Reforms to Increase Teacher Effectiveness in Developing Countries

A Systematic Review Protocol

Protocol written by Monazza Aslam (UCL Institute of Education), Shenila Rawal (University of Bristol), Geeta Kingdon (UCL Institute of Education), Bob Moon (The Open University) Rukmini Banerji (ASER Centre, Pratham Education Foundation), Sushmita Das (UCL Institute of Education), Manjistha Banerji (ASER Centre, Pratham Education Foundation), Shailendra K. Sharma

EPPI-Centre
Social Science Research Unit
UCL Institute of Education
University College London

December 2014

PROTOCOL

The authors are part of the UCL Institute of Education, University College London; University of Bristol; The Open University; and ASER Centre, Pratham Education Foundation; and were supported by the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre).

This protocol should be cited as: Aslam M, Rawal S, Kingdon G, Moon B, Banerji R, Das S, Banerji M, Sharma SK (2014) Reforms to Increase Teacher Effectiveness in Developing Countries: A Systematic Review Protocol. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London

Contents

1. Background

Section 1.1 introduces the basic principles that are discussed in more detail in the rest of the chapter.

1.1 Aims and rationale for review

Student learning in developing countries is persistently poor (Filmer, Hasan and Pritchett, 2006; Annual ASER reports on India and Pakistan; UWEZO in East Africa, 2011, 2012). Strong and consistent international evidence shows that teaching quality is probably the single most important institutional influence on student outcomes, with several studies strongly endorsing the need for interventions that focus on teachers and teaching quality (Goldhaber 1999, Clotfelter et al. 2006, Burgess et al. 2009, Hanushek and Woessmann, 2011).

Much of the high quality Randomised Control Trial (RCT) studies in various country contexts indicate that simply supplying more resources (more teachers or textbooks) is not the panacea. Deep-rooted distortions in developing country education systems – such as elite curricula and weak teacher incentives – undermine efforts to achieve desired objectives (Kremer and Holla 2009). Interventions and reforms that work around these distortions may, however, be able to achieve higher student achievement at low cost (Kremer and Holla 2009, Glewwe et al. 2013). This review will aim to identify quality evidence pertaining to reforms/interventions in education systems aimed at improving teacher effectiveness. The overall objective of such reforms can be seen to improve the overall quality of teaching and/or improve student achievement.

Effective education systems fundamentally build on good governance, robust public financial management and, inevitably, the effective management of teachers (including recruiting, training and deploying them)1. The major challenge of increasing teacher effectiveness lies in recruiting and training competent teachers and significantly improving the effectiveness of teachers already in post. There is also a strong sense that successful development involves taking good ideas and practice 'to scale' (AusAID, 2012). This review will aim to identify the literature that investigates all these aspects: what are the teacher effectiveness reforms at scale that have successfully improved teaching quality and student outcomes and what are the technical, political and financial barriers that have been overcome in the process.

The Intervention

It is possible to think of teacher effectiveness as a continuum from very direct attempts at impact (an in-service programme to improve the teaching of literacy or a bonus system applied to teacher personal performance) to more indirect interventions/reforms (paying recruitment grants to attract more effective mathematics or science teachers or generally improving teacher salaries to encourage them to work more productively). In the middle of the continuum might come the training of school leaders to promote teacher

[1]DFID Education Position Paper, 'Improving Learning, Expanding Opportunities', July 2013, Department for International Development.

effectiveness. The scope of this review will encompass a wide range of interventions that fall along this spectrum.

Some examples of interventions to improve teacher effectiveness at scale include (but are not limited to): contract teacher schemes such as those undertaken in many parts of Africa and Asia and while the narrative of these schemes is to overcome teacher shortages, by improving teacher accountability they are seen to improve teacher effectiveness and improve student outcomes (see Kingdon et al. 2013 for a systematic review and Bold et al.'s, 2013, study of scaling up of contract teachers in Kenyan primary schools); teacher training and education schemes such as TESSA (Teacher Education in Sub-Saharan Africa) which provides online teacher training/education and resources to teachers or the rolling out of INSET training under the SarvaShikshaAbhyaan (SSA) in India or the Read India campaign launched by Pratham in collaboration with the Government of India or the English in Action campaign launched in Bangladesh; teacher community assistant programs such as the Ghana Government's Teacher Community Assistant Initiative(TCAI), aimed at improving literacy and numeracy levels in basic schools with the view for national roll out; improved monitoring systems such as the those provided under -the Punjab Education Reform Roadmap in Pakistan which aim to improve the functioning of the education system (Sir Michael Barber, 2013), instituting merit-pay schemes, merit-hiring schemes (such as through the Teacher Eligibility Test in India), computer and technology assisted learning schemes (such as Text2Teach in the Philippines) and so on. Studies evaluating these schemes highlight the technical, financial and political barriers and 'drivers of change' in implementing these reforms. Experiences in other countries to scale-up reform programmes based on impact evaluation results have often been hindered by political-economy factors (Acemoglu, 2010) or been aided by strong political will (as discussed by Sir Michael Barber's think-piece on Punjab's education reforms), or mired by technical barriers (as highlighted in the TESSA evaluation report) and financial constraints (Bold et al., 2013).

1.1.1 Review questions

The review objectives will be addressed through answering the following review questions:

1. What is the evidence on the impacts of reforms/interventions of education systems at scale to increase teacher effectiveness on: the quality of teaching and on learning outcomes in low and middle income countries?

   This question will be answered by synthesising evidence from the experimental and quasi-experimental literature on effectiveness and will focus on 2 key outcomes. The first outcome relates to improvements in teaching quality. This will incorporate measures such as teacher credentials, effort, time-on-task, absenteeism, content-knowledge, improvements in pedagogy etc.

2. What is the evidence on the relationship between educational reforms/interventions for improving teacher effectiveness at scale and the quality of teaching and learning outcomes in low and middle income countries?

This question will draw on studies which focus on the relationship between teacher effectiveness on student outcomes. This might be numerical data (e.g. correlational studies using statistical analysis) or qualitative studies (e.g. drawing on participants perception that outcomes have improved as a result of participating in an intervention) but which do not establish causation or direction of impact. Outcomes of interest might include both quality (cognitive test scores) and quantity (years of schooling, completion and progression rates). However the focus of this review will still remain on studies that investigate this relationship with regards to cognitive educational outcomes of the students. It will exclude any studies pertaining to non-cognitive skills such as social, emotional and physical development skills.

3. Where reforms/interventions to education systems to increase teacher effectiveness at scale have occurred, what is the evidence on how technical, financial and political barriers have been overcome?

   From the studies that have been identified as answering question 1 and 2, we will also extract any relevant qualitative and descriptive evidence that examines the technical, financial and political economy issues that have either enhanced or hindered the implementation, progress and impact of teacher effectiveness reforms at scale.

The search strategy will be designed to ensure that we include a broad range of interventions aimed at improving teacher effectiveness. The issue of scale will then be analysed during the screening stage (see methodology section for further details). The question of scale is very context and programme specific and, therefore, using a stringent and quantified definition of 'scale' may limit our research. In examining this issue, we will consider aspects such as administrative scalability, functional scalability, geographical coverage, population coverage and/or schemes that have been implemented in a range of different contexts within the same country or across different countries etc.

1.2 Definitional and conceptual issues

Figure 1 below sketches the Theory of Change of teacher effectiveness reforms at scale as studied in this review. The aim is to identify the possible relationships between these programs and the outcomes of interest. Additionally, we aim to identify which assumptions/associations are supported by evidence and where the evidence is especially weak. For example, while it may be believed that teacher effectiveness interventions such as performance incentives, have a positive effect on student learning, there may not be robust evidence to support this assumption. This review will identify the different 'assumptions' or 'chains' and indicate clearly where the evidence is especially strong to support a causal relationship or where causality cannot be assumed.

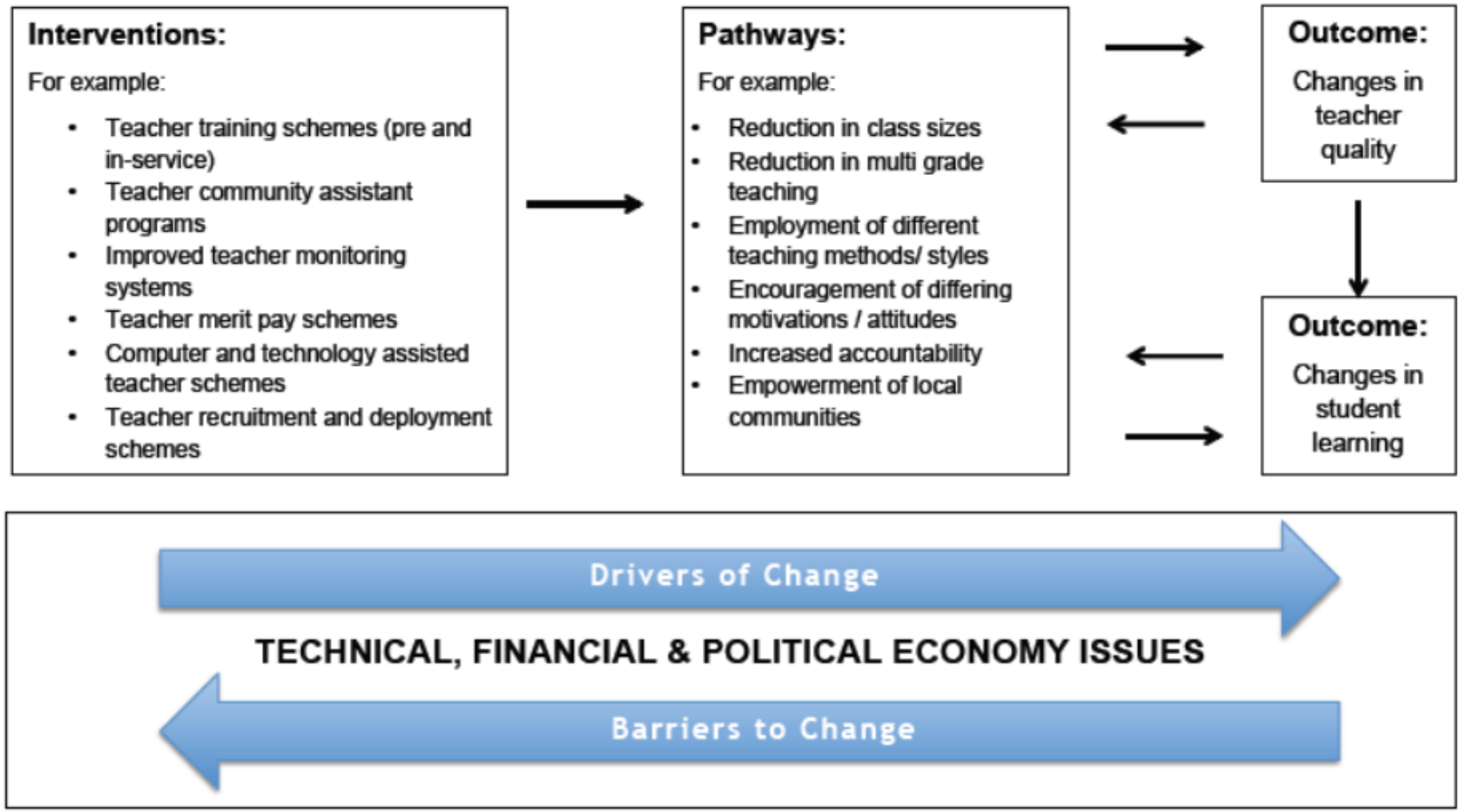
The left hand side of this diagram indicates the first step in this relationship namely the interventions. For example, these may include (and are not limited to) teacher training and education programmes, recruitment and retention programmes, remuneration programmes, teacher deployment programmes etc. Theoretically, the introduction of these programmes may improve teacher quality and/or student outcomes (specifically

student learning) through various pathways/channels of change as indicated in the middle column of the diagram. For example, a training intervention that supports teachers through in-service training, may alter pedagogical teaching styles which in turn may improve student learning either directly or indirectly through enhancing teacher quality (for instance through increased time on task or lower absenteeism). There is the possibility that the intervention may have either no effect or may negatively impact student learning and/or teacher quality. For example, a programme aimed at reducing teacher shortages (such as an intervention hiring teachers on contracts rather than on a permanent basis) may result in the hiring of a lower quality pool of teachers which may potentially negatively impact student outcomes.

The education system and the mechanisms around its organization do not exist in isolation and are often influenced by incentives and constraints operating within the wider environment. This means that educational reforms, no matter how well meaning, may be influenced at the design, financing, implementation or even the evaluation stages by factors that may enhance or hinder the effectiveness of the interventions themselves. The theory of change depicted below explicitly allows for the examination of these technical, financial and political economy issues with the view to identify situations where certain barriers have been overcome and allowed certain education system interventions to be achieved at scale.

It should be noted that this initial Theory of Change will be adapted and supplemented once the review of literature has been undertaken to comprehensively cover the key associations and identify specific barriers to reform and drivers of change within different education systems studied.

Figure 1: Theory of Change



1.3 Policy, practice and research background

The question posed in this review is unique in that it asks a very critical and policy relevant issue. This systematic review will have several contributions:

1) It will allow us to identify teacher effectiveness programmes that have occurred at scale and examine the literature thereof. While there is a wide range of literature examining different teacher effectiveness reforms such as in-service teacher training/education programmes and merit pay schemes etc. (at scale or otherwise) across several contexts individually, this systematic review will collate this evidence specifically for interventions at scale and identify robust findings from a widely dispersed literature base into a concise review.

2) By linking the theoretical framework to the literature base, we hope to provide guidance to policy makers and practitioners. This will be done by identifying the possible relationships that exist between different interventions and outcomes (and vice versa) and the channels and assumptions through which these work. In conducting this review, we will be able to provide a clear indication as to which of these assumptions and channels are supported or negated by robust evidence and where further research should be directed.

We are not aware of any systematic reviews that directly address this question. However, a recent systematic review undertaken by some of the authors of the current proposed review is based on looking at the evidence on one aspect – contract teachers - that will be important in the discussions surrounding the current review.

2. Methods used in the review

2.1 User involvement

2.1.1 Approach and rationale

The team members have experience of conducting Systematic Reviews and Rigorous Literature reviews and will follow the stringent steps needed to arrive at an evidence base that helps address the question in a systematic and rigorous manner. The literature is likely to be spread across different disciplines and the strength of the proposed team lies in combining experts from the quantitative and qualitative disciplines who will aim to ensure a full scoping out of the evidence base has been achieved and the findings appropriately retrieved, critically appraised and synthesised with a policy context in mind.

2.1.2 User Involvement in designing the review

With teacher effectiveness policies at the forefront of policy-makers agendas, it is expected that this review will be of substantial interest to key stakeholders. During the process of conducting this review, the authors will remain engaged with DFID advisers to allow them to have the opportunity to discuss the scope and content of the review to ensure it provides them meaningful policy insights. Once prepared, the review will be presented to DFID and policy-makers and other stakeholders through roundtables and discussions. Specifics sections of the review – in particular, the policy implications and recommendations section – may be most valuable for these discussions. The review will also be disseminated to authors of contributing studies who may be consulted during the review process. It is also the aim of the authors to publish this review in peer reviewed journals and to present the findings in seminars at universities in the UK and in other countries where the authors may be based to reach a wider academic and policy-making audience.

2.2 Identifying and describing studies

2.2.1 Defining relevant studies: inclusion and exclusion criteria

This review will include studies which:

Population: focus on DFID priority countries and will include all countries currently listed as low or lower middle income by the World Bank. (http://data.worldbank.org/about/country-classifications).

This does run the risk that some countries which have moved from lower middle income to higher middle income between 1990 and 2014 will be missed. However, as it is inevitable that countries change over time, the risk of missing 1 or 2 potentially relevant countries will be inherent in any country based exclusion criteria. We will not confine our countries to DFID aid recipients, as the choice of recipients is fluid and this excludes many countries where schemes exist or may be proposed in the future. The review will exclude: high-income countries, transition economies and upper middle-income countries.

Setting: primary and secondary schooling including government schools.

Intervention: investigate teacher effectiveness interventions (such as contract teacher schemes, pre-service training, in-service training, merit pay, computer assisted teaching and learning etc.)

Comparison:

To answer review question 1 on the effectiveness of teach effectiveness interventions studies need to have a comparison group. For example, treated teachers versus non-treated teachers or, students taught by treated teachers vs. those taught by non-treated teachers

Or

Reporting data:

To answer review question 2 studies need to report evidence on the relationship between educational reforms/interventions for improving teacher effectiveness at scale on relevant outcomes (see below) using either numerical or narrative data.

Outcomes: report learning outcomes, e.g. academic achievement tests and/or teacher quality e.g. time on task, teacher motivation, competence, absence, skills, effort, qualifications, credentials, teacher test scores, and characteristics.

Date: are published from 1990 onwards in order to maintain policy relevance.

Language: studies written in English only, as the scope of this review does not extend to sourcing and translating non-English language

Scale: provide data on the impact of the intervention 'at scale'.

2.2.2 Identification of potential studies: search strategy

2.2.2.1 Search terms

A 'search strategy' will be devised and search terms arrived at. An iterative procedure will be used to search for the relevant literature using a number of key words and synonyms to ensure that all possible evidence has been covered. Search terms will be agreed among team members and all effort will be made to ensure that the final terms allow for different kinds of literature to be covered including quantitative, qualitative and grey literature. Systematic reviews and rigorous reviews that address the question of interest will also be reviewed.

The key search terms will be determined by the review question and the inclusion and inclusion criteria as outlined above. The search strategy will involve developing strings of terms and synonyms to denote five key aspects of the review, namely:

Concept 1: synonyms of **"interventions"** and **"reforms"** with a focus on education. Initial searches will be done only with this concept. If an un-manageable number of hits are obtained within the databases then search strings will include the remaining concepts.

Concept 2: - synonyms which capture aspects of **teacher quality** only

Concept 3: key terms to capture a mix of aspects of **student outcomes** using the synonyms of "student" and "outcomes".

Concept 4: search terms will include aspects of **political economy** issues prefixed with synonyms for reforms/interventions where possible

Concept 5: A list of **low and lower middle income countries** as defined by the world bank (see earlier for definition and link to countries).

The search strategy that will be adjusted according to individual databases and web-based interface capabilities as required;

- The searches in each database will initially begin with (CONCEPT 1: interventions/reforms) AND (CONCEPT 5: LMIC). This is to ensure that the search is as broad as possible as a start point.
- If the initial search yields an unmanageable number of papers, then three separate searches will be run and the outputs from each will be saved for screening. The searches will be as follows;

  (CONCEPT 1: interventions/reforms) AND (CONCEPT 2: teacher quality) AND (CONCEPT 5: LMIC)

  (CONCEPT 1: interventions/reforms) AND (CONCEPT 3: student outcomes) AND (CONCEPT 5: LMIC)

  (CONCEPT 1: interventions/reforms) AND (CONCEPT 4: Political economy) AND (CONCEPT 5: LMIC)

  [Note: This may have to be adjusted for some of the databases depending on their scope].

- Where date settings are available, the searches in the database will be restricted to literature between 1990 and 2014. If date settings are not available in the databases, we will screen out literature pre-1990 during the title and abstract screening stage.
- Wild cards, proximity searches and thesaurus terms will be used as appropriate.

Search notes and search strings for each database will be maintained and logged within supporting documentation. Examples of search terms and strings applied to an example database can be found in Appendix 2.2

Table 1: Search Databases

Platform	Database	Details
Databases for published papers and reports		
EBSCO	**TRC -** Teacher Reference Center	Indexing & abstracts for 280 of the most popular teacher and administrator journals and magazines
	eBook Collection	Search and view the full text of eBooks.
	ECONLIT - Economic Literature	EconLit with Full Text contains all of the indexing available in *EconLit,* plus full text for nearly 600 journals
	ERC - Education Research	Provides indexing and abstracts for more than 2,100 journals, as well as full text for more than 1,200 journals
PROQUEST	**ASSIA -** Applied Social Sciences Index and Abstracts	Health services, social work, sociology and psychology - journal articles
	ERIC - Education Resources Information Center	Education Resources Information Center
	IBSS - International Bibliography of the Social Sciences	
	ProQuest Dissertations & Theses	UK & Ireland
	AEI	Australian Education Index
	PSYCINFO	Psychology journals database
	BEI	British Education Index
WEB OF KNOWLEDGE	**CPCI-SSH**	Conference Proceedings Citation Index - Social Science & Humanities
	WEB OF Science	All sciences and humanities
JSTOR	**JSTOR**	Social sciences
SCIENCE DIRECT	**SCIENCEDIRECT**	All sciences and humanities

AJOL	**AJOL**	Africa Journals Online
ASIAJOL	**ASIAJOL**	Asia Journals Online
Databases for working papers and reports (grey-literature)		
SSRN	**SSRN**	Social Science Research Network
NBER	**NBER**	NBER working papers
REPEC	**Econpapers**	Research Papers in Economics
CID	**CID**	Center for International Development of Harvard University
3ie International Initiative for Impact Evaluation	**3ie**	Impact evaluation reports (both peer reviewed and working papers), systematic reviews
WORLD BANK	**WORLD BANK**	Working papers, reports (including DIME)
JPAL	**Poverty Action Lab**	Working papers, published articles, reports
IMF	**IMF**	Working papers, reports
UNDP	**UNDP**	Research papers, reports
UNESCO	**UNESCO**	Research papers, reports
ILO	**ILO**	Working papers and reports
CREATE	**CREATE**	Consortium for Research on Educational Access, Transitions and Equity
British Library	**Index of Conference Proceedings**	
Additional sources for grey literature (eg. conferences) and grey literature itself to be included by team-members.		
Databases for Theses & Other		
SIGLE	**SIGLE**	
GOOGLE SCHOLAR	**GOOGLE SCHOLAR**	
Hand Searching		
Hand Searching		This involves searching manually through references of shortlisted papers. This will need to continue even after full-text screening as we will possibly need to locate additional papers from shortlisted bibliograhies/references

2.2.3 Screening studies: applying inclusion and exclusion criteria

Title and abstract screening

Once the studies have been identified, they will be uploaded to EPPI-reviewer and screened for their relevance to the systematic review. The inclusion criteria are formally stated using the "PICOS" method (Population, Intervention, Comparison and Outcomes,), to which we added time. These inclusion and exclusion criteria will be applied to title and abstract.

Excluded studies will be those that fail to satisfy at least one inclusion criteria or meet at least one of the exclusion criteria. Studies that fail to meet the inclusion criteria will be coded as such that it can be reported how many inclusion criteria each study failed to meet and what these criteria were. If a study scores "Yes" for the relevance criteria it will be brought forward to the full-text screening stage.

Full text screening: Inclusion/Exclusion criteria

At the full text screening stage, full reports will be obtained for those studies that appeared to meet the criteria or where there was insufficient information to decide. This will involve reviewing the full text and re-applying the PICOS framework. The included studies will be coded by various indicators (region, setting, sample size etc.) and will be taken forward to the final stage, the in-depth review. At this point in the review, we propose to distinguish between interventions of the following types: i) Interventions at scale and and ii) interventions not at scale. Only studies that fall in category (1) will be carried forward to the synthesis stage, and studies categories as not at scale will be 'described only' (see below).

2.2.4 Characterising included studies

The review team will organise studies by type. This means, each study will be assigned a code based on the intervention/s being studied, the outcomes being analysed, level of schooling, country, potential for effect size analysis, and findings. At the final stage, the studies will then be coded based on whether the intervention is 'at scale' or not. All the keyworded studies will be added to the larger EPPI-Centre database, for others to access via the website. Firstly, this will allow recurring themes to be identified. Additionally, by having the literature-base coded in such a functional and constructive manner, will allow the research term to capitalize on the extensive research base that is covered to the benefit of the funder. For example, at some stage, should DFID wish to broaden the scope of this research and investigate non-scale interventions or pursue any key sub-themes, this coding strategy will allow for that specific literature base to be easily accessible and extracted from the current review.

2.2.5 Identifying, describing and synthesising studies: quality assurance process

The systematic review will follow the standard EPPI-Centre procedures for maintaining quality. At the scoping review stage, to ensure consistency in application of the selection criteria, reviewers will undertake double screening on a sample of papers to pilot the inclusion/exclusion criteria. The remainder of the screening will be carried out by individual reviewers. Where there was uncertainty, reports will be marked for discussion and at the end of the screening process these reports will be considered by two or three reviewer, as required. As a final check, all reports selected for inclusion will be checked

by the second reviewer to confirm their relevance. At the synthesis stage, data extraction and quality assessment processes will have been undertaken by two researchers working independently, in order to achieve a high level of consistency.

2.3 In-depth review

2.3.1 Assessing quality of studies and weight of evidence for the review question

Studies identified as meeting the inclusion criteria will be analysed in depth using a consistent and detailed data extraction methodology (DfID, 2014, please see appendix 2.4). The dimensions used for this critical appraisal will include assessing the methodological quality of the study, the relevance and appropriateness of the research design and the relevance of the focus of the study. These will feed into judgments () to provide an overall Weight of Evidence to determine if the findings of studies are trustworthy (see table 3 in appendix 2.4).

The validity, reliability and applicability of each study will be examined in a consistent and comprehensive manner. For example, a hierarchy of evidence will be used to evaluate the validity of quantitative studies ranging from Randomized Control Trials (high quality) to less rigorous methodologies such as simple descriptive statistics that do not allow causal interpretations (such as comparison of means). The validity of studies with narrative/qualitative data will be analysed by choosing those that give relevance to the wider context and based on factors such as the extent to which the study employs a methodology that minimizes the risk of bias. This will include assessment of risk of bias in relation to: selection bias, selective reporting bias, placement bias, consideration of intervention integrity, consideration of differences within groups, explaining variations in outcomes amongst other factors (Stewart et al. 2012).

In relation to reliability, all studies will be judged on the basis of whether their findings are reproducible while the applicability of the studies will be judged on the basis of whether their findings are appropriately applicable. Quality assurance and consistency will be ensured by assessing studies independently but also by using a prior agreed-upon approach.

All studies, irrespective of design, will be assessed according to the following criteria:

Completeness of reporting: this entails assessing transparency, reporting bias and publication bias. We expected a good quality study to have a description of the intervention and the participants (children), clear account of methods of data collection and analysis and consideration of confounding factors along with complete reporting in relation to measured results. A study will be considered of poor quality if it fails to meet one or more of these requirements.

Feasibility of assumptions: If the reviewers remain unconvinced about the assumptions made within the study on which the conclusions were based, the study will be classified as of low or medium quality. Whilst studies that do not specifically articulate their assumptions will not automatically be excluded, the assessment of quality will be affected by whether or not a study articulates its assumptions clearly or not.

Appropriateness of methodology: Methodology will be analysed to ensure trustworthiness, reliability and validity. Assessment of the appropriateness of the methodology depends on

whether a study is quantitative or qualitative in nature. These will be assessed according to the approaches discussed below.

Consideration of confounding factors: these will include (when necessary) assessing sampling bias, attrition bias, detection bias, endogeneity bias, ability to address heterogeneity effectively and so on. Confounding factors can be controlled for at the sampling stage or at the analysis stage. If studies take no consideration of confounding factors at either stage, they will be considered of poor quality and will be excluded from the in-depth review. Studies that control for confounding factors at any one stage will be considered of medium quality and be included in the in-depth review. Studies controlling for these factors at both stages will be considered of high quality and will be used for in-depth review.

Comprehensive reporting of findings: Were the studies finding apparent and comprehensively reported? For example, if the study initially aimed to measure certain outcomes and did not report on all of the outcomes, it will be judged of poor quality and be excluded.

Evaluating the quality of methodology of quantitative studies

We will use the following critical appraisal approach to evaluating the methodology of quantitative studies:

(i) How was the intervention assigned? i.e. was assignment random or non-random? If random, the study will be judged of high quality, if non-random or medium or poor depending on how the intervention was further assigned.

(a) If randomized, is the counter-factual clearly stated?

(1) teacher training/education versus un-trained teachers?

(2) Contract teacher versus regular teacher?

(3) Classroom assistant versus no assistant?

Etc.

Yes/No/Partly, if yes clearly stated, the study will be considered of high quality, partly will be medium and when it is not clearly stated the study will be considered of poor quality.

(b) If non-random, is selection bias a threat to internal validity? Yes or No. If selection bias threatens internal validity, then:

(1) Is the selection explicitly modelled or controlled for? If not, the study will be considered of poor quality and excluded from the in-depth review.

(ii) What question is being asked in the study? Do they evaluate the ‘as is’ effect of the intervention, or do they evaluate a conditional effect? A study that controls for the 'as is' effect will be considered medium quality while one controlling for the conditional effects will be considered of high quality. Both will be included in the in-depth review.

(iii) Is the intervention effect homogenous across different student-types? Studies that consider the intervention affect across different student types will be considered high quality, those that do not will be considered medium quality and both will be included in the final review.

(iv) Is the cohort representative of the population? If not, does the sample have any characteristics which may affect the external validity of results? If so, the study will be considered of poor quality and be excluded from the in-depth review.

Evaluating the quality of methodology of qualitative studies

1) Is the epistemological approach clearly stated? Yes/No/partly - studies where it is clearly stated will be of high quality, partly stated are of medium quality and where it is not stated at all are of poor quality.

2) Was sampling appropriate? Yes/No/partly? Studies where it was appropriate were considered of high quality, partly stated are of moderate quality and where it is not stated at all are of poor quality.

3) Was data collection appropriate/repeatable and trustworthy? Yes/No/partly - studies where it was appropriate/trustworthly were considered of high quality, partly stated are of medium quality and where it is not stated at all are of poor quality.

4) Was the approach to data analysis appropriate/repeatable and trustworthy? Yes/No/partly - studies where it was appropriate/trustworthy were considered of high quality, partly stated are of medium quality and where it is not stated at all are of poor quality.

Based on the findings of the above, studies will be judged to be either High, Moderate of Low Quality. In order to ensure rigor, judgments relating to the above will be done independently.

2.3.2 Synthesis of evidence

2.3.2.1 Overall approach to synthesis

A preliminary instrument has been designed to guide the researchers in retrieving the information from each study. These forms will be filled for all studies that make it through the screening stage (see Appendix 2.1). Although preliminary, the instrument indicates the types of key questions we will aim to answer for each study. Many of the open-ended questions will then be collated into smaller categories and appropriate tables generated. The forms will ask questions pertaining to type of study (study design), the research question addressed, sample size, methodology used, contextual factors and so on.

Following this stringent process will lead us to a final set of quality studies that provide robust evidence that either supports, counters or is neutral with respect to the different proposed relationships between interventions and outcomes. A diagram to identify the flow of studies will be included in the review, which will map out the process and indicates how the final set of studies will be arrived at to ensure transparency.

2.3.2.2 Process of synthesis; and deriving conclusion and implications

This systematic review includes three distinct questions, directly aligned to different studies types to inform the in-depth review synthesis.

In answering the first: "*what is the evidence on the impacts of reforms of education systems at scale to increase teacher effectiveness on the quality of teaching and on*

learning outcomes in low and middle income countries? we will include studies that measure the effects of interventions using experimental and quasi-experimental study designs. Specifically we will include (1) Studies where participants are randomly assigned to treatment and comparison groups; (2) Studies where assignment to treatment and comparison group is based on other known allocation rules, including a threshold on a continuous variable (regression discontinuity designs) or exogenous geographical variation in the treatment allocation (natural experiments); (3) Studies with non-random assignment to treatment and comparison group, provided they include pre-and post-test measures of the outcome variables of interest to ensure equity between groups on the baseline measure, as well as use appropriate methods to control for selection bias and confounding, such as: statistical matching (e.g. propensity score matching, or covariate matching), regression adjustment (e.g. difference-in-differences, and single difference regression analysis, instrumental variables, and 'Heckman' selection models).

In answering the second question: what evidence is available on the relationship between educational reforms for improving teacher effectiveness at scale and the quality of teaching and learning outcomes in low and middle income countries?, we propose to include studies such as those without random allocation to treatment and comparison group with only post-test measures of the outcome variables but that attempt to use statistical methods to control for selection bias and confounding factors. For example, in studying a contract teacher intervention, it is crucial to understand that contract teachers are often appointed to schools with fewer resources in more remote areas and often serve more disadvantaged children, so any valid estimate of the contract teacher effect must take account of the wider social and economic context in which these contract teachers are employed and also the potential non-random matching of contract teachers to particular children/schools on the basis of unobserved characteristics of both the teachers and the students. Similarly, contract teachers may be systematically assigned to less able children within a school. A study that evaluates this intervention and finds that contract teachers are not as 'effective' as regular government school teachers in imparting learning for instance, may therefore be largely due to the low-ability profile of the students they teach rather than a pure contract teacher effect. It may also be that contract teachers are systematically different in their unobserved characteristics from regular government school teachers. It is therefore very important to control for the observed and unobserved student, school and teacher characteristics in a study that aims to estimate true contract teacher effects. Quantitative studies such as descriptive data analysis will not take this into account as effectively, will be excluded from this review. Only studies that attempt to 'control' for the wider social and economic context and provide generalizability more accurately will be retained in answering this second question (see Kingdon et al. 2013). Quantitative studies that do not effectively control for confounding factors or self-selection will be excluded completely. However, this review also aims to assess grey literature and qualitative literature to ensure comprehensive coverage of specific interventions and studies using case study designs, ethnographies and interviews and focus-groups may be included in the final set for in-depth review in answering the questions provided they meet the stringent quality assessment criteria posed by the reviewers.

The final question: where reforms to education systems to increase teacher effectiveness at scale have occurred, what is the evidence on how technical, financial and political

barriers have been overcome?, will be addressed using evidence from the literature base identified in answering questions 1 and 2.

Results of this in-depth review will be collated to provide a Weight of Evidence on the overall evidence based, directly informing a synthesis of evidence. Once we have identified and assessed the studies, the research will be collated. This synthesis will be presented in the form of structured thematic narrative, summary tables and where possible, quantitative studies will be synthesised using meta-analysis. In addition to a synthesis of the established body of evidence, a main objective of this review will be to identify the key gaps in the literature, and derive policy and practice implications. Studies that do not meet quality assurance procedures may still be discussed in the review as they may nonetheless provide useful insights and context.

The review team will also prepare an Evidence Brief. Depending on the evidence available and the findings of the review, the team will work with DFID and the EPPI centre to produce an Evidence Brief that communicates the findings of this review in the most effective manner.

Proposed Time Line

Protocol Submission	25 September 2014
First Draft	30 April 2015
Final Report and Evidence Brief	30 June 2015

3. References

Acemoglu, D (2010) 'Theory, General Equilibrium and Political Economy in Development Economics', MIT (mimeo).

Annual Status of Education Report (ASER, Pakistan): http://www.aserpakistan.org/

Annual Status of Education Report (ASER, India): http://www.asercentre.org/#6v2dd

AusAid (2012) 'Scaling Up Development Results: A Literature Review and Implications for Australia's Aid Program', Ian Anderson, Commonwealth of Australia, February, 2012.

Barber, M (2013) 'The Good News from Pakistan, How a Revolutionary New Approach to Education Reform in Punjab shows the Way Forward for Pakistan and Development Aid Everywhere', Reform 2013

Bold, T, Kimenyi, M, Mwabu, G, Ng'ang'a, A, and Sandefur, J (2013) 'Scaling up What works: Experimental Evidence on External Validity in Kenyan Education', Centre for Global Development, January 2013.

Burgess, S, Davies, N, Slater, H (2009) 'Do teachers matter? Measuring teacher effectiveness in England', CMPO Working Paper No. 09/212

Clotfelter, CT, Ladd, H and Vigdor, J (2006) 'Teacher-Student Matching and the Assessment of Teacher Effectiveness', NBER Working Paper Series: National Bureau of Economic Research, Cambridge MA

DFID (2014) *How To Note* on *Assessing the Strength of Evidence* London: p.1-21

Filmer, D, Hasan, A and Pritchett, L (2006) "A Millennium Learning Goal: Measuring Real Progress in Education," Working Papers 97, Center for Global Development.

Glewwe, P, Hanushek, EA, Humpage, S and Ravina, R (2013) 'School Resources and Educational Outcomes in Developing Countries: A Review of the Evidence from 1990-2010', NBER Working Paper Series, WP No. 17554, http://www.nber.org/papers/w17554

Goldhaber, D (1999) 'School Choice: an examination of the empirical evidence on achievement, parental decision making and equity', Educational Researcher, 28, (9), 16-25

Hanushek, E. and Woessmann, L (2011) 'Overview of the symposium for performance pay for teachers'. Economics of Education Review, 30, 391-393

Kremer, M, and Holla, A (2009) Improving Education in the Developing World: What Have We Learned from Randomized Evaluations? Annual Review of Economics 1: 513-542.

Kingdon, G (with M. Aslam, S. Rawal and S. Das) (2013) 'Are contract teachers and para-teachers a cost-effective intervention to address teacher shortage and improve learning outcomes?' London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Stewart, R, van Rooyen, C, Korth, M, Chereni, A, Rebelo Da Silva, N, de Wet, T (2012) Do micro-credit, micro-savings and micro-leasing serve as effective financial inclusion interventions enabling poor people, and especially women, to engage in meaningful economic opportunities in low- and middle-income countries. A systematic review of the evidence. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

UWEZO, http://www.uwezo.net/

Appendices

Appendix 1.1: Authorship of this report

Lead review author:

The lead author is the person who develops and co-ordinates the review team, discusses and assigns roles for individual members of the review team, liaises with the editorial base and takes responsibility for the on-going updates of the review.

Name: Monazza Aslam
Title: Dr.
Affiliation: Institute of Education, University of London (Visiting Research Fellow)
Address: 5 Downside End
City, State, Province or County: Oxford
Postal Code: OX3 8JF
Country: United Kingdom
Phone: + 44 (0) 7956970628
Email: monazza.aslam@wolfson.oxon.org

Co-author(s):

Name: Shenila Rawal	
Title: Dr	
Affiliation: University of Bristol (Visiting Researcher)	
Address: Venture House, 42 London Road	
City, State, Province or County:	Staines
Postal Code:TW18 4HF	

Country: UK

Phone:

Email: shenilarawal@aol.com

Co-author(s):

Name: Geeta Kingdon

Title: Professor

Affiliation: Institute of Education, University of London

Address: 3 Zoar Cottages,

Mary Tavy

City, State, Province or County: Tavistock

Postal Code: PL19 9ND

Country: United Kingdom

Phone: +44 (0) 1822 810265

Email: g.kingdon@ioe.ac.uk

Co-author(s):

Name: Bob Moon

Title: Emeritus Professor of Education

Affiliation: The Open University (UK)

Address: c/o 17 Beaumont Buildings

City, State, Province or County: Oxford

Postal Code: OX12LL

Country: United Kingdom

Phone: +44 (0)1865 554497

Email: r.e.moon@open.ac.uk

Co-author(s):

Name: Rukmini Banerji

Title: Dr.

Affiliation: Director, ASER Centre, Pratham Education foundation

Address: B 4/54 Sardarjang Enclave

City, State, Province or County: New Delhi, India

Postal Code: 110029

Country: India

Phone: +91-11-26716084

Email: Rukmini.banerji@pratham.org

Co-author(s):

Name: Sushmita Das

Title: Ms.

Affiliation: Department of Quantitative Social Science, Institute of Education, London, UK

Address: 25 Moat Lodge, London Road, Harrow on the Hill,

City, State, Province or County:
Middlesex

Postal Code: HA1 3LU

Country: UK

Phone: +44 7903257885

Email: sushmita.nalini@gmail.com

Co-author(s):

Name: Manjistha Banerji

Title: Dr.

Affiliation: ASER Centre, India

Address: B- 4/ 54 Safdarjung Enclave

City, State, Province or County: Delhi

Postal Code: 110029

Country: India

Phone: + 91- 11- 46023612

Email: manjistha.banerji@asercentre.org

Monazza Aslam, Shenila Rawal and Geeta Kingdon were the Principal Investigators in a DFID/3ie funded Systematic Review 'Are contract and para-teachers a cost effective intervention to address teacher shortages and improve learning outcomes?' This SR has been published, see: *Aslam M, (with G. Kingdon, S. Rawal and S. Das) (2013), 'Are contract teachers and para-teachers a cost-effective intervention to address teacher shortage and improve learning outcomes?' London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London*. These researchers have also been engaged in 2 rigorous reviews funded by DFID. See: *Aslam M. (with G. Kingdon and S. Rawal) (2013): DFID rigorous literature review of 'Private Schools for the Economically Disadvantaged', DFID, UK*. And *Aslam, M. (with T. Beteille, G. Kingdon, A., Little, , T. Moe, H. Patrinos and S. Rawal) (2013): DFID rigorous literature review of 'The Political Economy of Education Systems in Developing Countries', DFID, UK*.

Dr. Aslam and Dr.Rawal have both obtained training from the EPPI-centre at the Institute of Education, University of London on how to use the EPPI-Reviewer software and on conducting meaningful Systematic Reviews. The EPPI-reviewer software has been used both in conducting the Systematic Review and for undertaking the Political Economy rigorous review. Professor Kingdon is well-versed in the use of meta-analyses to systematically assess literature. The three researchers are also quantitative economists with significant expertise in the use of statistical techniques.

Content:	Professor Geeta Kingdon, Dr.Monazza Aslam, Dr.Shenila Rawal, Professor Bob Moon, Dr.Rukmini Banerji: all have expertise in this area and will contribute in this regard.
Systematic review methods:	**Search terms**: will be arrived at by the Lead Researchers and Advisors. **Search**: Ms.Sushmita Das and a Research Assistant from Pratham (to be confirmed) will undertake the initial searches. Ms. Das has previously worked with some of the team members in undertaking a Systematic Review (see above) and has used the EPPI reviewer. She has significant expertise in this area. The Research Assistant will work in collaboration with her to undertake the searches and will be closely guided by Dr. Aslam and Dr.Rawal (who will consult the other team members closely to ensure full coverage of the literature base). **Review (screening and in-depth review)**: This will mainly be the responsibility of Dr. Aslam and Dr.Rawal with guidance from the advisers and with input from Professor Kingdon. **Write-up**: Mainly undertaken by Dr. Aslam and Dr.Rawal with input from the rest of the team members.
Statistical analysis (if relevant):	Professor Kingdon has significant expertise in this regard and will provide inputs in this aspect of research while closely guiding Dr. Aslam and Dr.Rawal.
Information retrieval:	**Information retrieval**: retrieving the documents etc. to be uploaded onto the EPPI reviewer will be the responsibility of Ms. Das and the RA at Pratham. **Content retrieval**: retrieving information from studies will be undertaken by Dr. Aslam and Dr.Rawal with guidance provided by advisors.

Conflicts of interest

All team members in this review have been engaged in research on teacher effectiveness. Professor Bob Moon is a founding director of TESSA and all of the researchers have been significantly involved in primary research in the area of teacher effectiveness. There are no financial conflicts of interest.

Appendix 2.1: Inclusion and exclusion criteria

Table 1: PICOST inclusion/exclusion criteria for defining studies (quantitative and qualitative)

	Included	Excluded
Population	Lesser-developed countries Middle-income countries Primary and secondary school children in government schools General schooling	High-income countries Transition economies Upper middle-income countries Tertiary schooling Private or aided school children Vocational and technical education, Non-formal education
Intervention	Teacher effectiveness interventions (such as contract teacher schemes, pre-service training, in-service training, merit pay, computer assisted teaching and learning etc.)	Any interventions that do not target teacher effectiveness directly: for example, class size reduction interventions, community involvement schemes etc.
Comparison	**Question 1 must include a comparison group**: For example, treated teachers versus non-treated teachers or, students taught by treated teachers vs. those taught by non-treated teachers. **Question 2, if there is no comparison group, but studies provide data on the intervention and outcomes (see below)**	
Outcome	Academic achievement tests (learning outcomes) Teacher quality (time on task, teacher motivation, competence, absence, skills, effort, qualifications, credentials, teacher test scores, and characteristics etc.)	Self-reported happiness, measures of well-being Non-cognitive scores School enrolment, attendance, completion, transition.

Appendix 2.2: Search strategy

Concepts & Search Terms

Five separate concepts are identified in order to construct the search strategy and manage the search terms. The main concept (i.e. the main "input") here is **Interventions/Reforms that influence teacher quality** and student outcomes.

Concept 1: Intervention/Reforms:

academic reform(s), academic intervention(s), academic incentive(s), academic initiative(s) academic program(s), academic scheme(s), child reform(s), child intervention(s), child incentive(s), child initiative(s), child program(s), child scheme(s), classroom reform(s), classroom intervention(s), classroom incentive(s), classroom initiative(s) classroom program(s), classroom scheme(s), education reform(s), education intervention(s), education incentive(s),), education initiative(s) education program(s), education scheme(s),learning reform(s), learning intervention(s), learning incentive(s), learning initiative(s), learning program(s), learning scheme(s), pupil reform(s), pupil intervention(s), pupil incentive(s), pupil initiative(s), pupil program(s), pupil scheme(s), school reform(s), school intervention(s), school incentive(s), school initiative(s), school program(s), school scheme(s), student reform(s), student intervention(s),student incentive(s), student initiative(s), student program(s), student scheme(s),

teacher reform(s), teacher intervention(s), teacher incentive(s), teacher initiative(s), teacher program(s), teacher scheme(s), teaching reform(s), teaching reform(s), teaching intervention(s), teaching incentive(s), teaching initiative(s), teaching program(s), teaching scheme(s), teacher training reform(s), teacher education reform (s) teacher training intervention(s), teacher education intervention (s), teacher training incentive(s), teacher education initiative (s), teacher training initiative(s), teacher training program(s), teacher education program (s), teacher training scheme(s), teacher education scheme (s) teacher pay reform(s), teacher pay intervention(s), teacher pay incentive(s), teacher pay initiative(s), teacher pay program(s), teacher pay scheme(s)

Concept 2: Teacher Quality

pedagogical improvement(s), pedagogical method(s), pedagogical resource(s), pedagogical skill(s), pedagogical strategy(ies), pedagogical style(s), teacher absenteeism, teacher attendance, teacher accountability, teacher competence, teacher content knowledge, teacher characteristics, teacher development, teacher effectiveness, teacher effort(s), teacher motivation(s), teacher method(s), teacher practice(s), teacher resources, teacher subject knowledge, teacher skill(s), teacher observation(s), teacher quality teaching competence, teaching characteristics, teaching development, teaching effort, teaching method(s), teaching practice(s), teaching resource(s), teaching strategy(ies), teaching skill(s), teaching observation(s), teaching quality,

Concept 3: Student Outcomes – *for this concept the search terms are a mix of aspects of student outcomes and synonyms of "student" and "outcomes".*

academic achievement(s), academic attainment, academic assessment(s), academic attendance, academic evaluation(s), academic enrolment, academic performance(s), academic progress, academic skill(s), academic test(s), academic test score(s) academic mark(s), academic result(s), academic retention, academic outcome(s)

child achievement(s), child attainment, child assessment(s), child attendance, child evaluation(s), child enrolment, child performance(s), child progress, child schooling, child skill(s), child test(s), child test score(s), child mark(s), child result(s), child retention, child outcome(s), classroom achievement(s), classroom attainment, classroom assessment(s), classroom attendance, classroom evaluation(s), classroom performance(s), classroom progress, classroom skill(s), classroom test(s), classroom test score(s), classroom mark(s), classroom result(s), classroom retention, classroom outcome(s), cognitive achievement(s), cognitive attainment, cognitive assessment(s), cognitive performance(s), cognitive progress, cognitive skill(s), cognitive test(s), cognitive test score(s), cognitive mark(s), cognitive result(s), cognitive retention, cognitive outcome(s), education achievement(s), education attainment, education assessment(s), education attendance, education evaluation(s), education enrolment, education performance(s), education progress, education test(s), education test score(s) education mark(s), education result(s), education retention, education outcome(s), learning achievement(s), learning attainment, learning assessment(s), learning performance(s), learning progress, learning skill(s), learning test(s), learning test score(s), learning mark(s), learning result(s), learning outcome(s), pupil achievement(s), pupil attainment, pupil assessment(s), pupil attendance, pupil evaluation(s), pupil enrolment, pupil performance(s), pupil progress, pupil test(s), pupil test score(s) pupil mark(s), pupil result(s), pupil retention, pupil outcome(s), scholastic achievement(s), scholastic attainment, scholastic assessment(s), scholastic evaluation(s), scholastic performance(s), scholastic progress, scholastic skill(s), scholastic test(s), scholastic test score(s), scholastic mark(s), scholastic result(s), scholastic retention, scholastic outcome(s), student achievement(s), student attainment, student assessment(s), student attendance, student evaluation(s), student enrolment, student performance(s), student progress, student test(s), student test score(s), student mark(s), student result(s), student retention, student outcome(s)

Concept 4: Political economy issues – *for this concept search terms include aspects of political economy prefixed with synonyms for reforms/interventions where possible*.

advanc(ing) reform(s), advanc(ing) intervention(s), advanc(ing) incentive(s), advanc(ing) initiative(s) advanc(ing) program(s), advanc(ing) scheme(s), allow(ing) reform(s), allow(ing) intervention(s), allow(ing) incentive(s), allow(ing) initiative(s) allow(ing) program(s), allow(ing) scheme(s), assist(ing) reform(s), assist(ing) intervention(s), assist(ing) incentive(s), assist(ing) initiative(s) assist(ing) program(s), assist(ing) scheme(s), authoriz(ing) reform(s), authoriz(ing) intervention(s), authoriz(ing) incentive(s), authoriz(ing) initiative(s) authoriz(ing) program(s), authoriz(ing) scheme(s), block(ing) reform(s), block(ing) intervention(s), block(ing) incentive(s), block(ing) initiative(s) block(ing) program(s), block(ing) scheme(s), enabl(ing) reform(s), enabl(ing) intervention(s), enabl(ing) incentive(s), enabl(ing) initiative(s) enabl(ing) program(s), enabl(ing) scheme(s), encourag(ing) reform(s), encourag(ing) intervention(s), encourag(ing) incentive(s), encourag(ing) initiative(s) encourag(ing) program(s), encourag(ing) scheme(s), implement(ing) reform(s), implement(ing) intervention(s), implement(ing) incentive(s), implement(ing) initiative(s) implement(ing) program(s), implement(ing) scheme(s), oppose(ing) reform(s), oppose(ing) intervention(s), oppose(ing) incentive(s), oppose(ing) initiative(s) oppose(ing) program(s), oppose(ing) scheme(s),

partner(ing) reform(s), partner(ing) intervention(s), partner(ing) incentive(s), partner(ing) initiative(s) partner(ing) program(s), partner(ing) scheme(s),

resist(ing) reform(s), resist(ing) intervention(s), resist(ing) incentive(s), resist(ing) initiative(s) resist(ing) program(s), resist(ing) scheme(s), reinforce reform(s), reinforce intervention(s), reinforce incentive(s), reinforce initiative(s) reinforce program(s), reinforce scheme(s), support(ing) reform(s), support(ing) intervention(s), support(ing) incentive(s), support(ing) initiative(s) support(ing) program(s), support(ing) scheme(s), sanction(ing) reform(s), sanction(ing) intervention(s), sanction(ing) incentive(s), sanction(ing) initiative(s) sanction(ing) program(s), sanction(ing) scheme(s), politics, political, political economy, politics of education, politicization of education, politics of schools, politics of teachers, politicization of teachers, teacher strike(s), teacher unions, teacher organizations, education and principal agent theory teacher(s) and principal agent theory, rent seeking and education

Concept 5: Countries

Afghan* OR Armen* OR Bangladesh* OR Benin* OR Bhutan* OR Burkina Faso* OR Burund* OR Bolivia* OR Cambodia* OR Cameroon* OR Verde* OR Central African Republic OR Chad* OR Comoros OR Congo* OR Côte d'Ivoire OR Ivory Coast OR Djibouti* OR Eritrea* OR Ethiopia* OR Egypt* OR El Salvador* OR Georgia* OR Gambia* OR Ghan* OR Guinea* OR Guatemal*OR Haiti* OR Hondura* OR Guyan*OR India* OR Indonesia* OR Kenya* OR Kiribati* OR Kyrgyz* OR Lao* OR Kosov* OR Lesotho OR Liberia* OR Madagasca* OR Malawi* OR Mali* OR Marshall Islands OR Mauritania* OR Micronesia* OR Moldova* OR Mongoli*OR Mozambi* OR Moroc* OR Nepal* OR Nicaragua* OR Niger* OR Myanmar OR Pakistan* OR Papua New Guinea* OR Paraguay* OR Philippin* OR Rwanda* OR Samoa* OR São Tomé and Principe OR Senegal* OR Sierra Leon* OR Solomon Islands OR Somalia* OR Sudan* OR Swazi* OR Syria* OR Sri Lank* OR Tajik* OR Tanzania* OR Timor-Leste OR Togo* OR Tonga* OR Ukrain* OR Palestin* OR West Bank OR Gaza OR Turkmenistan* OR Tuvalu* OR Uganda* OR Uzbek* OR Vanuatu* OR Vietnam* OR Yemen* OR Zambia* OR Zimbabwe*

[Note: * indicates truncation eg. for Bangladesh and Bangladeshi]

Search Strings

A search of concept 1 and 5 alone yielded an un-manageable number of hits. Search strings have been constructed such that they combine concepts.

#CONCEPT 1 AND CONCEPT2 AND CONCEPT 5

((academic OR child* OR classroom OR education OR learning OR pupil* OR school* OR student* OR teach*) AND (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative***)) AND ((**teach*) AND (improvement* OR method* OR quality OR resource* OR skill OR style OR strategies OR practice* OR effectiveness OR observation* OR absenteeism OR attendance OR accountability OR competence OR knowledge**)) AND** ((Afghan* OR Armen* OR Bangladesh* OR Benin* OR Bhutan* OR Burkina Faso* OR Burund* OR Bolivia*

OR Cambodia* OR Cameroon* OR Verde* OR Central African Republic OR Chad* OR Comoros OR Congo* OR Côte d'Ivoire OR Ivory Coast OR Djibouti* OR Eritrea* OR Ethiopia* OR Egypt* OR El Salvador* OR Georgia* OR Gambia* OR Ghan* OR Guinea* OR Guatemal*OR Haiti* OR Hondura* OR Guyan*OR India* OR Indonesia* OR Kenya* OR Kiribati* OR Kyrgyz* OR Lao* OR Kosov* OR Lesotho OR Liberia* OR Madagasca* OR Malawi* OR Mali* OR Marshall Islands OR Mauritania* OR Micronesia* OR Moldova* OR Mongoli*OR Mozambi* OR Moroc* OR Nepal* OR Nicaragua* OR Niger* OR Myanmar OR Pakistan* OR Papua New Guinea* OR Paraguay* OR Philippin* OR Rwanda* OR Samoa* OR São Tomé and Principe OR Senegal* OR Sierra Leon* OR Solomon Islands OR Somalia* OR Sudan* OR Swazi* OR Syria* OR Sri Lank* OR Tajik* OR Tanzania* OR Timor-Leste OR Togo* OR Tonga* OR Ukrain* OR Palestin* OR West Bank OR Gaza OR Turkmenistan* OR Tuvalu* OR Uganda* OR Uzbek* OR Vanuatu* OR Vietnam* OR Yemen* OR Zambia* OR Zimbabwe*)) **16 hits**

#CONCEPT 1 AND CONCEPT3 AND CONCEPT 5

((academic OR child* OR classroom OR education OR learning OR pupil* OR school* OR student* OR teach*) AND (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)) AND (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative***)) AND ((**academic OR child* OR classroom OR cognitive OR education OR grade OR learning OR pupil* OR schola* OR student*) AND (achievement* OR attainment OR assessment* OR attendance OR evaluation* enrolment* OR performance* OR progress OR skill* test* OR test-score* OR mark* OR result* or retention OR outcome***)) AND** ((Afghan* OR Armen* OR Bangladesh* OR Benin* OR Bhutan* OR Burkina Faso* OR Burund* OR Bolivia* OR Cambodia* OR Cameroon* OR Verde* OR Central African Republic OR Chad* OR Comoros OR Congo* OR Côte d'Ivoire OR Ivory Coast OR Djibouti* OR Eritrea* OR Ethiopia* OR Egypt* OR El Salvador* OR Georgia* OR Gambia* OR Ghan* OR Guinea* OR Guatemal*OR Haiti* OR Hondura* OR Guyan*OR India* OR Indonesia* OR Kenya* OR Kiribati* OR Kyrgyz* OR Lao* OR Kosov* OR Lesotho OR Liberia* OR Madagasca* OR Malawi* OR Mali* OR Marshall Islands OR Mauritania* OR Micronesia* OR Moldova* OR Mongoli*OR Mozambi* OR Moroc* OR Nepal* OR Nicaragua* OR Niger* OR Myanmar OR Pakistan* OR Papua New Guinea* OR Paraguay* OR Philippin* OR Rwanda* OR Samoa* OR São Tomé and Principe OR Senegal* OR Sierra Leon* OR Solomon Islands OR Somalia* OR Sudan* OR Swazi* OR Syria* OR Sri Lank* OR Tajik* OR Tanzania* OR Timor-Leste OR Togo* OR Tonga* OR Ukrain* OR Palestin* OR West Bank OR Gaza OR Turkmenistan* OR Tuvalu* OR Uganda* OR Uzbek* OR Vanuatu* OR Vietnam* OR Yemen* OR Zambia* OR Zimbabwe*)) **235 hits**

#CONCEPT 1 AND CONCEPT4 AND CONCEPT 5

((academic OR child* OR classroom OR education OR learning OR pupil* OR school* OR student* OR teach*) AND (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)) **AND ((**advanc* OR authoriz* OR assist* OR block* OR resist* OR enabl* OR encourag*OR oppos* OR advanc* OR partner* OR reinforc* OR support* OR sanction* OR implement*) AND (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative***)) AND** ((Afghan* OR Armen* OR Bangladesh* OR Benin* OR Bhutan* OR Burkina Faso* OR Burund* OR Bolivia* OR Cambodia* OR Cameroon* OR Verde* OR Central African Republic OR Chad* OR Comoros OR Congo* OR Côte d'Ivoire OR Ivory Coast OR Djibouti* OR Eritrea* OR Ethiopia* OR Egypt* OR El Salvador* OR Georgia* OR Gambia* OR Ghan* OR

Guinea* OR Guatemal*OR Haiti* OR Hondura* OR Guyan*OR India* OR Indonesia* OR Kenya* OR Kiribati* OR Kyrgyz* OR Lao* OR Kosov* OR Lesotho OR Liberia* OR Madagasca* OR Malawi* OR Mali* OR Marshall Islands OR Mauritania* OR Micronesia* OR Moldova* OR Mongoli*OR Mozambi* OR Moroc* OR Nepal* OR Nicaragua* OR Niger* OR Myanmar OR Pakistan* OR Papua New Guinea* OR Paraguay* OR Philippin* OR Rwanda* OR Samoa* OR São Tomé and Principe OR Senegal* OR Sierra Leon* OR Solomon Islands OR Somalia* OR Sudan* OR Swazi* OR Syria* OR Sri Lank* OR Tajik* OR Tanzania* OR Timor-Leste OR Togo* OR Tonga* OR Ukrain* OR Palestin* OR West Bank OR Gaza OR Turkmenistan* OR Tuvalu* OR Uganda* OR Uzbek* OR Vanuatu* OR Vietnam* OR Yemen* OR Zambia* OR Zimbabwe*))

23 hits

Total Number of hits from Econpapers = 274 hits.

The majority of searches will take place online using key word searches on databases such as those indicated below. In addition to the published literature, other sites such as REPEC will be searched for working papers, the Conference Proceedings Citation Index – Social Science and Humanities (CPCI-SSH) and Index of Conference Proceedings will be searched for conference papers and the Dissertation Abstracts database searched for PhD. and Masters Dissertations. Grey literature will also be sourced through manual searches and may be provided by DFID. We will also carry out manual searches to ensure comprehensive coverage of the entire literature base (grey + non-grey). Table 1 below depicts the key databases to be searched.

Appendix 2.3: Example of data extraction form

Title of Study:

Type of study (dissertation, journal article, book chapter etc.)

Authors:

Publication Date:

Purpose of Study:

Type of intervention:

Context/setting:

At scale? (extent of intervention, discuss):

Methodology:

Outcomes measured:

Findings:

Research Question addressed?

Were there any technical, financial or political economy factors that hindered or enhanced the intervention?

Quality Assurance (include here limitations of study):

Any additional/related issues that arise that may be interesting/relevant for the readers of this SR:

Appendix 2.4: Assessing the quality of evidence: example form

Please refer to the DFID *How To Note* on *Assessing the Strength of Evidence*, March 2014, p.10-13 for explanations of terms.

Principles of quality	Associated principles	YES/NO
Conceptual framing	Does the study acknowledge existing research?	
	Does the study construct a conceptual framework?	
	Does the study pose a research question?	
	Does the study outline a hypothesis?	
Transparency	Does the study present or link to the raw data it analyses?	
	What is the geography/context in which the study was conducted?	
	Does the study declare sources of support/funding?	
Appropriateness and rigour	Does the study identify a research design?	
	Does the study identify a research method?	
	Does the study demonstrate why the chosen design and method are good ways to explore the research question?	
Cultural sensitivity	Does the study explicitly consider any context-specific cultural factors that may bias the analysis/findings?	
Validity	To what extent does the study demonstrate measurement validity?	
	To what extent is the study internally valid?	
	To what extent is the study externally valid?	
	To what extent is the study ecologically valid?	
Reliability	To what extent are the measures used in the study stable?	
	To what extent are the measures used in the study internally reliable?	
	To what extent are the findings likely to be sensitive/changeable	
Cogency	Does the author 'signpost' the reader throughout?	

	To what extent does the author consider the study's limitations and/or alternative interpreations of the analysis?	
	Are the conclusions clearly based on the study's results?	

(Source: DFID, 2014, *How To Note* on *Assessing the Strength of Evidence.*, p.14)

When you have completed the checklist in Table 2, use the following table to grade the quality of the study.

Table 3

Study quality	Abbreviation	Definition
High	↑	Demonstrates adherence to principles of appropriateness/rigour, validity and reliability; likely to demonstrate principles of conceptual framing, openness/ transparency and cogency
Moderate*	→	Some deficiencies in appropriateness/rigour, validity and/or reliability, or difficulty in determining these; may or may not demonstrate principles of conceptual framing, openness/transparency and cogency
Low	↓	Major and/or numerous deficiencies in appropriateness/rigour, validity and reliability; may/may not demonstrate principles of conceptual framing, openness/ transparency and cogency

(Source: DFID, 2013, *How To Note* on *Assessing the Strength of Evidence.*, p.15)

First produced in 2014 by:

Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)
Social Science Research Unit
Institute of Education, University of London
18 Woburn Square
London WC1H 0NR

Tel: +44 (0)20 7612 6397

http://eppi.ioe.ac.uk/
http://www.ioe.ac.uk/ssru/

The **Evidence for Policy and Practice Information and Co-ordinating Centre** (EPPI-Centre) is part of the Social Science Research Unit (SSRU), Institute of Education, University of London.

The EPPI-Centre was established in 1993 to address the need for a systematic approach to the organisation and review of evidence-based work on social interventions. The work and publications of the Centre engage health and education policy makers, practitioners and service users in discussions about how researchers can make their work more relevant and how to use research findings.

Founded in 1990, the Social Science Research Unit (SSRU) is based at the Institute of Education, University of London. Our mission is to engage in and otherwise promote rigorous, ethical and participative social research as well as to support evidence-informed public policy and practice across a range of domains including education, health and welfare, guided by a concern for human rights, social justice and the development of human potential.

This document is available in a range of accessible formats including large print. Please contact the Institute of Education for assistance:

telephone: +44 (0)20 7947 9556 email: info@ioe.ac.uk