Rapid evidence assessment on online distance learning for school aged pupils
Protocol for a rapid evidence assessment
Principal investigator(s): Steve Higgins, Jennifer Stevenson, Jonathan Kay, Amy Ellis-Thompson, Mohammad Zaman

Table of contents
Table of contents .................................................................................................................. 1
Background and review rationale ......................................................................................... 1
Objectives .............................................................................................................................. 1
Methodology .......................................................................................................................... 2
Reporting ................................................................................................................................. 8
Peer review ............................................................................................................................. 8
Personnel ................................................................................................................................. 8
Conflicts of interest ............................................................................................................... 9
Timeline ................................................................................................................................ 9
References ............................................................................................................................... 10

Background and review rationale
The COVID19 pandemic has led to school closures across the UK and many countries across the world. This means that the majority of pupils in these systems are taught through various distance and online learning techniques and approaches.

There is an urgent need for an accessible overview of the evidence on effective distance learning and online learning practices and information on how to improve learning for pupils that are not able to attend classrooms. This review aims to be the basis for an accessible report that gives an overview of the efficacy of online distance learning practices.

There are existing meta-analyses and systematic reviews on the different approaches to online distance learning such as synchronous online learning (Martin et. al. 2017), and general reviews of online education (Sun et. al. 2016)

The urgency of the pandemic means that this review needs to be conducted quickly. The review aims to summarise the evidence on online distance learning through a rapid evidence assessment of systematic reviews and meta-analyses on the topic.

Objectives
The aim of the review is to summarise the efficacy of distance learning approaches compared to within school learning and, where evidence is available, find which distance learning approaches are most likely to improve pupil outcomes. We will also look for evidence on the characteristics of effective implementation of distance learning. Our aim is to produce a school facing publication summarising the findings of the rapid evidence assessment (REA) within six weeks of starting the review.

Research questions:

1. How effective are distance learning and online learning approaches in comparison to usual schooling?
2. Does pupil or school level disadvantage moderate pupil achievement outcomes in distance learning approaches?
3. How do different distance learning approaches moderate pupil achievement outcomes?
4. What are the characteristics of effective distance learning implementation?
5. Which EEF-funded programmes that could be or have been delivered remotely have shown promise in terms of improving pupil achievement compared to business as usual?

**Methodology**

We will conduct a rapid evidence assessment of existing systematic reviews and meta-analyses. The Cochrane Collaboration Rapid Reviews Methods Group has recently published interim guidance on producing rapid reviews, motivated by the COVID-19 pandemic and the need to provide answers to relevant time-sensitive questions (Garrity et al. 2020). We will draw on this methodological guidance for this REA, as well as the Civil Service REA methodological guidance (Government Social Research Service, 2009) and the Cochrane Collaboration’s guidance on overviews of reviews (Pollock et al. 2020). The scope has been limited to systematic review and meta-analyses in order to respond quickly to the policy challenge of COVID19. The review is also limited to studies from 2005 or later. Technology based approaches from pre-2005 are unlikely to be relevant to current online learning approaches.

**Inclusion and exclusion criteria for the review**

<table>
<thead>
<tr>
<th>Include</th>
<th>Exclude</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
<td>Primary or secondary aged pupils. Reviews of studies from any country will be included.</td>
</tr>
<tr>
<td><strong>Interventions</strong></td>
<td>We define distance learning approaches as methods of teaching that take place entirely outside of the classroom environment. They can be synchronous or asynchronous. The primary objective needs to be pupil learning outcomes. Examples of included approaches are: - Recreating the classroom environment through online platforms such as “Google Classrooms” or equivalents - Online tutoring that provides intensive support to individual pupils through a two-way link - Digital platforms/education software that are used independently by pupils</td>
</tr>
</tbody>
</table>

¹ Reviews focusing on students in higher education will be labelled during the screening process and reviewed if limited evidence is found for primary or secondary aged pupils.
Rapid evidence assessment on online distance learning for school aged pupils  
Protocol for a rapid evidence assessment  
Principal investigator(s): Steve Higgins, Jennifer Stevenson, Jonathan Kay, Amy Ellis-Thompson, Mohammad Zaman

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Classroom or school-based learning(^3); other types of distance / online based approaches.</th>
<th>Pure control (that is, compared to no learning activities)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome</td>
<td>Reviews covering any measure of academic achievement or cognitive measure of ability of pupils in any subject</td>
<td>Reviews that only examine behaviour, attendance, or other non-cognitive outcomes. Reviews that only focus on teacher outcomes.</td>
</tr>
<tr>
<td>Study design</td>
<td>Meta-analyses or systematic reviews of distance learning effectiveness or implementation of distance learning</td>
<td>Single studies, narrative reviews. Systematic reviews addressing research questions other than effectiveness or implementation</td>
</tr>
<tr>
<td>Other criteria</td>
<td>Published since 2005. Published in English. Reviews published in peer-reviewed journals or grey literature.</td>
<td>Published before 2005. Published in languages other than English.</td>
</tr>
</tbody>
</table>

**Study designs**

We will include any review that identifies as a systematic review or meta-analysis, if they also describe methods used for the search, data collection and synthesis. A systematic review is the process of searching for and selecting evidence using pre-specified criteria, appraising and synthesising it and

---

\(^2\) We will use the same classification as Acquah et al (2020), Hainey et al. (2016) and others, which distinguishes between games for learning (GL) and entertainment games (EG). EGs as “pre-made, COTS [commercial off-the-shelf] games that are used in [school] for the purposes of learning, teaching a particular subject or promoting engagement,” while, “games for learning ... is the production of a specially implemented application for the purposes of learning, teaching a particular subject of promoting engagement,” (Hainey et al., 2016, p. 203).

---

\(^3\) In practice, it may be difficult to apply this as an inclusion / exclusion criteria as systematic reviews and meta-analysis may not specify the comparison or may include a mix of both types of comparison. In those cases where it is difficult to apply this criterion, but if a review meets all other criteria, it should be included.
reaching conclusions about the body of evidence to answer a specific research question. Meta-analysis is a statistical method for combing the results from multiple studies. We will include systematic reviews and meta-analyses that address the effectiveness of distance learning or systematic reviews that address the barriers and / or facilitators to effective distance learning implementation.

**Population**
We will initially include only systematic reviews and meta-analyses that cover distance learning for school-aged pupils. However; we are also aware that the literature on this topic may be limited. If we identify few reviews that cover distance learning for school aged pupils, we will consider drawing on systematic reviews and meta-analyses covering distance learning at higher levels of education and assess whether there is relevant information on effectiveness or implementation of distance learning for schools. We will identify these reviews during the screening process, see more below, as they will be picked up in our search and revisit them depending on the extent of the synthesis literature that we identify on school distance learning.

**Search strategy for identification of studies**

**Systematic reviews and meta-analyses**
Searches will be conducted using a combination of search systems and bibliographic databases, including Web of science, Microsoft Academic and ERIC, and hand searches of known sources of systematic reviews such as the Campbell Library. We will also screen studies for inclusion from two existing umbrella reviews in digital technology, the Teaching and Learning Toolkit review of Digital Technology (Education Endowment Foundation, 2018) and ‘Using Digital Technology to Improve Learning: Evidence Review’ (Lewin et. al 2019).

**Search Systems and databases to be searched:**
- Web of science
- ERIC
- Google Scholar

**Other sources:**
- EPPI Centre library of reviews: https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=62
- Open Science Framework: https://osf.io/registries?view_only=

---

Google scholar has a 256 character limit and does not automatically searches for truncations. A more limited search string will be used for the google scholar search. The search will then be filtered to limit the results to studies that are published since 2005. We will look at the first 200 results in Google Scholar, in line with the recommendation of Haddaway et al. 2015.

**Adapted search:** ("Distance learning" | "distance education" | "remote learning" | "blended learning" | "hybrid learning" | e-learning | "Internet-based learning" | "intelligent tutoring" | "virtual learning environment") ("systematic review" OR meta-analysis)
Rapid evidence assessment on online distance learning for school aged pupils
Protocol for a rapid evidence assessment
Principal investigator(s): Steve Higgins, Jennifer Stevenson, Jonathan Kay, Amy Ellis-Thompson, Mohammad Zaman

- 3ie database of systematic reviews: https://developmentevidence.3ieimpact.org/
- Review of Education Research: https://journals.sagepub.com/home/rer
- Education Research Review: https://www.journals.elsevier.com/education-research-review

Once we have screened the search results from the databases above and have a set of included studies, we will also use Microsoft Academic to identify similar studies that might have been missed by the main search. Microsoft Academic is a large open access repository containing more than 228 million records. We will access it through the EPPI-Reviewer 4 software and use the EPPI Reviewer user guide. This process is conceptually similar to forward and backwards citation tracking.

Search terms:

We have drawn on the search terms used in the digital technology evidence review (Lewin et al., 2019) and combined with new search terms to cover the intervention areas not covered by that review. The terms will be used to search on titles and abstracts and adapted as necessary depending on the search functions of the search systems and databases. Where it is possible to refine searches using filters such as categories on web of science, we will exclude categories that are not related to education, digital technology or communications. Where filters on sites correspond to inclusion criteria we will also filter during the search – for example, only searching studies published in 2005.

(“Distance learning” OR “distance education” OR “remote learning” OR “blended learning” OR “hybrid learning” OR e-learning OR “Internet-based learning” OR “internet-based education” OR “mobile learning” OR m-learning OR “online learning” OR “online education” OR “online tutor” OR “virtual schools” OR “virtual learning” OR “virtual learning platforms” OR “web-based learning” OR WebQuests OR “digital learning resource” OR “virtual learning environment” OR VLE OR "learning platform" OR online quiz* OR digital quiz* OR "computer supported collaborative learning" OR CSCL OR “technology-supported collaborative learning” OR “e-textbook” OR "technology enhanced learning" OR "TEL" OR "computer-based teaching" OR "computer-based learning" OR "computer-based instruction" OR "computer-assisted training" OR "computer-assisted teaching" OR "computer-assisted learning" OR "computer-assisted instruction" OR "intelligent tutoring systems" OR "intelligent tutoring system" OR "computer game*" OR "computer gaming") AND ((systematic* or synthes*) adj (narrative or meta* or review* or literature or evidence or quantitative or study or studies or paper or impact or impacts or effect* or compar*)) OR ("meta regression" or "meta synth*" or "meta-synth*" or "meta analy*" or "metanaly*" or "meta-analy*" or "metanalysis*" or "metaregression")

Where databases allow Boolean operators will be used to exclude any studies with: “medical” or “health” included. Many of the studies in the existing evidence focus on e-learning for medical practitioners.

Primary studies
All EEF funded studies will be reviewed for inclusion in the school facing publication as examples of distance learning approaches that have been implemented within English schools. We will not include

1 More information is available at: http://eppi.ioe.ac.uk/CMS/Portals/35/MAG%20Browser%20v_1_0_User%20Guide.pdf
these in the main synthesis but will present these in the school facing publication as examples of programmes that schools may wish to consider.

Selection of studies
The results of the search will be imported into EPPI reviewer and duplicates removed. Each search result will be screened twice, first on abstract and title only, then on the full text. After initial calibration, each screening stage will be completed by one reviewer only due to the timeline for this project. However, we will take a “safety first” approach at both screening stages (Shemilt et al., 2016); that is, the reviewer will have the option of marking a search result as unclear for review by a second reviewer.

At the title and abstract stage, every reviewer will begin by screening the same 30 search results. The results of this screening will be compared to ensure that the inclusion and exclusion criteria are being interpreted and applied in the same way. The priority screening tool within EPPI-reviewer (Thomas et al., 2010) will be used for title and abstract screening to order results by probability of inclusion and stop screening once we reach a certain point when relevant studies are no longer being identified. The priority screening function orders the results based on the words in the title and abstract of the included and excluded papers from a training set of screening. It does this using machine learning text mining technology. We will screen a random set of 10 percent of the search results as the training set. Reviewers will stop screening after 100 studies are rejected in a row using the tool. As a check on this approach, we will randomly sample a number of the unscreened titles to see if this approach has missed any relevant studies.

The results of this process will be documented using a flow chart generated from EPPI-reviewer.

Data extraction and management
We will systematically extract data in Microsoft Excel using the templates included in appendix A. We will extract descriptive data about the type of review, broad intervention area and delivery and pedagogical features; outcomes covered by the review; description of individual effect sizes or results from meta-analysis and any information about effective implementation of distance learning. A team will be responsible for extracting information from the included studies using the data extraction tool. The core team will do double data extraction on 10% of studies (randomly selected).

Appraisal of included systematic reviews
We will undertake a critical appraisal of each of the included systematic reviews and meta-analysis in the REA, to make a judgement about the quality of the systematic review process of each. We will use an adapted version of the SURE checklist for how much confidence to place in a systematic review of effectiveness, drawing primarily on the version published in Snilstveit et al. (2014). The full checklist including decision rules, is included in Appendix B. The checklist is split into three sections:

- Methods used to identify, include and critically appraise studies
  - Were the criteria used for deciding which studies to include in the review reported?
  - Was the search for evidence reasonably comprehensive?
Rapid evidence assessment on online distance learning for school aged pupils
Protocol for a rapid evidence assessment
Principal investigator(s): Steve Higgins, Jennifer Stevenson, Jonathan Kay, Amy Ellis-Thompson, Mohammad Zaman

- Was bias in the selection of articles avoided?
- Did the authors use appropriate criteria to assess the quality and risk of bias in analysing the studies that are included?
- Overall – how much confidence do you have in the methods used to identify, include and critically appraise studies?

• Methods used to analyse the findings
  - Were the characteristics and results of the included studies reliably reported?
  - Are the methods used by the review authors to analyse the findings of the included studies clear, including methods for calculating effect sizes if applicable?
  - Did the review describe the extent of heterogeneity?
  - Were the findings of the relevant studies combined (or not combined) appropriately relative to the primary question the review addresses and the available data?
  - Does the review report evidence appropriately?
  - Did the review examine the extent to which specific factors might explain differences in the results of the included studies?
  - Overall - how much confidence do you have in the methods used to analyse the findings relative to the primary question addressed in the review?

• Overall assessment of the reliability of the review

Each study will receive an overall assessment of low, medium or high confidence, representing the quality of the systematic review process. This is distinct from the process of rating certainty in the effect estimates, for example GRADE assessments, which we are unable to do in this review due to the limited timeframe of the project. We will use the results of this critical appraisal to be more tentative with conclusions when the systematic review evidence base is weak. Depending on the number of reviews that we identify, we may also use the critical appraisal results to conduct sensitivity analysis to consider the potential impact of reviews of lower quality on overall conclusions. This will specifically be where included systematic reviews answer the same research questions.

The critical appraisal of each included review will be completed by one reviewer and checked by another. The full final REA report will include a table that provides a breakdown of how each systematic review was rated on each question of the tool and the overall confidence rating.

Data synthesis
Questions 1 - 3
We will undertake a narrative synthesis of the included systematic reviews and meta-analyses to answer review questions 1, 2 and 3, presenting pooled and individual effect sizes and associated measures of uncertainty where presented in the original reviews. We will not compute pooled estimates for the impact of distance learning approaches, that is, undertake meta-meta-analysis. The methodological limitations of synthesising meta-analyses and the expected heterogeneity of distance learning approaches make the calculation of an effect size inappropriate for this study. We will also not re-analyse the underlying studies included in the systematic reviews and meta-analyses that we
identify due to time restraints. The review will instead summarise the findings from each of the included reviews and will present implications to current practice in distance learning.

To answer research question 2 on disadvantage, we will report any sub-group analysis or meta-regression results reported in the original systematic reviews by any indicator of pupil or school level socio-economic disadvantage.

To answer research question 3, about how different distance learning approaches moderate outcomes, we will only use comparative results reported within an included systematic review. That is, we will not directly make comparisons of the relative effectiveness of different distance learning approaches reported in different systematic reviews, given that the studies included across different reviews will most likely have had different comparisons (Pollock et al., 2020).

**Question 4**
To answer research question 4, we will extract both qualitative and quantitative information where available from the included systematic reviews on barriers and facilitators to effective implementation of distance learning approaches. We will use thematic synthesis to synthesise the results, following Thomas and Harden (2008).

**Reporting**
The technical report will use the EEF review reporting template for evidence reviews.

A school facing publication will summarise the evidence for different distance learning practices, describing the impact and implementation challenges of different learning approaches, e.g. online tutoring, online quiz platforms (for full list of approaches see data extraction tool).

**Peer review**
Due to the limited timeframe of this REA, this protocol was not formally peer reviewed.

**Personnel**

**Core team:**
- Steve Higgins – Durham University - synthesis lead
- Jennifer Stevenson – EEF - search strategy and quality assessment lead
- Jonathan Kay – EEF - screening and retrieval lead
- Amy Ellis Thompson – EEF - publication lead
- Mohammad Zaman – EEF - project manager

**Data extraction and screening team:**
- Igraine Rhodes - EEF
- Shelby Roberts - EEF
- Trent Grassian – EEF
Rapid evidence assessment on online distance learning for school aged pupils

Protocol for a rapid evidence assessment

Principal investigator(s): Steve Higgins, Jennifer Stevenson, Jonathan Kay, Amy Ellis-Thompson, Mohammad Zaman

- Sue Morgan - EEF
- Liberty King - EEF
- Pauline Brown - EEF
- Harry Madgwick – EEF

Support from Durham University team:
- Jade Stafford
- Berenice
- Emma Dobson

Conflicts of interest
No conflicts of interest

Timeline

<table>
<thead>
<tr>
<th>Task</th>
<th>Deadline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protocol development</strong></td>
<td></td>
</tr>
<tr>
<td>Protocol finalised</td>
<td>31st March</td>
</tr>
<tr>
<td>School facing report template finalised</td>
<td>3rd April</td>
</tr>
<tr>
<td>Data extraction tool finalised</td>
<td>31st March</td>
</tr>
<tr>
<td><strong>Search</strong></td>
<td></td>
</tr>
<tr>
<td>Academic search</td>
<td>3rd April</td>
</tr>
<tr>
<td>Search grey literature and organisational websites</td>
<td>3rd April</td>
</tr>
<tr>
<td>Identify relevant EEF studies</td>
<td>3rd April</td>
</tr>
<tr>
<td>Citation tracking (checking included studies in the digital feedback review)</td>
<td>3rd April</td>
</tr>
<tr>
<td><strong>Screening</strong></td>
<td></td>
</tr>
<tr>
<td>Screening at title (and abstract)</td>
<td>8th April</td>
</tr>
<tr>
<td>Full text retrieval (assuming xx papers)</td>
<td>8th April</td>
</tr>
<tr>
<td>Full text screening</td>
<td>9th April</td>
</tr>
<tr>
<td><strong>Data Extraction</strong></td>
<td></td>
</tr>
<tr>
<td>Data extraction (decriptive, intervention, implementation, effect sizes), assuming xx per day</td>
<td>14th April</td>
</tr>
<tr>
<td>Check data extraction</td>
<td>15th April</td>
</tr>
<tr>
<td>Data extraction (critical appraisal of reviews)</td>
<td>15th April</td>
</tr>
<tr>
<td><strong>Synthesis and write-up</strong></td>
<td></td>
</tr>
<tr>
<td>Narrative synthesis</td>
<td>16th April</td>
</tr>
<tr>
<td>Write up of new synthesis</td>
<td>17th April</td>
</tr>
<tr>
<td>Write up of relevant EEF studies</td>
<td>17th April</td>
</tr>
<tr>
<td>Compilation of school facing findings</td>
<td>17th April</td>
</tr>
<tr>
<td>Review of findings document by D&amp;I, Becky, Stephen, Robbie (who else?)</td>
<td>20th April</td>
</tr>
<tr>
<td>Publish on EEF website</td>
<td>22nd April</td>
</tr>
<tr>
<td>Write up draft technical report using REA template</td>
<td>TBC</td>
</tr>
</tbody>
</table>
References

https://doi.org/10.1016/j.compedu.2019.103667

https://educationendowmentfoundation.org.uk/resources/teaching-learning-toolkit


Government Social Research Service (2009). GSR Rapid Evidence Assessment Toolkit (Online),


Appendix A: Data extraction tool

- Intervention name (open response)
  *Some studies will concern specific named interventions.*
- Technology used (open response)
  *For example, iPad, phone, computer, variable*
- Who delivers the distance learning?
  - Teacher
  - Tutor
  - Automated
  - Other (describe)
- Overall duration of distance learning (number of days/months/years)
- Is it new learning or consolidating existing learning? (select one)
  - New learning
  - Consolidating existing learning
- How do interactions take place? (select one)
  - Not interactive - independent use
  - Interactive - sequential/asynchronous
  - Interactive - live/synchronous
  - Mixed
- Age range (open response)
  *In pupil ages rather than school years*
- Country (country name)
  *Include if most or all of the studies are from one country*
- Other context notes (open response)
  *Is the review limited to a specific population? E.g. rural schools? Disadvantaged pupils?*
- Pooled effect (numerical value)
- Effect size type (open response)
- Outcome measure (open response)
- What is the comparison (select one)?
  - In-school teaching
  - Other distance learning intervention (describe)
  - Other (describe)
- Standard error (numerical value)
- Standard deviation (numerical value)
- Confidence interval lower (numerical value)
- Confidence interval higher (numerical value)
- Minimum effect size (numerical value)
- Maximum effect size (numerical value)
- Moderators (open response)
  *List any moderator analysis included in the review*
- Number of pupils (numerical value)
- Has the review searched grey literature? (select one)
  - Yes
Protocol for a rapid evidence assessment

Principal investigator(s): Steve Higgins, Jennifer Stevenson, Jonathan Kay, Amy Ellis-Thompson, Mohammad Zaman

- Is the review limited to randomised controlled trials? (select one)
  - Yes
  - No

- Number of effects (numerical value)
- Number of studies (numerical value)
- Does the report say anything about disadvantage? (open response)
- Does the report say anything about implementation? (open response)
- Other notes (open response)
Appendix B: Critical appraisal checklist

**Section A: Methods used to identify, include and critically appraise studies**

| Were the criteria used for deciding which studies to include in the review reported? | □ Yes | □ Partially | □ No |
| Did the authors specify: | | | |
| ☐ Types of studies | | | |
| ☐ Participants/ settings/ population | | | |
| ☐ Intervention(s) | | | |
| ☐ Outcome(s) | | | |

*Coding guide - check the answers above:

**YES:** All four should be yes

**NO:** All four should be no

**PARTIALLY:** Any other

| Was the search for evidence reasonably comprehensive? | □ Yes | □ Partially | □ No | □ Can’t tell |
| Were the following done: | | | | |
| ☐ No restriction of inclusion based on publication status | | | | |
| ☐ Relevant databases searched (Minimum criteria: All reviews should search at least one source of grey literature such as Google; at least one database of general social science literature and one subject specific database) | | | | |
| ☐ Reference lists in included articles checked | | | | |

*Coding guide - check the answers above:

**YES:** All should be yes

**PARTIALLY:** Relevant databases and reference lists are both reported

**NO:** Any other
| Was bias in the selection of articles avoided? | □ Yes □ Partially □ No  
Coding guide: 
YES: All should be yes 
PARTIALLY: Independent screening not done 
NO: All other. If list of included studies provided, but the authors do not report whether or not the screening has been done by 2 reviewers review is downgraded to NO. |
|---------------------------------------------|-------------------------------------------------|
| Did the authors use appropriate criteria to assess the quality and risk of bias in analysing the studies that are included? | □ Yes □ Partially □ No  
Coding guide: 
YES: All three should be yes 
PARTIALLY: The first and third criteria should be reported. If the authors report the criteria for assessing risk of bias and report a summary of this assessment for each criterion, but the criteria may be only partially sensible (e.g. do not address all possible risks of bias, but do address some), we downgrade to PARTIALLY. 
NO: Any other |
| --------------------------------------------|-------------------------------------------------|
| Did the authors specify: | □ Independent screening of full text by at least 2 reviewers or single screening with at least a 10% proportion of double screening to align screeners  
□ List of included studies provided |
| --------------------------------------------|-------------------------------------------------|
| Coding guide: | YES: All should be yes 
PARTIALLY: Independent screening not done |

**Was bias in the selection of articles avoided?**

- Did the authors specify:
  - Independent screening of full text by at least 2 reviewers or single screening with at least a 10% proportion of double screening to align screeners
  - List of included studies provided

**Coding guide:**

- YES: All should be yes
- PARTIALLY: Independent screening not done
- NO: All other. If list of included studies provided, but the authors do not report whether or not the screening has been done by 2 reviewers review is downgraded to NO.

**Did the authors use appropriate criteria to assess the quality and risk of bias in analysing the studies that are included?**

- The criteria used for assessing the quality/ risk of bias were reported
- A table or summary of the assessment of each included study for each criterion was reported
- Sensible criteria were used that focus on the quality/ risk of bias (and not other qualities of the studies, such as precision or applicability/external validity). “Sensible” is defined as a recognised quality appraisal tool/ checklist, or similar tool which assesses bias in included studies. Please see footnotes for details of the main types of bias such a tool should assess.

**Coding guide:**

- YES: All three should be yes
- PARTIALLY: The first and third criteria should be reported. If the authors report the criteria for assessing risk of bias and report a summary of this assessment for each criterion, but the criteria may be only partially sensible (e.g. do not address all possible risks of bias, but do address some), we downgrade to PARTIALLY.
- NO: Any other
A. Overall – how much confidence do you have in the methods used to identify, include and critically appraise studies?

Summary assessment score A relates to the 5 questions above.

High confidence applicable when the answers to the questions in section A are all assessed as ‘yes’

Low confidence applicable when any of the following are assessed as ‘NO’ above: not reporting explicit selection criteria (A1), not conducting reasonably comprehensive search (A2), not avoiding bias in selection of articles (A4), not assessing the risk of bias in included studies (A5)

Medium confidence applicable for any other – i.e. section A3 is assessed as ‘NO’ or can’t tell and remaining sections are assessed as ‘partially’ or ‘can’t tell’.

<table>
<thead>
<tr>
<th>Yes</th>
<th>Low confidence (limitations are important enough that the results of the review are not reliable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Medium confidence (limitations are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously, if a better review cannot be found)</td>
</tr>
<tr>
<td>Partially</td>
<td>High confidence (only minor limitations)</td>
</tr>
</tbody>
</table>

Section B: Methods used to analyse the findings

B. Were the characteristics and results of the included studies reliably reported?

Was there:

- Independent data extraction by at least 2 reviewers or single data extraction with at least a 10% proportion of studies with independent data extraction
- A table or summary of the characteristics of the participants, interventions and outcomes for the included studies
- A table or summary of the results of all the included studies

Coding guide:

YES: All three should be yes

PARTIALLY: Criteria one and three are yes, but some information is lacking on second criteria.

No: None of these are reported. If the review does not report whether data was independently extracted by 2 reviewers (possibly a reporting error), we downgrade to NO.

NOT APPLICABLE: if no studies/no data
Rapid evidence assessment on online distance learning for school aged pupils
Protocol for a rapid evidence assessment
Principal investigator(s): Steve Higgins, Jennifer Stevenson, Jonathan Kay, Amy Ellis-Thompson, Mohammad Zaman

### B.2 Are the methods used by the review authors to analyse the findings of the included studies clear, including methods for calculating effect sizes if applicable?

- [ ] Yes
- [ ] Partially
- [ ] No
- [ ] Not applicable (e.g. no studies or no data)

**Coding guide:**

**YES:** Methods used clearly reported. If it is clear that the authors use narrative synthesis, they don’t need to say this explicitly.

**PARTIALLY:** Some reporting on methods but lack of clarity

**NO:** Nothing reported on methods

**NOT APPLICABLE:** if no studies/no data

### B.3 Did the review describe the extent of heterogeneity?

- [ ] Did the review ensure that included studies were similar enough that it made sense to combine them, sensibly divide the included studies into homogeneous groups, or sensibly conclude that it did not make sense to combine or group the included studies?

- [ ] Did the review discuss the extent to which there were important differences in the results of the included studies?

- [ ] If a meta-analysis was done, was the $I^2$, chi square test for heterogeneity or other appropriate statistic reported? If no statistical test was reported, is a qualitative justification made for the use of random effects?

- [ ] Yes
- [ ] Partially
- [ ] No
- [ ] Not applicable (e.g. no studies or no data)

**Coding guide:**

**YES:** First two should be yes, and third category should be yes if applicable should be yes

**PARTIALLY:** The first category is yes

**NO:** Any other

**NOT APPLICABLE:** if no studies/no data
B.4 Were the findings of the relevant studies combined (or not combined) appropriately relative to the primary question the review addresses and the available data?

How was the data analysis done?

- Descriptive only
- Vote counting based on direction of effect
- Vote counting based on statistical significance
- Description of range of effect sizes
- Meta-analysis
- Meta-regression
- Other: specify
- Not applicable (e.g. no studies or no data)

Coding guide:

**YES:** If appropriate table, graph or meta-analysis AND appropriate weights (if appropriate).

**NO:** If narrative OR vote counting (where quantitative analyses would have been possible) OR inappropriate reporting of table, graph or meta-analyses.

**NOT APPLICABLE:** if no studies/no data

**CAN’T TELL:** if unsure (note reasons in comments below)

How were the studies weighted in the analysis?

- Equal weights (this is what is done when vote counting is used)
- By quality or study design (this is rarely done)
- Inverse variance (this is what is typically done in a meta-analysis)
- Number of participants (sample size)
- Other: specify
- Not clear
- Not applicable (e.g. no studies or no data)
B. 5 Does the review report evidence appropriately?

☐ The review makes clear which evidence is subject to low risk of bias in assessing causality (attribution of outcomes to intervention), and which is likely to be biased, and does so appropriately

☐ Where studies of differing risk of bias are included, results are reported and analysed separately by risk of bias status

☐ Yes

☐ No

☐ Partially

☐ Not applicable

Coding guide:

YES: Both criteria should be fulfilled (where applicable)

NO: Criteria not fulfilled

PARTIALLY: Only one criteria fulfilled, or when there is limited reporting of quality appraisal (the latter applies only when inclusion criteria for study design are appropriate)

NOT APPLICABLE: No included studies

Note on reporting evidence and risk of bias: For reviews of effects of ‘large n’ interventions, experimental and quasi-experimental designs should be included (if available). For reviews of effects of ‘small n’ interventions, designs appropriate to attribute changes to the intervention should be included (e.g. pre-post with assessment of confounders)
### B.6 Did the review examine the extent to which specific factors might explain differences in the results of the included studies?

- Yes
- Partially
- No
- Not applicable

**Coding guide:**

- **YES:** Explanatory factors clearly described and appropriate methods used to explore heterogeneity
- **PARTIALLY:** Explanatory factors described but for meta-analyses, sub-group analysis or meta-regression not reported (when they should have been)
- **NO:** No description or analysis of likely explanatory factors
- **NOT APPLICABLE:** e.g. too few studies, no important differences in the results of the included studies, or the included studies were so dissimilar that it would not make sense to explore heterogeneity of the results

### B. Overall - how much confidence do you have in the methods used to analyse the findings relative to the primary question addressed in the review?

*Summary assessment score B relates to the 5 questions in this section, regarding the analysis.*

- **Low confidence** (limitations are important enough that the results of the review are not reliable)
- **Medium confidence** (limitations are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously, if a better review cannot be found)
- **High confidence** (only minor limitations)

*Use comments to specify if relevant, to flag uncertainty or need for discussion*
### Section C: Overall assessment of the reliability of the review

<table>
<thead>
<tr>
<th>C.1 Are there any other aspects of the review not mentioned before which lead you to question the results?</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Additional methodological concerns – only one person reviewing</td>
</tr>
<tr>
<td>- Robustness</td>
</tr>
<tr>
<td>- Interpretation</td>
</tr>
<tr>
<td>- Conflicts of interest (of the review authors or for included studies)</td>
</tr>
<tr>
<td>- Other</td>
</tr>
<tr>
<td>- No other quality issues identified</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C.2 Are there any mitigating factors which should be taken into account in determining the reviews reliability?</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Limitations acknowledged</td>
</tr>
<tr>
<td>- No strong policy conclusions drawn (including in abstract/ summary)</td>
</tr>
<tr>
<td>- Any other factors</td>
</tr>
</tbody>
</table>

Use comments to specify if relevant, to flag uncertainty or need for discussion

<table>
<thead>
<tr>
<th>C.3 Based on the above assessments of the methods how would you rate the reliability of the review?</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Low confidence in conclusions about effects:</td>
</tr>
<tr>
<td>- Medium confidence in conclusions about effects:</td>
</tr>
<tr>
<td>- High confidence in conclusions about effects:</td>
</tr>
</tbody>
</table>

If applicable: The review has the following minor limitations...

**Coding guide:**

**High confidence in conclusions about effects:** high confidence noted overall for sections A and B, unless moderated by answer to C1.

**Medium confidence in conclusions about effects:** medium confidence noted overall for sections A or B, unless moderated by answer to C1 or C2.

**Low confidence in conclusions about effects:** low confidence noted overall for sections A or B, unless moderated by answer to C1 or C2.

Limitations should be summarised above, based on what was noted in Sections A, B and C.

---

1. **Risk of bias** is the extent to which bias may be responsible for the findings of a study.