regained with a small amount of practice. If students have not used a particular technique or procedure for a few months they are unlikely to perform it fluently if tested on arrival back in school. But if they had previously learnt it well, they might well regain that state quickly.

In other words, to understand fully the implications of learning loss, we need to know something about the process of learning regain. If that process is slow and effortful then the loss is painful. However, if the regain is quick and easy then we probably should not even call it 'loss'. Unfortunately, none of the studies we have reviewed tells us anything about the trajectory of learning regain.

One study that might inform this issue is Kuhfeld and Soland (2020). They find that when test scores are available at three points in the year, rates of growth are higher at the beginning of the year and slow towards the end. A possible explanation is that part of the growth at the beginning of the year is 'regain' that is acquired more quickly. Kuhfeld and Soland also show that this departure from linear growth has implications for estimates of summer learning loss that may be only half those derived from assuming linear growth.

Methods

Aims

The aims for the rapid evidence assessment are captured by the final research questions:

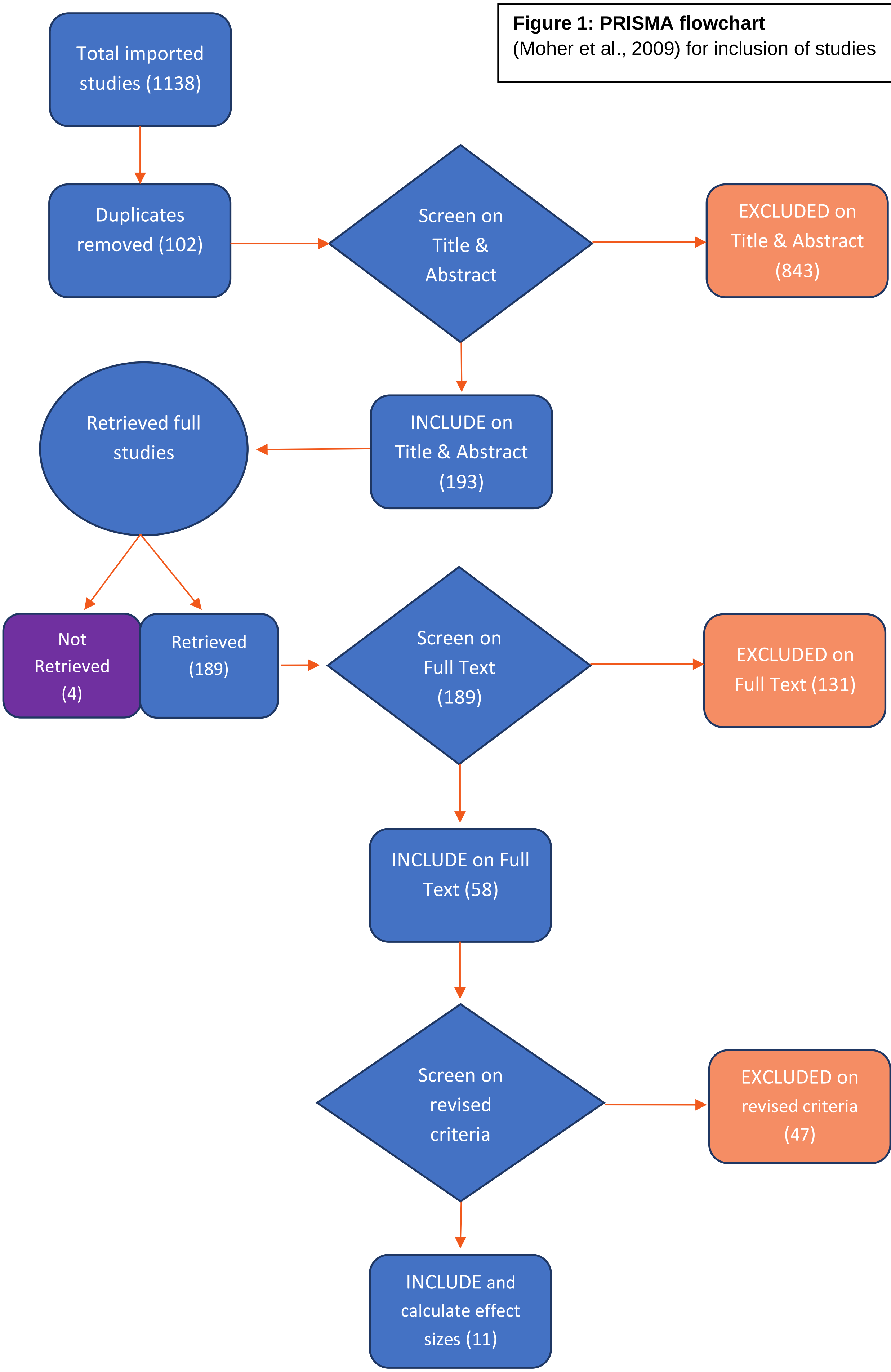
1. What evidence currently exists about the impact of different kinds of school closure (e.g. due to summer holidays; adverse weather, natural disasters)?
    a. On differential academic attainment for disadvantaged/others?
    b. On other outcomes related to education (e.g. impact on IQ or lifetime earnings)?
2. What factors moderate the impact? (e.g. age of pupils; subject/content area; types of attainment measure/methodology; length of closure; timing/conditions of testing)?
3. What evidence and theory helps us to understand the mechanisms by which school closure leads to learning loss and widening of attainment gaps (if it does)?

In the original design, there were additional research questions about the overall impact of closures on learning loss and about the estimate of the absolute effects of schooling. However, these foci were dropped at an early stage because of constraints of time. This process is described in the protocol which is publicly available on the EEF website.

Search and screening of studies

Although the rapid evidence assessment followed an explicit and transparent search process, we do not describe it as a systematic review: the process was systematic, but far from comprehensive, given constraints of time. Searches were run through Web of Science, ERIC and Google Scholar. Additional references were also found through the reference lists of included studies. The exact search terms, and inclusion/exclusion criteria can be found in the protocol. The PRISMA flowchart (Moher et al., 2009) for inclusion and exclusion of studies at each stage of the process is shown in Figure 1.

Figure 1: PRISMA flowchart
(Moher et al., 2009) for inclusion of studies



The original full-test screening process delivered 58 eligible studies. Given the time and resource available, this was more than we would be able to extract data from. At this point, we made a decision to focus the review on estimates of the effect of school closure on the gap between disadvantaged pupils and others, rather than on estimating the overall effect on learning loss. As many of the remaining studies neither attempted to estimate this gap nor provided enough detail for us to be able to extract a quantitative estimate, the scale of the review was reduced considerably.

Data extraction and coding

Data extraction was conducted on the selected set of 11 studies, in accordance with the protocol. The main variables extracted were:

- Reason for school closure
- Year in which closure occurred
- Country
- Age of pupils affected
- Number of pupils affected
- Subject or curriculum area in which learning assessed, and the test used
- Duration of the closure, and whether the analysis adjusted for imperfect overlap between that and the time between tests
- Whether the sample contained a full range of attainment
- How the study defined and operationalised the groups being compared (e.g. disadvantaged vs not), whether according to individual or group variables
- The dataset or study used
- Estimate of delta, the effect size change in the gap in population standard deviation units, per month of closure
- Estimate of the standard error of delta
- If there were duplicates in the dataset or sample analysed, whether that study provided the best estimate available

Analysis

Calculating comparable estimates

For each study-outcome pair we sought to calculate Δ, defined as *the rate of change for the gap between "disadvantaged" students and their peers, measured in effect-size units per month of closure.*[1]

In some instances, this required us to:

1. Rescale reported effects, so that the units were "effect-size per month". Equivalent rescaling was also applied to estimates of uncertainty.
2. Rescale reported effects so that they reflected a socioeconomic gap that was as-close-as-possible to the "disadvantage gap" in England. This is discussed below.

The studies in our meta-analysis operationalized disadvantage in different ways. We included studies where attainment gaps were defined at the student level. Gaps were defined by: income, poverty (e.g. Free or Reduced Price Lunch status), parental occupational, parental education, or some combination of these variables. This information is provided for each study in Table 2.

In England, the educational disadvantage gap at KS2 in 2018-19 is defined by comparing the mean attainment rank of the 30.5% of disadvantaged pupils, to the mean attainment rank of the other 59.5%

[1] We followed the definition of "disadvantage" used by the Department for Education, i.e. ""[d]isadvantaged pupils are defined as: those who were registered as eligible for free school meals at any point in the last six years, children looked after by a local authority or have left local authority care in England and Wales through adoption, a special guardianship order, a residence order or a child arrangements order."

of students.[2] The midpoint of the disadvantaged group is the 15th percentile of the "disadvantage" distribution, while the midpoint of the non-disadvantaged group is the 65th percentile. In terms of disadvantage, 'the gap' represents 50 percentiles. Assuming a normal distribution, this is equivalent to $\Phi(0.65) - \Phi(0.15) = 1.45$ standard deviation units.[3] If a research study presented the impact of school closures on a disadvantage gap defined as the difference between the mean attainment of the 10th and 90th percentile (which represents a gap of 2.46 standard deviation units) then we divided the reported estimate by $\frac{2.46}{1.45}$, and made equivalent changes to uncertainty estimates.

Meta-analysis procedure

Let $\widehat{\Delta}_i$ be the i^{th} estimate (for $i = 1, \ldots, 15$) of the rate at which the gap changes. Assume that these estimates have some distribution with $mean(\Delta_i)=\mu$, and $\text{var}(\Delta_i)=\tau^2$. At this stage, we make no assumptions about the shape of the distribution of Δ_i. We do, however, make a distributional assumption about the sampling variance: $\widehat{\Delta}_i | \Delta_\text{i} \sim \text{N}(\Delta_\text{i}, \sigma_i^2)$.

We estimate τ^2 using method of moments:[4]

$$\hat{\tau}^2 = \max\left\{0, \frac{Q-(k-1)}{\sum_i \hat{\sigma}_i^{-2} - \frac{\sum_i \hat{\sigma}_i^{-4}}{\sum_i \hat{\sigma}_i^{-2}}}\right\} \quad (1)$$

Where

$$Q = \sum_i \left(\widehat{\Delta}_i - \widehat{\Delta}\right)^2 \hat{\sigma}_i^{-2} \quad (2)$$

And

$$\widehat{\Delta} = \frac{\Sigma_i \widehat{\Delta}_i \hat{\sigma}_i^{-2}}{\Sigma_i \hat{\sigma}_i^{-2}} \quad (3)$$

In equation (1), k is the number of independent estimates.[5] In the interests of conservatism, we limit ourselves to setting k equal to the number of unique datasets available for analysis (n=8). This widens the predictive interval, described below.

Following Higgins et al. (2009), we estimate $\hat{\mu}$ as follows:

$$\hat{\mu} = \frac{\Sigma_i \Delta_i \omega_i}{\Sigma_i \omega_i} \quad (4)$$

Where

$$\omega_i = (\hat{\sigma}_i^2 + \hat{\tau}_i^2)^{-1} \quad (5)$$

---

[2] Department for Education data "National curriculum assessments at key stage 2, 2019 (revised)" Available at: https://www.gov.uk/government/statistics/national-curriculum-assessments-key-stage-2-2019-revised.

[3] Φ represents the normal CDF.

[4] Higgins, J. P., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 172*(1), 137-159.

[5] This could also be estimated from the data, although estimates would be highly uncertain. See Killip, S., Mahfoud, Z., & Pearce, K. (2004). What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Annals of Family Medicine, 2*(3), 204-208. doi:10.1370/afm.141.

Next, we generate empirical Bayes estimates of Δ for each study-outcome pair:

$$\Delta_i^* = \lambda_i \hat{\mu} + (1-\lambda_i)\widehat{\Delta}_i, \quad where\ \lambda_i = \frac{\sigma_i^2}{\sigma_i^2 + \tau^2} \quad (6)$$

Finally, we generate a distribution of Δ^{New}. This distribution is used in Figure 1 to communicate uncertainty about what the impact of a lengthy school closure might be. Again, we are guided here by Higgins et al. (2009). Specifically, we make the following distributional assumption:[6]

$$\frac{\Delta^{New} - \hat{\mu}}{\sqrt{\hat{\tau}^2 + \left(S\hat{E}(\hat{\mu})\right)^2}} \sim t_{k-2} \quad (7)$$

Figure 1 contains historical data on the attainment gap between disadvantaged pupils[7] and their peers, at the end of Key Stage 2. We chose the KS2 measure because primary schools are the overwhelming focus of research into school closures and attainment gaps. We selected KS2, rather than Early Years, based on our judgement that this was a more reliable measure over time.

The primary data historical data source in Figure 1 is the Department for Education data "National curriculum assessments at key stage 2, 2019 (revised)" (DfE, 2019b). We use table N5. The table reports the disadvantage index in which:

> *Comparisons are made by ordering pupil scores in reading and maths assessments at end of key stage 2 and assessing the difference in the average position of disadvantaged pupils and others. The mean rank of pupils in the disadvantaged and other pupils groups are subtracted from one another and multiplied up by a factor of 20 to give a value between -10 and +10 (where 0 indicates an equal distribution of scores).*

Let D_t be the disadvantage index in year t, and p_t^d be the proportion of disadvantaged children in year t. Next, let r_t^d be the mean rank of disadvantaged children (on a scale of 0-100), and $r_t^{\bar{d}}$ be the equivalent rank for non-disadvantaged pupils.

Based on the above description: $D_t = \frac{r_t^{\bar{d}} - r_t^d}{5}$. We use reported values of D_t and p_d to calculate the attainment gap in effect size units. The expression for D_t implies:

$$r_t^{\bar{d}} = 5 \cdot D_t - r_t^d \qquad (8)$$

The weighted average rank must be 50, implying:

$$50 = p_t^d r_t^d + (1 - p_t^d) r_t^{\bar{d}} \quad (9)$$

Combining (8) and (9) and rearranging, we have:

$$r_t^d = 50 - 5D_t(1 - p_t^d) \quad (10)$$

[6] A simple alternative would be to use the observed distribution of $\widehat{\Delta}$. However, this has two shortcomings: first the distribution of $\widehat{\Delta}$ is overdispersed (as it contains both τ^2 and σ_i^2); second, $\hat{\tau}$ is estimated with uncertainty, which isn't accounted for in the empirical distribution of $\widehat{\Delta}$, as per Higgins et al. (2009).

[7] We follow the Department for Education definition: "[d]isadvantaged pupils are defined as: those who were registered as eligible for free school meals at any point in the last six years, children looked after by a local authority or have left local authority care in England and Wales through adoption, a special guardianship order, a residence order or a child arrangements order."

To calculate the attainment gap in effect size units, we assume that attainment follows a normal distribution. Using (8) and (10):

$$\delta_t = \Phi\left(\frac{\mathrm{r}_t^{\bar{\mathrm{d}}}}{100}\right) - \Phi\left(\frac{\mathrm{r}_t^{\mathrm{d}}}{100}\right)$$

Where Φ is the normal CDF, and δ_t is the attainment gap defined in terms of effect size.

Moderator analysis

Our systematic review only found 15 comparable estimates of $\widehat{\Delta}_i$, making it difficult to perform useful moderator analysis. In particular, existing literature contained limited variation in outcomes (which were almost all "reading" and "maths") and age.

That said, we tested whether there were any clear differences in $\widehat{\Delta}_i$ for subject and age.

First, we fit a simple linear model to see if reading gaps seemed to grow faster than those maths:

$$\widehat{\Delta}_i = \beta_0 + \beta^{Read} Read_i + e_i \quad (model\ 1)$$

Where $Read_i$ is a binary indicator equal to one if the outcome of Δ_i was reading. We fit model 1 to the 13 estimates of maths/reading, using inverse-variance weights. The point estimate was $\hat{\beta}^{Read} = -0.012$. We then conducted a simple randomization inference. The null hypothesis being examined was $\beta^{Read} = 0$. The test statistic was the t-statistic associated with $\hat{\beta}^{Read}$from model 1. The observed test statistic was t^{obs}= -1.17. To generate a single draw under the null, we randomized the "subject" variable, re-fit model 1 and captured the t-statistic. We repeated this process 10000 times. The p-value can be defined as the proportion of draws under the null with an absolute value greater than the t^{obs}.[8] In this case, p=0.332. In short, we find no evidence of an association between outcome-type and the rate at which gaps. Given our power to detect such an association, this comes as no surprise.

We conducted a similar procedure in terms of age. This time we fit model 1 to our full meta-analytic sample of 15 estimates:

$$\widehat{\Delta}_i = \beta_0 + \beta^{Age} Age_i + e_i \quad (model\ 2)$$

Age is defined by the average age of the children reported in the study. Again, this model was fit with inverse-variance weights. There was no evidence of association between Age_i and $\widehat{\Delta}_i$.

[8] Davison, A. C., & Hinkley, D. V. (1997). Bootstrap methods and their application. Cambridge University Press.

Results

The extracted effect size estimates and other relevant variables from all the studies that provided quantitative estimates are shown in Table 1.

Table 1

Study_year	Cohort	Country	Grade_range	n_student	Subject	Break length	Dataset	Delta_gap	SE_Delta_gap
Studies where we could calculate FSM6 gap									
burkham_2004	1999	USA	K-1	3664	Other	2.6	ecls_k99	0.049	0.011
davies_2013	2011	Canada	1-3	1376	Reading	2.2	davies	0.011	0.004
dumont_2020/quinn_2016	2012	USA	K-1	3740	Maths	2.6	ecls_k11	0.043	0.006
dumont_2020/quinn_2016	2012	USA	K-1	3750	Reading	2.6	ecls_k11	0.016	0.006
dumont_2020/quinn_2016	2013	USA	1-2	3630	Maths	2.6	ecls_k11	-0.001	0.005
dumont_2020/quinn_2016	2013	USA	1-2	3630	Reading	2.6	ecls_k11	-0.021	0.005
lindahl_2001	1998	Sweden	5-6	556	Maths	2.2	lindahl	0.009	0.024
meyer_2017	2013	Germany	2-3	51	Reading	1.4	meyer	0.113	0.076
meyer_2017	2013	Germany	2-3	51	Other	1.4	meyer	-0.019	0.057
paechter_2015	2013	Austria	5-6	180	Maths	2.1	paechter	0.073	0.013
verachtert_2009	2003	Belgium	K-1	829	Maths	2.0	verachtert	0.012	0.029
vonhippel_2019	1999	USA	K-1	17779	Maths	2.6	ecls_k99	0.014	0.005
vonhippel_2019	1999	USA	K-1	17779	Reading	2.6	ecls_k99	0.015	0.004
vonhippel_2019	1986	USA	1-6	790	Maths	2.6	bss	0.047	0.012
vonhippel_2019	1986	USA	1-6	790	Reading	2.6	bss	0.033	0.014
Studies with other measures of disadvantage (not comparable)									
Study_year	**Cohort**	**Country**	**Grade_range**	**n_student**	**Subject**	**Break length**	**Dataset**	**Delta_gap**	**SE_Delta_gap**
campbell_2019	2016	USA	3-4	5513	Reading	2.6	campbell	-0.027	0.007
meyer_2017	2013	Germany	2-3	78	Reading	1.4	meyer	0.163	0.079
meyer_2017	2013	Germany	2-3	78	Other	1.4	meyer	0.054	0.065
meyer_2020	2014	NZ	4-7	4390	Other	1.4	nz_govt	0.166	0.029
vonhippel_2019	1999	USA	K-1	17779	Maths	2.6	ecls_k99	0.014	0.007
vonhippel_2019	2009	USA	K-1	177549	Maths	2.6	nwea	-0.005	0.005
vonhippel_2019	1999	USA	K-1	17779	Reading	2.6	ecls_k99	0.026	0.007
vonhippel_2019	2009	USA	K-1	177549	Reading	2.6	nwea	-0.019	0.005
vonhippel_2019	2009	USA	1-6	177549	Maths	2.6	nwea	-0.011	0.002
vonhippel_2019	1986	USA	1-6	790	Maths	2.6	bss	0.034	0.012
vonhippel_2019	2009	USA	1-6	177549	Reading	2.6	nwea	-0.007	0.002
vonhippel_2019	1986	USA	1-6	790	Reading	2.6	bss	0.057	0.014
vonhippel_2019	2009	USA	1-8	177549	Reading	2.6	nwea	-0.014	0.002
vonhippel_2019	2009	USA	1-8	177549	Maths	2.6	nwea	-0.011	0.002

Table 1 contains information on the following variables:

- Study_year: first author, and the year in which the study was published
- Cohort: year in which data was collected
- Country
- Grade range: reported grade range of students
- n_students: number of students reported in the study
- Subject: “Other” represents either the general knowledge test in the ECLS-K study, or writing tests
- Break length: how long were schools closed? (measured in months)
- Dataset: indication about the underlying dataset used. bss = Beginning School Study; ecls = early childhood longitudinal study; nwea = GRD study, maintained by the northwest evaluation association.
- Delta_gap [Δ]: main outcome variable defined as the rate of change in the gap between FSM6 and nonFSM6 pupils, measured in effect-size units per month
- SE_Delta_gap: (SE(Δ)): standard error of Δ

The set of studies that appear in Table 1 are:

- All studies we use in our meta-analysis. This includes all studies where we have been able to calculate an estimate of Δ. In cases where multiple authors analysed the same sample of children, we have removed the studies with clear methodological deficiencies. For one dataset, ecls_k11, we found two analyses of excellent quality (Dumont 2020 and Quinn 2016). In order to avoid double counting, we took an average of the estimates from these analyses.
- All other studies where we found quantitative estimates of how SES gaps changed during school closures, but where disadvantage was measured at the school level, or in some way that could not be credibly converted into measure comparable with the FSM gap.

Table 2 partially repeats information from Table 1, but is limited to the nine studies that provided the 15 estimates that are comparable and of high quality, and provides additional information about the definition of disadvantage used in each study. It also shows the Bayesian shrunken estimates, Δ_i^*, that represent our best estimate of the likely contribution of each study to an overall measure. These results are also shown graphically in a forest plot, in Figure 2.

Table 2

Study_year	Country	n_student	Subject	Disadvantage Definition	$\widehat{\Delta}_i$ (delta_gap)	$SE(\widehat{\Delta}_i)$ (standard error)	Δ_i^* (shrunken estimates)
burkam_2004	USA	3664	Other	SES[1]	0.049	0.011	0.043
davies_2013	Canada	1376	Reading	SES[2]	0.011	0.004	0.011
dumont_2020/quinn_2016	USA	3630	Reading	SES[3]	-0.021	0.005	-0.018
dumont_2020/quinn_2016	USA	3630	Maths	SES[3]	-0.001	0.005	0.000
dumont_2020/quinn_2016	USA	3750	Reading	SES[3]	0.016	0.006	0.017
dumont_2020/quinn_2016	USA	3740	Maths	SES[3]	0.043	0.006	0.041
lindahl_2001	Sweden	556	Maths	SES[4]	0.009	0.024	0.017
meyer_2017	Germany	52	Other	Parental occupation[5]	-0.019	0.057	0.018
meyer_2017	Germany	51	Reading	Parental occupation[5]	0.113	0.076	0.028
paechter_2015	Austria	182	Maths	Mother's education[6]	0.073	0.013	0.059
verachtert_2009	Belgium	829	Maths	SES[7]	0.012	0.029	0.019
vonhippel_2019	USA	17779	Maths	FRPL status[8]	0.014	0.005	0.014
vonhippel_2019	USA	17779	Reading	FRPL status[8]	0.015	0.004	0.015
vonhippel_2019	USA	790	Reading	FRPL status[8]	0.033	0.014	0.030
vonhippel_2019	USA	790	Maths	FRPL status[8]	0.047	0.012	0.041

[1]Composite measure of parents' education, parents' occupational prestige, and household income; [2]composite measure of parent education, other parent education, income each standardized and summed; [3]NCES-created socioeconomic status (SES) variable, which is a composite of family income, parental education, and occupational prestige; [4]Census-based measure, combining the mean income and mean parental years education on the block of the relevant student, among households on that block where parents are aged 28-54 and kids are aged 10-12; [5]Highest Socio-Economic Index of Occupational Status; [6] Binary indicator for whether or not a mother sat the university entrance exam; [7]Composite measure including the educational level of both parents, the professional status of both parents, and the household income; [8]Free and Reduced Price Lunch status.

Figure 2: Change in attainment gap (by study and outcome)

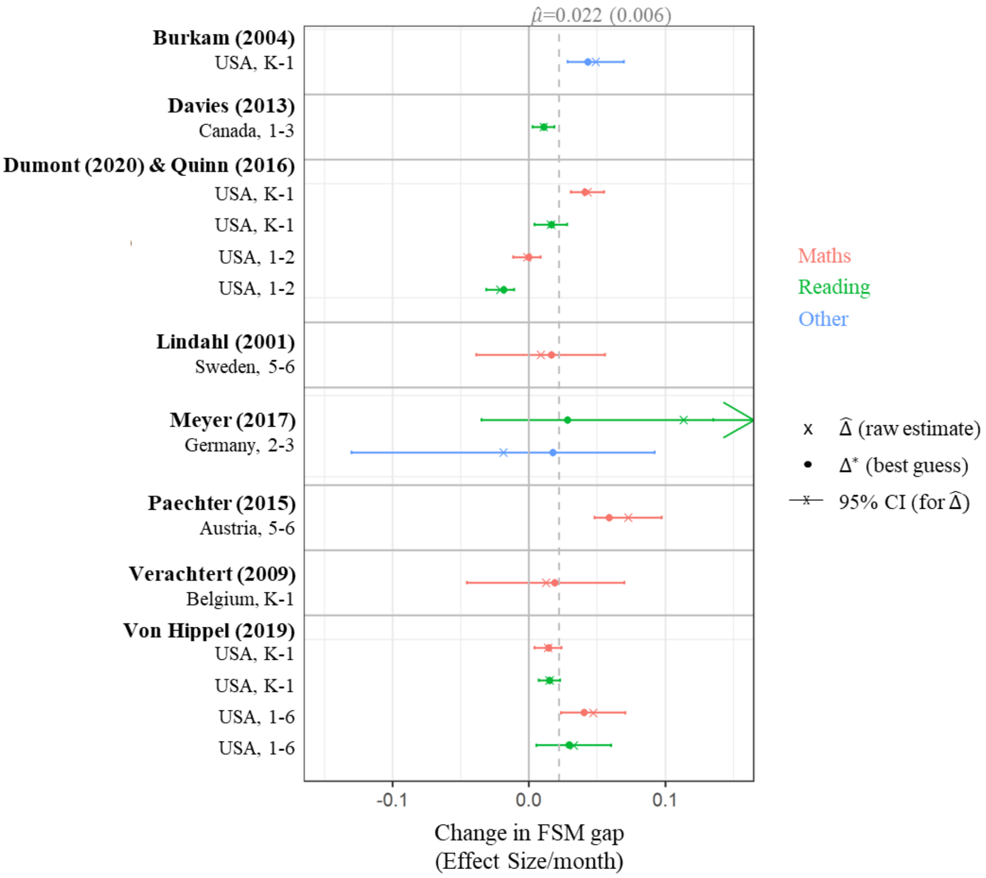
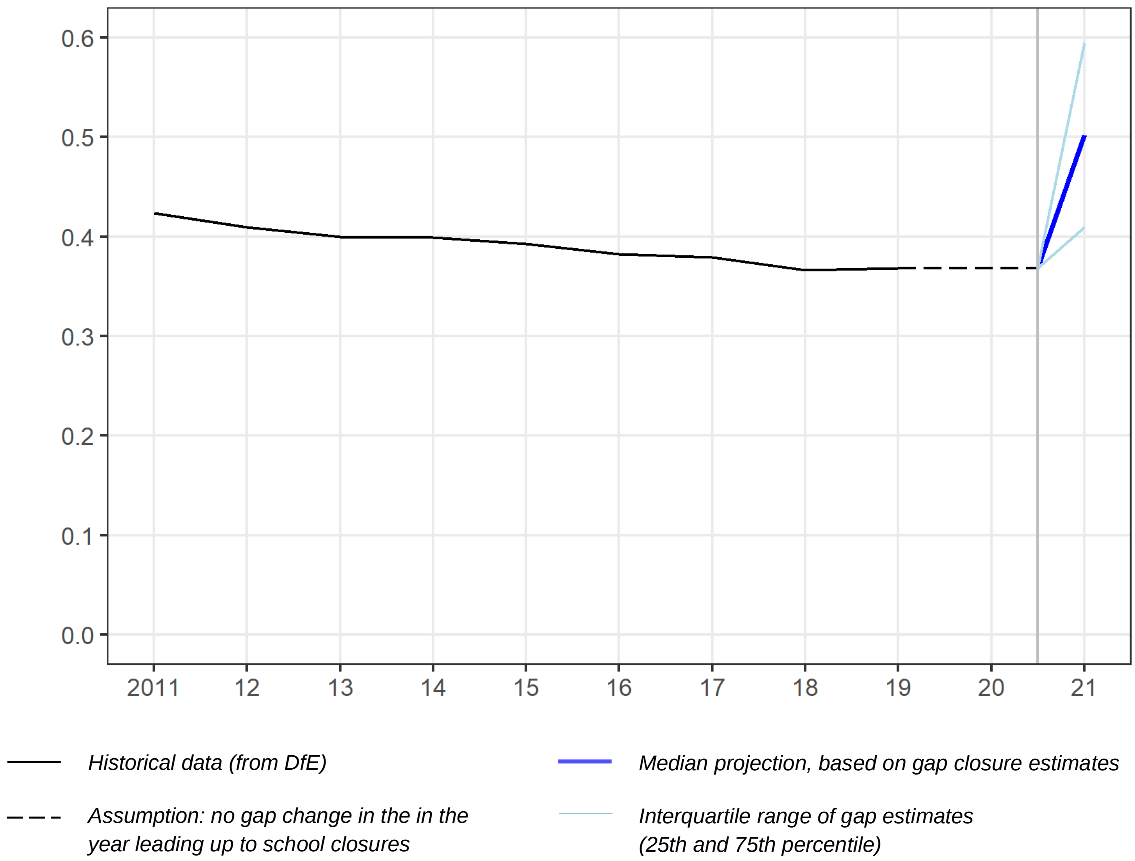


Figure 2 presents the overall estimate from the meta-analysis of 0.022 (standard error = 0.006) standard deviations per month, as the amount by which the attainment gap between disadvantaged pupils and their peers may be expected to grow.

Figure 3 presents this estimate in the context of the existing and historical gap in attainment at Key Stage 2. The thick blue line is the median projection and covers the second half of the 2019-20 academic year. The uncertainty in our projection is illustrated by the two light blue lines. These represent the 25th and 75th percentile and can be thought of as plausible "good" and "bad" cases. The median estimate, based on the existing literature on school closures, indicates that the gap would widen by 36% by September 2020. The range displayed stems from 11% to 75%. The projections suggest that school closures will widen the attainment gap between disadvantaged children and their peers, likely reversing progress made to narrow the gap since 2011.

It should be acknowledged that this range of estimates only contains 50% of our uncertainty. In other words, if we were to observe a new estimate of Δ^{New} (for example from a new study that we had not previously seen) we would expect the effect size to fall outside of the 25-75th percentile range half of the time. In addition, it should be noted that some studies contained estimates indicating that in certain subjects and for some age groups the gap may narrow.

Figure 3: historical estimates of the FSM gap at KS2



Wider literature and limitations

It is important to stress that the current school closures due to coronavirus are different to the closures included in our systematic review, meaning that the estimates above should be viewed as an imperfect guide. For example, our search provided no examples of unplanned closures of the length already experienced by schools in England and the existing evidence on school closures almost exclusively focuses on summer holidays and younger children.

One clear difference between closures due to coronavirus and the closures analysed above is that schools have been providing substantial support to enable remote learning, including by providing resources and online teaching. In addition, national initiatives such as Oak Academy have been watched by millions of pupils (Schools Week, 2020).

A recent review of remote teaching conducted by the EEF found that remote learning can be effective, given the right conditions (EEF, 2020a). A review of the impact of online schools in the US found that although pupil outcomes online were, on average, poorer for all groups compared to in-school learning, attainment gaps between disadvantaged students and their peers were no wider (Woodworth, 2015). This might suggest that, if online schooling were operating well for all children, then the projection of gap widening above may be overstated.

It is also likely that some pupils will return to school earlier than September. For these pupils, the projection of gap widening in Figure 3 may also be overstated.

However, while efforts to support remote learning are likely to have been of considerable benefit to many children, and are likely to have reduced the overall amount of learning loss due to closures, there are indications that, overall, the remote learning that has taken place during school closures is likely to have further widened rather than narrowed the gap.

Surveys of teachers and parents in England in 2020 show that many pupils are not engaging in high-quality home learning and that disadvantaged pupils appear to be learning less than their peers (e.g. Sutton Trust, 2020; Institute for Fiscal Studies, 2020). Findings from these studies suggest that children from the most disadvantaged families are spending less time on learning activities, are submitting less work and typically have access to fewer resources at home.

There are several additional reasons why the studies reviewed here may underestimate the impact of school closures. For example, the estimates do not capture the fact that coronavirus has had a differentially large economic and health impact on disadvantaged families (e.g. Office for National Statistics, 2020; Douglas et al., 2020), which may in turn affect educational outcomes.

Given the lack of evidence about how coronavirus-specific factors might affect the rate at which the attainment gap will widen, we have not attempted to include these factors in our quantitative projections.

Estimates by subject, phase and prior attainment

Testing for differences across subject and age is severely limited by two factors. First, there is a lack of variation in existing research. All but two of the studies focus on either reading or maths, and all the estimates in our meta-analysis come from research conducted in primary schools. Second, we have a small set of estimates to draw on, as our systematic review only yielded 15 comparable estimates.

With those important caveats in mind, we note that we did not find any evidence of an association between gap estimates and age or subject. It is worth stressing that these findings are an “absence of evidence”, rather than clear evidence of no difference. The EEF hopes to address this shortage of evidence in future work.

Similarly, there is little evidence related to differences between pupils with high or low prior attainment. Gershenson (2017) finds that over the summer higher-attaining disadvantaged children fall behind other higher attainers at a faster rate than other groups. However, this result was not replicated in maths, and overall there was not enough evidence to draw clear conclusions.

Discussion and implications

Given the impact on the gap identified by the rapid evidence assessment, we now explore evidence to inform efforts to mitigate the extent to which the gap widens and to compensate for lost learning, including by drawing on wider literature on effective approaches (e.g. EEF, 2020a).

Supporting learning at home

Two factors affecting learning while pupils are at home are *remote learning* and *parental involvement.*

Remote learning

It is very hard to use technology to replace the learning relationships that exist between teachers and pupils in the classroom. However, providing access to teaching via technology has the potential to make a small-to-moderate positive impact on learning during school closures.

A key challenge is ensuring that access to teaching is provided to all pupils. There is a significant risk that disadvantaged children have less access to teaching than their peers, in part due to having reduced access to technology, exacerbating the impact of school closures on the attainment gap.

A rapid evidence assessment on remote learning conducted by the EEF (available here) also emphasised that the pedagogical quality of remote learning is more important than how lessons are delivered. Ensuring the elements of effective teaching are present – for example; clear explanations, scaffolding and feedback – is more important than how or when they are provided (EEF, 2020a). It is unlikely that providing pupils with access to resources without support will improve learning.

To increase access to teaching, it would also be valuable to test the feasibility of online tuition as a way to supplement the support provided by teachers to disadvantaged children.

In addition to providing access to technology, ensuring that teachers and pupils are provided with support and guidance to use specific platforms is essential, particularly if new forms of technology are being implemented (EEF, 2020a).

Parental involvement

Parental engagement in children's learning and the quality of the home learning environment are associated with improved academic outcomes at all ages (EEF, 2020b).

However, the evidence indicates that it is very challenging for schools to increase levels of parental engagement successfully. Schools may need support in communicating effectively with parents and in helping parents understand specific ways to help their child learn.

It is likely to be particularly valuable to focus on developing and maintaining two-way communication with parents and promoting the development and maintenance of reading habits.

The effectiveness of strategies will differ by age group. For example, in primary schooling, shared book reading and linked activities such as building vocabulary and practising spellings are valuable, while in secondary schools parents can support children to read independently and create study routines (e.g. Meyer et al., 2015).

Parents can support their children by encouraging them to set goals, plan, and manage their time, effort, and emotions. This type of support can help children, in particular older children, to regulate their own learning and will often be more valuable than direct help with schoolwork.

EEF resources for schools on supporting parents during school closures are available here.

Supporting catch-up after pupils return to school

It is highly likely that the gap will have widened when pupils return to school, even if the strongest possible mitigatory steps are put in place. Approaches that could help pupils catch up include:

- Targeted support
- Professional development for teachers

Key risks related to *pupil absence* and *sustained support* are also highlighted.

Targeted support

The EEF has identified a list of 18 promising projects that have been evaluated and shown to have positive impacts on learning, with particularly strong effects for disadvantaged children in most cases.

Tuition is likely to be a particularly effective catch up approach. The EEF estimates that the average impact of one-to-one tuition is five additional months' progress (EEF, 2020b). An evaluation of low-cost tutoring delivered by university students showed a positive impact on learning of three additional months' progress (Torgerson, 2018).

Professional development

Alongside targeted interventions, improving the quality of teaching is the strongest lever schools have to improve pupil outcomes, particularly for disadvantaged students.

The EEF recommends that when spending the pupil premium schools take a tiered approach, starting with efforts to improve teaching quality.

Priorities for professional development might include: ensuring high-quality materials are available for early career teachers linked to the Early Career Framework; online courses linked to the best available evidence on improve literacy and maths; and online courses linked to pedagogical approaches that are likely to be particularly effective for disadvantaged learners, e.g. metacognition.

Pupil absence

A key risk relates to the distinction between school closures and pupil absence. Notwithstanding the overall projections above Goodman (2014) emphasises that schools are typically able to deal relatively effectively with school closures, be they planned or unplanned. In contrast, missing school due to absence is typically associated with a substantially greater negative effect.

Part of this difference is likely to be driven by methodological challenges, i.e. there are likely to be unobservable factors that are associated with being absent that lead to low attainment, even when pupils compared to apparently similar peers. However, it is also likely that it is easier for teachers to respond to closures — for example, by repeating key content as a class — than it is to support individual children who have been absent (e.g. Department for Education, 2016; Gottfried, 2010).

The severe negative effect of absence poses a particularly high risk for disadvantaged children (Department for Education, 2019b), who typically have lower rates of attendance and whose families have indicated that the they would be substantially less likely to send their child back to school if given the choice (IFS, 2020).

Sustained support

Sustained support will be required to help disadvantaged pupils catch-up after they return to school. While a focused catch-up programme – including assessment and targeted support – would be beneficial when pupils first return to school, it is unlikely that a single catch-up strategy will be sufficient to compensate for lost learning due to school closures.

Additional resources

The Education Endowment Foundation has created a number of resources that are relevant to supporting learners during the Covid-19 outbreak. All resources can be found here. Some of the resources that directly relate to the findings of this rapid evidence assessment are detailed below:

Resource	Description	Link
Guidance reports	EEF guidance reports provide clear and actionable recommendations for teachers on a range of high-priority issues based on the best available evidence.	https://educationendowmentfoundation.org.uk/tools/guidance-reports/
Parental engagement guidance report	Four recommendations on working with parents to support their child's learning.	https://educationendowmentfoundation.org.uk/tools/guidance-reports/working-with-parents-to-support-childrens-learning/
Parental engagement evidence review	The underlying evidence review for the parental engagement guidance report.	https://educationendowmentfoundation.org.uk/evidence-summaries/evidence-reviews/parental-engagement/
Digital technology guidance report	Four recommendations on using digital technology to improve children's learning.	https://educationendowmentfoundation.org.uk/tools/guidance-reports/using-digital-technology-to-improve-learning/
Digital technology evidence review	The underlying evidence review for the digital technology guidance report.	https://educationendowmentfoundation.org.uk/evidence-summaries/evidence-reviews/digital-technology-2019/
Metacognition guidance report	Seven recommendations for teaching self-regulated learning and metacognition,	https://educationendowmentfoundation.org.uk/tools/guidance-reports/metacognition-and-self-regulated-learning/
Teaching and Learning Toolkit	The Teaching and Learning Toolkit provides an accessible summary of the evidence across 35 different approaches aimed at improving pupil outcomes	https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/
Peer tutoring	Toolkit summary of peer tutoring approaches	https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/peer-tutoring/
Metacognition and self-regulation	Toolkit summary of metacognition and self-regulation approaches	https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/meta-cognition-and-self-regulation/
Parental engagement	Toolkit summary of parental engagement approaches	https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/parental-engagement/
Homework	Toolkit summary of homework (primary and secondary)	https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/homework-primary/ https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/homework-secondary/
Digital technology	Toolkit summary of digital technology approaches	https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/digital-technology/
EEF-funded evaluations	This is the full list of evaluations that have been funded by the EEF.	https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/
What Works Clearinghouse list of studies	A list of studies that examine the impact of remote learning approaches, identified by the What Works Clearinghouse	https://ies.ed.gov/ncee/wwc/distancelearningstudy

References

References from systematic search (*included in quantitative analysis)

Alexander K., Entwisle D., & Olson L. (2001). Schools, achievement, and inequality: A seasonal perspective. Educational evaluation and policy analysis, 23(2), pp.171-191.

Alexander K., Entwisle D., & Olson L. (2007). Summer Learning and Its Implications: Insights from the Beginning School Study. New Directions for Youth Development 114(114):11-32.

Alexander K., Entwisle D., & Olson L. (2007). Lasting consequences of the summer learning gap. American Sociological Review, 72(2), pp.167-180.

Baker J. (2018). Summer Learning Loss in Reading Achievement: Effects of Demographic Variables and Summer Activity. Available at: https://conservancy.umn.edu/handle/11299/206340.

Baggerly J., & Ferretti, L. (2008). The Impact of the 2004 Hurricanes on Florida Comprehensive Assessment Test Scores: Implications for School Counselors.

*Benson J & Borman G. (2010). Family, neighborhood, and school settings across seasons: When do socioeconomic context and racial composition matter for the reading achievement growth of young children. Teachers College Record, 112(5), pp.1338-1390.

*Burkam D., Ready D., Lee V., & LoGerfo L. (2004). Social-class differences in summer learning between kindergarten and first grade: Model specification and estimation. Sociology of Education, 77(1), pp.1-31.

*Campbell, L., Sutter, C. & Lambie, G. (2019). An investigation of the summer learning effect on fourth grade students' reading schools. Reading Psychology, 40(5), pp.465-490.

Cooper H., Nye B., Charlton K., Lindsay J., & Greathouse S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. Review of educational research, 66(3), pp.227-268.

Cooper H., Valentine J., Charlton K. & April, M. (2003). The effects of modified school calendars on student achievement and on school and community attitudes. Review of Educational Research, 73(1), pp.1-52.

*Davies, S. & Aurini, J. (2013). Summer Learning Inequality in Ontario. Canadian Public Policy-Analyse De Politiques, 39(2), pp.287-307.

*Dumont, H. & Ready, D. (2020). Do Schools Reduce or Exacerbate Inequality? How the Associations Between Student Achievement and Achievement Growth Influence Our Understanding of the Role of Schooling. American Educational Research Journal, 57(2), pp.728-774.

Entwisle D. & Alexander, K. (1995). A Parent's Economic Shadow – Family-Structure Versus Family Resources As Influences On Early School-Achievement. Journal Of Marriage And The Family, 57(2), pp.399-409.

Fitzpatrick M., Grissmer, D. & Hastedt, S. (2011). What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment. Economics of Education Review, 30(2), pp.269-279.

Georges A. & Pallas, A. (2010). New Look at a Persistent Problem: Inequality, Mathematics Achievement, and Teaching. Journal of Educational Research, 103(4), pp.274-290.

Gershenson S. & Hayes, M. (2017). The summer learning of exceptional students. American Journal of Education, 123(3), pp.447-473.

Gershenson S. & Tekin, E. (2018). The Effect of Community Traumatic Events on Student Achievement: Evidence from the Beltway Sniper Attacks. Education Finance and Policy, 13(4), pp.513-544.

Henry, G., Henderson L., Ponder Bentley D., Gordon C., Mashburn A. & Rickman, D. (2003). Report of the Findings from the Early Childhood Study: 2001-02.

Hornack, D. (2016). Impact of Summer Recess on Mathematics Learning Retention. Available at: https://files.eric.ed.gov/fulltext/EJ1116885.pdf.

Jaume, D. & Willén, A. (2018). The Long-Run Effects of Teacher Strikes: Evidence from Argentina. Research Briefs in Economic Policy. Number 123.

Lamb J., Lewis, M. & Gross, S. (2013). The Hurricane Katrina Effect on Mathematics Achievement in Mississippi. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/ssm.12003.

Lindahl, M. (2001). Summer Learning and the Effect of Schooling: Evidence from Sweden. IZA Discussion Paper No. 262. Available at: https://www.iza.org/publications/dp/262/summer-learning-and-the-effect-of-schooling-evidence-from-sweden.

Kerry, T. & Davies, B. (1998). Summer learning loss: The evidence and a possible solution. Support for Learning, 13(3), pp.118-122.

Kuhfeld M. (2019). Surprising new evidence on summer learning loss. Phi Delta Kappan, 101(1), pp.25-29.

McCoach B., O'Connell A., Reis S. & Levitt H. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. Journal of educational psychology, 98(1), pp.14-14.

*Meyer F. (2014). The summer learning effect in Germany: Home and school contributions to summer learning patterns in the communities of two primary schools.

*Meyer F., Meissel, K., & McNaughton, S. (2017). Patterns of literacy learning in German primary schools over the summer and the influence of home literacy practices. Journal of Research in Reading 40(3), pp.233-253.

*Paechter M., Luttenberger S., Macher, D., Berding, F., Papousek, I., Weiss E., & Fink, A. (2015). The Effects of Nine-Week Summer Vacation: Losses in Mathematics and Gains in Reading. Eurasia Journal of Mathematics Science And Technology Education, 11(6), pp.1399-1413.

*Quinn D., Cooc, N., McIntyre, J. & Gomez, C. (2016). Seasonal dynamics of academic achievement inequality by socioeconomic status and race/ethnicity: Updating and extending past research with new national data. Educational Researcher, 45(8), pp.443-453.

Ready, D. (2010). Socioeconomic Disadvantage, School Attendance, and Early Cognitive Development: The Differential Effects of School Exposure. Sociology of Education, 83(4), pp.271-286.

Sacerdote, B. (2008). When the Saints Come Marching In. Effects of Hurricanes Katrina and Rita on Student Evacuees. NBER Working Paper No. 14385.

*Verachtert, P., Van Damme, J., Onghena, P., & Ghesquiere, P. (2009). A seasonal perspective on school effectiveness: evidence from a Flemish longitudinal study in kindergarten and first grade. School Effectiveness and School Improvement, 20(2), pp.215-233.

*von Hippel, P. & Hamrock, C. (2019). Do Test Score Gaps Grow Before, During, or Between the School Years? Measurement Artifacts and What We Can Know in Spite of Them. Sociological Science, 6, pp.43-80.

Zajic, J. (2017). Summer Regression & Rural Students: The Effects of The Extended School Year on Mathematics & Reading for Elementary-Age Students with a Low-SES Background.

Zvoch, K. (2009). A Longitudinal Examination of the Academic Year and Summer Learning Rates of Full- and Half-Day Kindergartners.

Wider references

Abott, C., Kogan, V., Lavertu, S., Peskowitz, Z. (2020). School district operational spending and student outcomes: Evidence from tax elections in seven states. Journal of Public Economics 183: 1. https://doi.org/10.1016/j.jpubeco.2020.104142.

Davison, A. C., & Hinkley, D. V. (1997). Bootstrap methods and their application. Cambridge University Press.

Department for Education (2016). Absence and attainment at key stages 2 and 4: 2013 to 2014. Available at: https://www.gov.uk/government/publications/absence-and-attainment-at-key-stages-2-and-4-2013-to-2014.

Department for Education (2019a). "National curriculum assessments at key stage 2, 2019 (revised)". Available at: https://www.gov.uk/government/statistics/national-curriculum-assessments-key-stage-2-2019-revised.

Department for Education (2019b). Pupil absence in schools in England: 2017 to 2018. Available at: https://www.gov.uk/government/statistics/pupil-absence-in-schools-in-england-2017-to-2018.

Department for Education (2020). Pupil premium: conditions of grant 2019 to 2020. Available at: https://www.gov.uk/government/publications/pupil-premium-allocations-and-conditions-of-grant-2019-to-2020/pupil-premium-conditions-of-grant-2019-to-2020

Dragoset, L., Thomas, J., Herrmann, M, Deke, J., James-Burdumy, S. and Luca, D. (2019). "The Impact of School Improvement Grants on Student Outcomes: Findings From a National Evaluation Using a Regression Discontinuity Design." Journal of Research on Educational Effectiveness 12 (2): 215-50. https://doi.org/10.1080/19345747.2019.1571654.

Dynarski, M. (2017). It's not nothing: The role of money in improving education. Available at: https://www.brookings.edu/research/its-not-nothing-the-role-of-money-in-improving-education/

Douglas, M., Katikireddi, S., Taulbut, M., McKee, M., McCartney, G. (2020). Mitigating the wider health effects of covid-19 pandemic response. *BMJ*, 369.

Downey, D. B., von Hippel, P. T., & Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. American Sociological Review 69: 613-635. https://doi.org/10.1177/000312240406900501.

Education Endowment Foundation (2020a). Remote Learning, Rapid Evidence Assessment, London: Education Endowment Foundation. Available at: https://educationendowmentfoundation.org.uk/coronavirus-resources/.

Education Endowment Foundation (2020b). Teaching and Learning Toolkit. Available at: https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/.

Education Policy Institute (2019). Education in England: Annual Report 2019. Available at: https://epi.org.uk/publications-and-research/annual-report-2019/.

Goodman, J. (2014). Flaking Out: Student Absences and Snow Days as Disruptions of Instructional Time. NBER Working Paper No. 20221. Available at: https://www.nber.org/papers/w20221.

Gottfried, M. A. (2010). Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach. American Educational Research Journal 47(2), 434–465.

Gu., Q., Hodgen, J., Adkins, M. & Armstrong, P. (2019). Incentivising Schools to Take Up Evidence-Based Practice to Improve Teaching and Learning: Evidence from the Evaluation of the Suffolk Challenge Fund. Available at: https://educationendowmentfoundation.org.uk/scaling-up-evidence/campaigns/suffolk-challenge-fund/evaluation/.

Higgins, J. P., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 172*(1), 137-159. 10.1111/j.1467-985X.2008.00552.x

Institute for Fiscal Studies (2020). Learning during the lockdown: real-time data on children's experiences during home learning. Available at: https://www.ifs.org.uk/publications/14848.

Killip, S., Mahfoud, Z., & Pearce, K. (2004). What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Annals of Family Medicine, 2*(3), 204-208. 10.1370/afm.141

Kirabo Jackson, C. (2016). The Effects of School Spending on Educational and Economic Outcomes. The Quarterly Journal of Economics, 157–218. https://doi.org/10.1093/qje/qjv036.

Kirabo Jackson, C. (2018). Does School Spending Matter? The New Literature on an Old Question." Northwestern Mimeo. Available at: https://www.nber.org/papers/w25368.

Kreisman, D., and Steinberg, M. (2019). "The Effect of Increased Funding on Student Achievement: Evidence from Texas's Small District Adjustment." Working Paper. Available at: http://www.dkreisman.com/research/Kreisman_Texas.pdf.

Kuhfeld, M., & Soland, J. (2020). The learning curve: Revisiting the assumption of linear growth across the school year. (EdWorkingPaper: 20-214). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/bvg0-8g17.

Moher D, Liberati A, Tetzlaff J, Altman D. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097.

Office for National Statistics (2020). Statistical bulletin: Deaths involving COVID-19 by local area and socioeconomic deprivation. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19bylocalareasanddeprivation/deathsoccurringbetween1marchand17april.

Perline, R., Wright, B. D., & Wainer, H. (1979). The Rasch model as additive conjoint measurement. Applied Psychological Measurement, 3(2), 237-255.

Schools Week (2020). Oak National Academy delivers 2m lessons in a week. Available at: https://schoolsweek.co.uk/coronavirus-oak-national-academy-delivers-2m-lessons-in-first-week/

Sullivan, A., Moulton, V. & Fitzsimons, E. (2017). The intergenerational transmission of vocabulary. Centre for Longitudinal Studies Working Paper 2017/14. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2017/11/CLS-WP-201714-The-intergenerational-transmission-of-vocabulary.pdf.

Sutton Trust (2020). Coronavirus and Social Mobility Impact Brief #1: School Shutdown. Available at: https://www.suttontrust.com/our-research/coronavirus-and-social-mobility-impact-brief/.

Torgerson, C. et al. (2018). EEF Evaluation Report: The Tutor Trust. Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/tutor-trust-effectiveness-trial/.

Woodworth, J., Raymond, M., Chirbas, K., Gonzalez, M., Negassi, Y., Snow, W. & Van Donge, C. (2015). Online Charter School Study. Stanford: Centre for Research on Education Outcomes. Available at: https://credo.stanford.edu/sites/g/files/sbiybj6481/f/online_charter_study_final.pdf.