Education Endowment Foundation

Table of contents

Background and review rationale

When schools reopen in September it is likely that some staff and pupils will stay away from school due to health concerns. The need for distancing may pose challenges to delivering professional development within school. Staff may be also unable to attend traditional face-to-face training outside school as national or local guidance may restrict movement, staff may need to limit contact due to personal health concerns or schools may require teachers to be present in school in order to re-establish routines, support ongoing home learning or provide catch-up support to help to close the gaps that are likely to have emerged between the amounts different children have learnt.

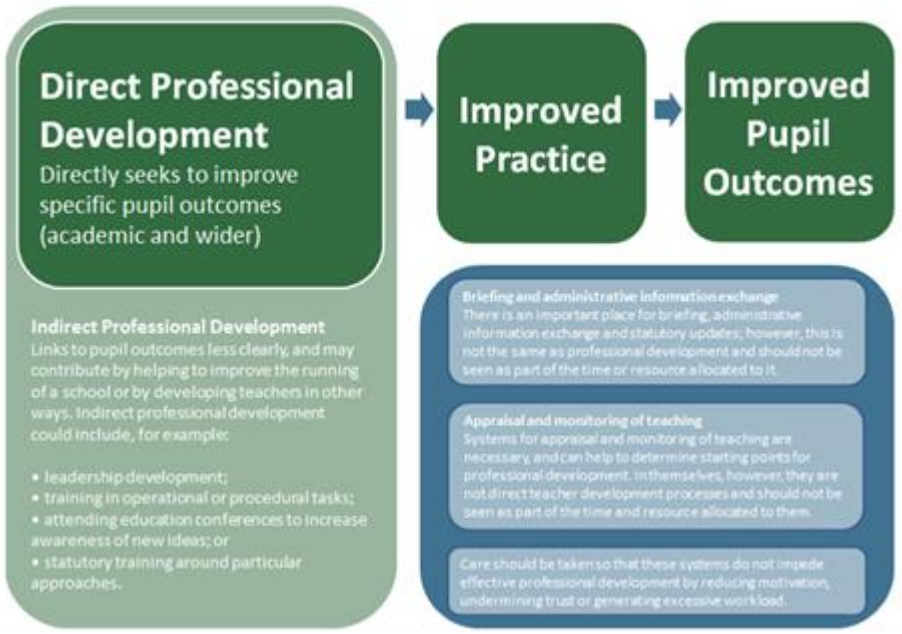
Schools and training providers, including EEF Research Schools, Promising Projects and other trial providers, are considering ways in which they can use distance learning techniques to continue providing professional development for teachers and other members of staff. There is no evidence review of remote professional development in this context for them to draw upon. There is an urgent need for evidence on:

- The qualities of successful distance learning techniques for professionals
- How to successfully implement distance learning to facilitate adult professional development
- How teachers and other school staff learn well when some or all training is remote

Ideally a review of this scope would be conducted using an exhaustive systematic review. Given the urgent need for evidence now, a rapid evidence assessment seems to be the most pragmatic solution as with two other recent reviews on remote learning (EEF, 2020a) and the attainment gap (EEF, 2020b). We will limit the scope of the review to existing systematic reviews and meta-analysis on the topic. We suggest remote CPD is also included in an upcoming systematic review of professional development that the EEF is commissioning.

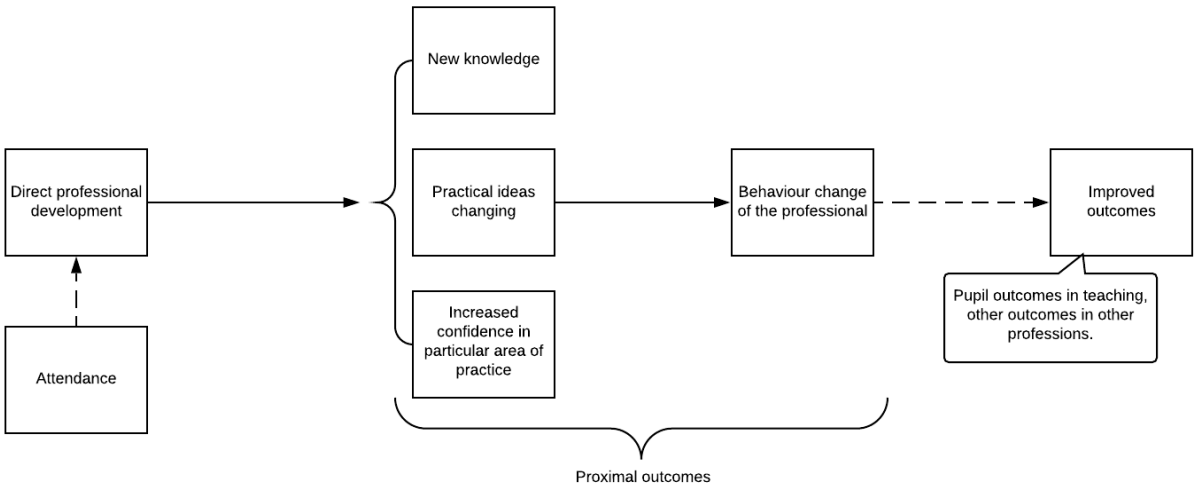
For the purposes of this rapid evidence assessment, the scope is limited to 'direct professional development' as defined by the DfE above and thus discounts simple procedural training (that does not aim to change pedagogical approaches) as well as leadership programmes.

Figure 1. DfE's Standard for Teachers' Professional Development: Definition



We produced a simple theory of change in order to inform our inclusion criteria and extract the key elements of professional learning relevant for our purposes. We will include studies of remote learning for professionals in education, welfare and public health as we think there may be lessons that can be learned from other sectors, but will code the studies by sector in order to identify e those relating to education professionals. This will allow us to comment on any differences between the broader set of studies and those relating specifically to schools and the teaching profession. Where studies are of remote learning outside of an education setting, 'pupil outcomes' are replaced with relevant service beneficiary outcomes, for example a client in psychology or the general public in public health.

Figure 2: theory of change



While pupil (or equivalent service user) outcomes are the ultimate goal, professional development success can be measured through two proximal goals: a change in knowledge of the professional and a change in their behaviour.

Objectives

The aim of the review is to summarise the efficacy of remote professional development approaches compared to face-to-face professional development or other types of remote professional development, and, where evidence is available, find which remote professional development approaches are most likely to improve pupil outcomes. We will also look for evidence on the characteristics of effective implementation of remote professional development for professionals. 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. What evidence is available on the effectiveness of remote professional development approaches, compared to face-to-face, blended approaches and other remote approaches?
2. Does some face-to-face contact moderate outcomes in remote professional development approaches? How do different remote professional development approaches moderate outcomes?
3. What can we learn from existing systematic reviews and meta-analyses about the characteristics of effective remote professional development implementation?

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 Covid-19. The review is also limited to studies from 1995 or later. Remote training approaches from pre-1995 are unlikely to be relevant to current approaches.

Inclusion and exclusion criteria for the review

	Include	Exclude
Population	A. Professionals in education, welfare and public health B. Teachers and teaching assistants of school-age pupils (nursery and school, aged *3-18)*	A. Children under 18 B. Non-professionals (those attending higher education courses)

		C. Studies relating primarily to pre-service teachers and other professionals
Interventions	Direct remote or partially-remote professional development. We define remote PD approaches as methods of training that take place without face-to-face contact. We are interested both in fully remote and blended learning approaches, where distance learning is used for some content delivery/instruction, as opposed to face-to-face models where independent study and practice is likely a part but content delivery is delivered through face-to-face sessions. Distance learning can be synchronous or asynchronous with remote elements accessed by phone, email or online platforms. The primary objective needs to be improving pupil learning outcomes in the teaching profession or equivalent outcomes in other fields (e.g. better patient mental health in psychology). Examples of included approaches are: - Recreating live training sessions through online platforms such as "Zoom" or equivalents - Remote coaching or mentoring that provides intensive support to individual teachers through a two-way link - Digital platforms/education software that are used independently by trainees -	**Purpose of intervention:** Leadership development, operational/procedural tasks, statutory training (e.g. safeguarding) **Format of intervention:** fully face-to-face, one-off idea-sharing conferences
Comparison	Any comparison including: blended approaches; fully face-to-face PD; Other fully remote PD; Business as usual (no training)	No comparison group
Outcome	1) Pupil (or equivalent service beneficiary) outcomes.	Reviews that only examine attendance, engagement with

	2) Behaviour change of the professional which is aimed at ultimately improving pupil outcomes (and the equivalent in other professions). 3) Knowledge change of the professional, which is aimed at ultimately changing behaviour in order to improve pupil outcomes (and the equivalent in other professions). 4) Other benefits for the professional such as job satisfaction	training sessions or satisfaction with/enjoyment of training.
Study design	Meta-analyses or systematic reviews of remote professional development effectiveness or implementation of distance professional development	Single studies, narrative reviews. Systematic reviews addressing research questions other than effectiveness or implementation.
Other criteria	Published since 1995. Published in English. Reviews published in peer-reviewed journals or grey literature.	Published before 1995. Published in languages other than English.

Study designs

We will include any review that identifies as a systematic review or meta-analysis. A systematic review is the process of searching for and selecting evidence using pre-specified criteria, appraising and synthesising it and reaching conclusions about the body of evidence to answer a specific research question. Meta-analysis is a statistical method for combining the results from multiple studies. We will include systematic reviews and meta-analyses that address the effectiveness of remote professional development or systematic reviews that address the barriers and / or facilitators to effective distance learning implementation for professionals.

Population

We will include systematic reviews and meta-analyses that cover distance learning for professionals in education, welfare and public health (see definition in Appendix C) as we believe that there are generalisable mechanisms for effectively training adults (online or otherwise) to change their practice that are relevant to look at across sectors. A more restricted population was considered but an initial review of evidence suggested it was unlikely that restricting to education professionals would produce enough material; furthermore, a recent Campbell systematic review (Filges et al., 2019) on CPD for education, social welfare and justice and crime (including social workers, psychologists and police officers) found 51 studies, all of which related to education, suggesting an expansion to these fields would be unlikely to improve the study yield substantially. Public health is a bigger field with training having similar complexities in terms of monitoring progress and measuring success. It is thought to be more relevant to education than medicine as it involves similarly complex and distal aims. However,

we will also code studies according to their population and will separately analyse those relating to the education and welfare of children and young people. 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.

We will define education and welfare professionals as in Filges (2019) to include those in education, criminal and justice and social welfare settings. We will define public health professionals according to the World Health Organisation's definition: "the art and science of preventing disease, prolonging life and promoting health through the organized efforts of society" (Acheson, 1988; WHO). See Appendix 3 for further exemplification of these definitions.

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, Google Scholar and ERIC, and hand searches of known sources of systematic reviews such as the Campbell Library.

Search Systems and databases to be searched:

- Web of science (Core Collection)
- ERIC (searched through ProQuest)
- Education Database (searched through ProQuest)
- Pubmed (limited to MEDLINE)
- PsycINFO
- Google Scholar[1] (to pick up both published and grey literature)

Other sources:

- Campbell Collaboration Library of Systematic reviews: https://onlinelibrary.wiley.com/journal/18911803
- Cochrane Library of Systematic Reviews: https://www.cochranelibrary.com/search?cookiesEnabled
- EPPI Centre library of reviews: https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=62
- Open Science Framework: https://osf.io/registries?view_only=

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

[1] 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)

million records. We will access it through the EPPI-Reviewer 4 software and use the EPPI Reviewer user guide[2]. This process is conceptually similar to forward and backwards citation tracking.

It is anticipated that there will be some broader reviews with important findings about the relative efficacy of remote professional development approaches, but where there are no references to online or remote training in the titles, abstracts and keywords of the reviews. A manual process will be employed to identify reviews included in the Cordingley (2015) umbrella review relating to teacher professional development, where the main focus may not have been remote professional development but where there might be relevant sections or sub-group analyses related to remote training. We will also aim to capture such reviews through our search, but only for education. This is described in more detail below.

Search terms:

We have drawn on the search terms used in the Cordingley (2015) and Perry et al. (2019) reviews, 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 our professional areas of interest. Where filters on sites correspond to inclusion criteria we will also filter during the search – for example, only searching studies published from 1995.

When searching the education databases, we will apply just the search terms below and then screen texts for findings relating to remote or blended professional development (which may not be in meta-tags, in the title or the abstract).

((“professional development” OR “professional education” OR “professional learning” OR “training” OR “CPD”)

AND (“systematic review” OR “narrative synthes*” OR “literature review” OR “evidence review” OR “quantitative synth*” OR “systematic study” OR "meta regression" or "meta synth*" or "meta-synth*" or "meta analy*" or "metaanaly*" or "meta-analy*" or "metanaly*" or "metaregression" or "metaregression"))

When searching the non-education databases, we will additionally apply the terms below to narrow the search only to those studies that focus on remote and blended learning approaches. Given the education findings are of most relevance, this differential approach to searching databases will limit the volume of studies while ensuring that relevant education findings are not lost.

AND (“Distance” OR “remote” OR “blended” OR “hybrid” OR “e-learning” OR “Internet-based” OR “online” OR “virtual” OR “virtual learning platforms” OR “web-based” OR "virtual learning environment" OR “VLE” OR "learning platform" OR "technology-supported collaborative learning" OR "e-textbook" OR "technology enhanced learning" OR "computer-based learning" OR "computer-assisted training" OR "computer-assisted learning"))

[2] More information is available at: http://eppi.ioe.ac.uk/CMS/Portals/35/MAG%20Browser%20v_1_0_User%20Guide.pdf

Where public health databases allow Boolean operators will be used to exclude any studies with: "medical" 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 remote professional learning approaches that have been implemented within English schools. We will not include 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 PRISMA-style 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 remote professional development approaches. A team will be responsible for extracting information from the included studies using the data extraction tool. The core team will check key data – participant numbers and effect sizes - on a small sample of studies (randomly selected). In cases where the review contains some studies falling into one category and some into another, the mix will be noted and a description included.

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 adapted 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?
    - 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

We will undertake a narrative synthesis (similar to in the recent Distance Learning Rapid Evidence Assessment) 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 remote professional development 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 remote professional development.

We will examine blended learning approaches separately to inform trainer decisions on whether to include some face-to-face elements if possible. Systematic reviews and meta-analyses will be included for review if they include an outcome relating to the efficacy or implementation of distance or blended learning. If there are studies involving distance learning included in the review but no outcomes relating to these studies specifically, reviews will not be included).

To answer research question 2, about how different remote professional development 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). Reviews that compare the effectiveness of remote professional development approaches to other approaches will be used even if only some of the studies contained in the review are remote approaches.

Reporting

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

A school facing publication will summarise the evidence for different remote professional development practices, describing the impact and implementation challenges of different approaches Including fully remote and blended learning approaches.

As well as the research questions stated above, the school-facing report will include a narrative summary of findings from relevant EEF studies to answer the further question, “What can we learn from EEF-funded programmes that have been delivered (fully or partially) remotely?”

Peer review

This REA protocol will be peer reviewed by one reviewer. The REA will be peer reviewed by two peer reviewers.

Personnel

Core team:

- Igraine Rhodes – EEF – principal investigator
- 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 – data extraction coordinator

Data extraction and screening team:

- Mattias Barker – EEF (intern)
- Caroline Bilton – EEF
- Pauline Brown – EEF
- Simon Cox – EEF
- Katie Evans – EEF (intern)
- Phoebe Fisher – EEF
- Niki Kaiser – EEF
- Siobhan Keddie – EEF
- Sue Morgan – EEF
- Lorwyn Randall – EEF

Conflicts of interest

No conflicts of interest

Timeline

	Task	Completion date
Protocol development	Protocol finalised	CoB – Wednesday 15th July
	School facing report template finalised	CoB – Wednesday 29th July
	Data extraction tool finalised	CoB – Wednesday 15th July
Search	Academic search	CoB – Friday 17th July
	Search grey literature and organisational websites	CoB – Friday 17th July
	Identify relevant EEF studies	CoB – Monday 20th July
	Citation tracking (checking included studies in the relevant reviews)	CoB – Monday 20th July
Screening	Screening at title (and abstract)	CoB – Monday 20th July
	Full text retrieval (assuming 50 papers)	CoB – Thursday 23rd July
	Full text screening	CoB – Thursday 23rd July
Data Extraction	Data extraction (dsecriptives, intervention, implementation, effect sizes), assuming 10 per day	CoB – Wednesday 29th July
	Check data extraction	CoB – Thursday 30th July
	Data extraction (critical appraisal of reviews)	CoB – Wednesday 5th August

Synthesis and write-up	Narrative synthesis	CoB – Monday 10th August
	Write up of new synthesis	CoB – Monday 10th August
	Write up of relevant EEF / IES studies	CoB – Monday 10th August
	Review of findings document by D&I, Becky, Stephen, Robbie (who else?)	midday – Thursday 13th August
	Publish on EEF website	Friday 4th September
	Write up draft technical report using REA template	Not immediate priority
Peer Review	Select and approach 2 peer reviewers	CoB - Wednesday 1st July
	Send to peer reviewer	CoB – Thursday 13th August
	Peer Reviews complete	CoB - Monday 24th August
	Integrate Peer Reviewer comments	CoB – Wednesday 26th August
Supportive resources for schools	Summary doc	CoB – Thursday 27th August
	Feedback from wider team	CoB - Monday 31st August
	Online briefing	CoB – Wednesday 2nd September
	Compilation of school facing findings	CoB – Wednesday 26th August

References

Boelens, R., De Wever, B., & Voet, M. (2017). Four key challenges to the design of blended learning: A systematic literature review. Educational Review Research, 22.

Brettschneider, A., & Mather, M. A. (2005). Improving online collaborative learning for teachers: How changes to design features of the adolescent literacy collaboratory influenced participant retention, overall satisfaction, and engagement. Retrieved from https://www.brown.edu/academics/education-alliance/sites/brown.edu.academics.education-alliance/files/publications/imponlcollab.pdf

Cordingley, P., Higgins, S., Greany, T., Buckler, N., Coles-Jordan, D., Crisp, B., & Coe, R. (2015). Developing great teaching: Lessons from the international reviews into effective professional development. London: Teacher Development Trust.

Darling-Hammond, L., Hyler, M., & Gardner, M. (2017). Effective Teacher Professional Development. Leaning Policy Institute.

Department for Education (2016) Standard for teachers' professional development https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/537031/160712_-_PD_Expert_Group_Guidance.pdf

Dunst, C. (2015). Improving the Design and Implementation of In-Service Professional Development in Early Childhood Intervention. Infants and young children 28(3):210-219

Education Endowment Foundation (2020a) Remote Learning, Rapid Evidence Assessment, London: Education Endowment Foundation.

Education Endowment Foundation (2020b) Impact of school closures on the attainment gap: Rapid Evidence Assessment, London: Education Endowment Foundation.

Filges, Trine et al. (2019) for Campbell Collaboration. Effectiveness of continuing professional development training of welfare professionals on outcomes for children and young people https://onlinelibrary.wiley.com/doi/epdf/10.1002/cl2.1060

Fishman et al. (2013). Comparing the Impact of Online and Face-to-Face Professional Development in the Context of Curriculum Implementation. Journal of Teacher Education 64(5) 426 –438

Garritty C., Gartlehner G., Kamel C., King V.J., Nussbaumer-Streit B., Stevens A., Hamel C., Affengruber, L. (2020). Cochrane Rapid Reviews. Interim Guidance from the Cochrane Rapid Reviews Methods Group.

Government Social Research Service (2009). GSR Rapid Evidence Assessment Toolkit (Online), https://webarchive.nationalarchives.gov.uk/20140402164155/http://www.civilservice.gov.uk/networks/gsr/resources-and-guidance/rapid-evidence-assessment

Haddaway, N. R., Collins, A. M., Coughlin, D., Kirk, S. (2015). The role of Google Scholar in evidence reviews and its applicability to grey literature searching. PloS one, Volume 10(9).

Hill, H.C., et al. (2020). Professional development that improves STEM outcomes. Phi Delta Kappan Volume: 101 issue: 5, page(s): 50-56

Kalinowski, E., et al. (2019). Effective Professional Development for Teachers to Foster Students' Academic Language Proficiency Across the Curriculum: A Systematic Review. AERA Open January-March 2019, Vol. 5, No. 1, pp. 1 –23

Kennedy, M. (2016). How Does Professional Development Improve Teaching? Review of Educational Research Vol. 86, No. 4, pp. 945-980

Perkins, J. (2010), Personalising teacher professional development: Strategies enabling effective learning for educators of 21st century students . Quick, (113), 15-19.

Perry, E. et al. (2019). Quality Assurance of Teachers' Continuing Professional Development Rapid Evidence Review. https://educationendowmentfoundation.org.uk/evidence-summaries/evidence-reviews/teachers-continuing-professional-development/

Pollock, M., Fernandes, R.M., Becker, L.A., Pieper, D., Hartling, L. (2020). Chapter V: Overviews of Reviews. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0.

Russell, M., et al. (2009). Face-to-face and online professional development for mathematics teachers: a comparative study. Journal of Asynchronous Learning Networks, v13 n2 p71-87

Shemilt, I., Khan, A., Park, S. Thomas, J. (2016). Use of cost-effectiveness analysis to compare the efficiency of study identification methods in systematic reviews. Systematic Reviews, Volume 5(140), pp. 1 – 13.

Sun, A., Chen, X. (2016). Online Education and Its Effective Practice: A Research Review. Journal of Information Technology Education: Research, Volume 15, pp. 157-190. https://doi.org/10.28945/3502

Thomas, J., Harden, A. (2008) Methods for the Thematic Synthesis of Qualitative Research in Systematic Reviews. BMC Medical Research Methodology, 8(45), pp. 1–10.

Thomas, J., Brunton, J., Graziosi, S. (2010). EPPI-Reviewer 4: software for research synthesis. EPPI-Centre Software. London: Social Science Research Unit, UCL Institute of Education.

Vrasidas, C., & Zembylas, M. (2004), "Online professional development: lessons from the field", Education + Training. Volume 46. Issue 6/7. Page 326 - 334.

World Health Organisation (2012). European Action Plan for Strengthening Public Health Capacities and Services. Accessed July 2020: https://www.euro.who.int/__data/assets/pdf_file/0005/171770/RC62wd12rev1-Eng.pdf

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 remote learning?
  - School staff
  - External agency
  - Peer-facilitated
  - Automated/individual study
  - Other (describe)
- Overall duration of training (number of days/months/years)
- Proportion of the instructional content delivered remotely
- 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
- Population (open response)
  - School and nursery teachers
  - School Leaders
  - Teaching Assistants
  - Other education professionals
  - Welfare professionals
  - Public health professionals
- 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?
- Pooled effect (numerical value)
- Effect size type (open response)
- Outcome measure (open response)
- What is the comparison (select one)?
  - Fully online
  - Blended
  - Face-to-face training
  - Pure control (no CPD)
- 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
  - No
- 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? Did the authors specify: ☐ Types of studies ☐ Participants/ settings/ population ☐ Intervention(s) ☐ Outcome(s)	☐ Yes ☐ Partially ☐ No *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? 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	☐ Yes ☐ Partially ☐ No ☐ Can't tell *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? 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	☐ Yes ☐ Partially ☐ No *Coding guide:* *YES: All should be yes* *PARTIALLY: Independent screening not done* *NO: All other.*
Did the authors use appropriate criteria to assess the quality and risk of bias in analysing the studies that are included?[i] ☐ The criteria used for assessing the quality/ risk of bias were reported ☐ A table or summary of the assessment of each included study 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.	☐ 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, 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'.*	☐ **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)

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 exctraction ☐ 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	☐ Yes ☐ No ☐ Partially ☐ Not applicable (e.g. no 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.* *NOT APPLICABLE: if no studies/no data*

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) 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)	☐ Yes ☐ No ☐ Not applicable (e.g. no studies or no data) ☐ Can't tell *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)*

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? ☐ Were factors that the review authors considered as likely explanatory factors clearly described? ☐ Was a sensible method used to explore the extent to which key factors explained heterogeneity? ☐ Descriptive/textual ☐ Graphical ☐ Meta-analysis by sub-groups ☐ Meta-regression ☐ Other	☐ 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.* *High confidence applicable when all the answers to the questions in section B are assessed as 'yes'.* *Low confidence applicable when any of the following are assessed as 'NO' above: critical characteristics of the included studies not reported (B1), not describing the extent of heterogeneity (B3), combining results inappropriately (B4), reporting evidence inappropriately (B5).* *Medium confidence applicable for any other: i.e. the "Partial" option is used for any of the 6 preceding questions or questions and/or B.2 and/ or B.6 are assessed as 'no'.*	☐ **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

C.1 Are there any other aspects of the review not mentioned before which lead you to question the results?	☐ Additional methodological concerns – only one person reviewing ☐ Robustness ☐ Interpretation ☐ Conflicts of interest (of the review authors or for included studies) ☐ Other ☐ No other quality issues identified
C.2 Are there any mitigating factors which should be taken into account in determining the reviews reliability?	☐ Limitations acknowledged ☐ No strong policy conclusions drawn (including in abstract/ summary) ☐ Any other factors
Use comments to specify if relevant, to flag uncertainty or need for discussion	
C.3 **Based on the above assessments of the methods how would you rate the reliability of the review?** ☐ **Low confidence in conclusions about effects**: ☐ **Medium confidence in conclusions about effects**: ☐ **High confidence in conclusions about effects**: 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.	

Appendix C: Exemplification of professions:

1) **Education and welfare professions**
   Examples of education and welfare professionals include teachers, teacher assistants, pre-school teachers (pedagogues), care providers, social workers, paraprofessionals, psychologists, police officers, family support providers, disability specialists, inclusion specialists. (Filges, 2019)

2) **Public health**

   Ten essential public health operations (EPHOs)

   1. Surveillance of population health and well-being
   2. Monitoring and response to health hazards and emergencies
   3. Health protection including environmental, occupational, food safety and others
   4. Health promotion including action to address social determinants and health inequity
   5. Disease prevention, including early detection of illness
   6. Assuring governance for health and well-being
   7. Assuring a sufficient and competent public health workforce
   8. Assuring sustainable organizational structures and financing
   9. Advocacy, communication and social mobilization for health
   10. Advancing public health research to inform policy and practice (WHO, 2012)

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