© IFAD/Franco Mattioli

# REFORMS TO INCREASE TEACHER EFFECTIVENESS IN DEVELOPING COUNTRIES

SYSTEMATIC REVIEW, SEPTEMBER 2016

**The authors of this report are:**

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 (ASER Centre, Pratham Education Foundation).

**Conflicts of interest**

There were no conflicts of interest in the writing of this report.

**Contributions**

The opinions expressed in this publication are not necessarily those of the EPPI-Centre/3IE or the funders. Responsibility for the views expressed remains solely with the authors.

**Citation**

This report should be cited as: Aslam M, Rawal S, Kingdon G, Moon, B., Banerji R, Das S, Banerji M, Sharma SK (2016) *Reforms to Increase Teacher Effectiveness in Developing Countries*. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London

**Picture**

© IFAD/Franco Mattioli

## SUMMARY

We sought high-quality evidence on reforms/interventions in education systems aimed at improving teacher effectiveness, at scale. This executive summary provides an overview of that key evidence to answer three review questions:

*RQ1. 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?*

*RQ2. 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?*

*RQ3. 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*?

Despite indications that teachers constitute one of the most critical institutional inputs into a child's learning experience, robust evidence that examines the relationship between teacher effectiveness and pupil outcomes tends to examine the *relationship* between teacher effectiveness and pupil outcomes in a non-causal manner. In light of this, this review has chosen two research questions (RQ1 and RQ2) instead of just one (RQ1), so that this non-causal evidence base may also be presented. The 'at-scale' nature of reforms being investigated in this review potentially makes them more vulnerable to political-economy factors, which may either hinder or promote the design or implementation, and, consequently, the efficacy of the reforms and, ultimately, their outcomes. Hence, RQ3 aims to uncover any evidence in this regard.

It is also worth noting that the availability and nature of teacher-training programmes, the quality of the existing and potential pool of teachers, the format and nature of existing teacher contracts, the pay-scale of existing teachers, etc., may differ across countries (and even across regions within one country), resulting in different outcomes for potentially very similar interventions. The evidence presented in this review is deeply contextual and this brief is not designed to provide specific advice on which interventions are more or less appropriate in particular contexts, but, rather, to summarise what is known in response to these questions, by outlining some of the key contextual factors, such as the role of teachers' unions; the knowledge base and power of stakeholders; bureaucratic and institutional factors in place due to historic reasons (for instance, the status of teachers within a country), and so on.

## APPROACH

We searched topic-specific bibliographic databases and websites to identify studies relevant to answering each of the three review questions on effectiveness, the relationship between interventions and outcomes, and studies on the contextual factors that may aid or hinder the efficacy of these reforms. After application of the inclusion and exclusion criteria, we built a 'map' of studies to provide a descriptive overview of the types of studies on teacher-reform interventions, at scale and not at scale (n=27). We then synthesised evidence on the 15 at-scale studies, in order to answer the in-depth review questions. As using a quantified definition of 'scale' may have limited our research, a wide range of 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., were considered when classifying a study as at-scale.

## SUMMARY MAP OF EVIDENCE

As stated, a total of 27 studies were identified; 15 studies reported on the impact of reforms at-scale and 12 studies not at-scale. The majority of studies focused on interventions relating to contract teachers (n=10) and monetary incentives for teachers (n=5). Evidence reported was mostly quantitative (n= 26), rather than qualitative, in nature (n=1), with the geographical spread of the studies covering Africa and Asia, as well as parts of Latin America.

## OUTLINE OF EVIDENCE

The review initially embarked on developing a theoretical framework, within which we aimed to capture the potential relationships between teacher-effectiveness reforms and student outcomes, as specified in the three research questions mentioned above. A key aspect of embedding these questions into a *theory of change* is seeking to identify the different chains and assumptions that underpin the relationships between teacher effectiveness and the outcomes of interest. Most importantly, this review aims to identify clearly where the evidence is especially strong to support a given assumption/chain and, therefore, where we can identify robust evidence to support a causal relationship, as identified within the research questions. This review places equal importance on shedding light on those assumptions/chains within the theoretical framework that are not well supported by the evidence and, consequently, do not allow strong causal relationships to be assumed.

The outcomes of interest that are examined in this research paper are learning outcomes and teacher quality. The theoretical framework developed during the course of this review process indicates that changes in teaching quality are also a route that, if taken, may affect student-learning outcomes, as well as being an outcome of interest as regards the research question. However, the ultimate policy goal is to improve learning outcomes and, to reflect this, the main outcome of interest, for the purposes of this review, is the assessment of changes in learning outcomes that result from policy interventions. Additionally, recognising the fact that evaluations are often limited in their ability to show causation between

interventions and student outcomes (due to data limitations, insufficient time having elapsed since the implementation of the policy, etc.), as well as robust existing literature showing the importance of teacher quality in impacting student outcomes, gives an additional argument as to why teacher quality should be included as an outcome in and of its own right.

Our stringent and iterative process resulted in an in-depth review of 15 studies. The summary of findings of these 15 studies report on four key interventions (summarised in detail below):

i) contract teachers;
ii) monetary incentives;
iii) teacher certification; and
iv) teacher training.

In assessing the overall strength of evidence, we have used the **quality** (based on assessment of cogency, reliability and methodology; see Appendix 2.4 for further details) of individual studies constituting the body of evidence, the **size** of the body of evidence (whether it is large: 30 studies or more; medium: 10-30 studies; or, small: fewer than 10 studies), the **context** they cover (global or context-specific) and the **consistency of findings** (a range of studies pointing to identical or similar conclusions versus different studies pointing to different findings) to conclude whether our review shows 'strong', 'modest' or 'insufficient' evidence of the relationship being studied.

## CONTRACT TEACHERS

- Eight studies (two high-quality, and the remainder moderate; only two studies were able to show impact) provide evidence of the relationship between contract teachers and (i) teacher attendance and (ii) student learning, measured using test scores. These studies cover a number of country contexts, including India, Pakistan, Kenya, Niger, Mali and Togo.
- There are two studies addressing RQ1, and both show a positive impact of contract-teacher interventions on student achievement. One of the studies also shows a positive impact of contract-teacher policy on teacher quality. The results from Muralidharan and Sundararaman (2013) suggest that the gains are broadly distributed among all students. However, those children in the more remote areas do appear to receive more benefits (context: India). Bold et al. (2012) find that, while implementation of the contract-teacher policy by an NGO in Kenya resulted in positive effects for English and Mathematics, there were no such effects when that same intervention was implemented by the Government.
- There are six studies that investigate the relationship between contract-teacher interventions and the quality of teaching and learning outcomes (RQ2). Overall, the findings are mixed. However, the evidence on contract-teacher reforms, as examined by these studies, appears to suggest that, in most cases, contract teachers do not perform any less well than regular teachers, and sometimes perform better

in relation to student performance, as measured by test scores. In terms of teaching quality, the evidence appears to indicate that, by and large, when compared to their regular counterparts, contract teachers exert more effort as measured by absence rates, as well as teaching activity in school. This increase in effort appears to be temporary, with research indicating that contract teachers provide diminishing returns through their display of lower effort levels in subsequent contract periods than in their initial contract periods.

- Findings show that results are context-specific, as well as being determined by the characteristics of the reform and its implementation.
- The evidence indicates that the pathways through which contract-teacher interventions can affect both outcomes include factors such as improved incentives, improved accountability, lower social distance, reducing the instances of multi-grade teaching and increased empowerment of local communities and school management.

## MONETARY INCENTIVES

- There are five studies (all of moderate quality), none showing impact (therefore, all answer RQ2).
- The studies cover the following country contexts: Chile, Pakistan, Mexico and The Gambia.
- The evidence of monetary incentives on student outcomes (four studies) is mixed, with two of the four studies showing a positive relationship and two showing a negative relationship.
- In respect of teacher quality (one study), in one context (The Gambia), the evidence is positive.
- Two of the five studies examine the SNED in Chile. This intervention rewards schools based on their pupils' performance. Both studies show the positive effects of the programme for a subset of schools (in particular, those closest to the cut-off point for attaining the award). The third study examines the Carrera Magisterial programme in Mexico, which was a teacher-wage reform that included wage increases for those teachers whose students performed well. The study does not find evidence of a positive impact of these monetary incentives on student achievement. The fourth study, which examines Foundation Assisted Schools in Pakistan, assesses the effectiveness of conditional cash subsidies to low-cost private schools. The paper takes the dual approach of examining whether positive incentives (group bonuses) or negative incentives (removal from the programme) can induce improvements in learning outcomes. The authors find that this set of incentives only leads to the maintenance of minimum levels of learning to remain in the programme. The final study in this category examines the Gambian Hardship Allowance, which aimed to improve the provision of teachers in remote rural locations, through salary incentives. There appears to be a suggestion emerging from this latter piece of evidence that, while the reform was generally 'successful', it did not succeed in reaching the most remote parts of the country.

- The studies support the view that monetary incentives could encourage different attitudes, as well as improve supply and deployment, while, at the same time, their positive effect could be hindered by corruption.

## TEACHER CERTIFICATION

- One (high-quality) study showing impact (answering RQ1).
- Covers the Indonesian context.
- Chang et al. (2014) show no evidence of teacher certification improving student outcomes or teacher quality. This piece of research focuses on the Teacher Law reform (2005) and, in particular, the aspects of that reform aimed at improving the status of the teaching profession. This process was meant to improve the quality of teaching, not only by aiming to attract a better cadre of entrants into the profession, but also by aiming to improve the skills and behaviours of those already in it through improvements in teachers' recognition and a doubling of their salaries. While, theoretically, this was intended to motivate teachers to become more productive, the authors argue that the nepotistic and non-meritocratic fashion in which this reform was initially implemented marred any potential gains in student outcomes and teacher quality.
- The studies within the book (Chang et al. 2014) confirm that, while certification can potentially improve the status of the teaching profession, as well as ensuring minimum standards, this is only the case if those reforms are appropriately implemented.

## TEACHER TRAINING

- One (high-quality) study covering the Ethiopian context and answering RQ2.
- The study examines an in-service teacher-training programme and finds that the programme was particularly effective in improving test scores (especially for girls).
- Whilst trained teachers displayed greater knowledge, they did not appear to adopt more student-centered pedagogy as compared to untrained teachers. The training did not appear to change the range of pedagogical methods used, however, the trained teachers appeared to make more effective choices of which methods to use and apply during teaching. This would suggest that it is not the methodologies that teachers use per se but how they use them that ultimately impact student learning.

## RESEARCH GAPS

- High-quality experimental evaluations on the impact of all four interventions at scale, in different contexts, are required to build a more extensive picture of the modest evidence found and to identify whether the impact of contract teachers is felt in other contexts and what, in the design or context, may contribute to any differences observed in their effectiveness.

- Further research is needed into the barriers and facilitators to implementing reforms and interventions, preferably nested within high-quality experimental evaluations of the impacts of those reforms and interventions, to increase teacher effectiveness, at scale, are required across all contexts, for all intervention types.

- Further exploration of the extent to which process results in change, how they are similar and/or vary from context to context, across student populations for different outcomes would also support the design and implementation of future teacher reforms and implementations.

- Evaluations would also benefit from greater collaboration between policymakers and researchers to ensure more policy-driven research is undertaken to investigate reforms and interventions already implemented. This could include multi-site evaluations to explore single intervention types 'across' contexts to support possible best practice and understand the extent to which findings may be generalised across those contexts.

# CONTENTS

# 1   BACKGROUND

## OUTLINE OF CHAPTER

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 CURRENT 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 RCT studies in various country contexts indicate that simply supplying more resources (more teachers or textbooks) is not a panacea (Kremer and Holla 2009, Glewwe et al. 2013). 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 produce higher student achievement at low cost (Kremer and Holla 2009, Glewwe et al. 2013).

This review aims to identify high-quality evidence pertaining to reforms/interventions in education systems aimed at improving teacher effectiveness. The ultimate objective of such reforms can be argued to improve the overall quality of teaching, with the end-goal of improving student achievement. Effective education systems fundamentally build on good governance, robust public financial management and, inevitably, the effective management of teachers (including recruitment, training and deployment) (DFID 2013). The major challenge in increasing teacher effectiveness lies in recruiting and training competent teachers and significantly improving the effectiveness of teachers already in post. There is, however, a strong sense that successful development involves taking good ideas and practice 'to scale' (AusAID 2012). While the body of literature on teacher effectiveness and interventions to improve said effectiveness is large, as expected, the literature on the effectiveness of such reforms, at scale, is more limited and harder to find. This is a critical shortcoming in light of the fact that educational goals are better attained on a widespread basis through large-scale interventions, the results of which are potentially driven by very different factors than can be observed even in the most successful small-scale projects. Unfortunately, the success of small-scale interventions has not been reproduced on a larger scale. Therefore, it is not only important to identify at-scale interventions, but it is crucial also to ascertain the factors (be they contextual, political or other) that drive or hinder the design and implementation of these interventions, along with their resultant outcomes. This review aims to synthesise evidence investigating these aspects to answer the following overall review question: **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?** As mentioned below (section 1.5), this review question will be split into three sub-questions, to ensure comprehensive coverage of the literature, which will allow meaningful conclusions to be drawn for future policy and research.

## 1.2 DEFINITIONAL AND CONCEPTUAL ISSUES

### THE INTERVENTION: TEACHER

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). The training of school leaders to promote teacher effectiveness could be located in the middle of the continuum.

The scope of this review encompasses a wide range of interventions that fall at various points along this spectrum. Some examples of interventions to improve teacher effectiveness, at scale ,include (but are not limited to):

- **Contract-teacher schemes.** For example, those undertaken in many parts of Africa and Asia. While the narrative of these schemes is to overcome teacher shortages, by improving teacher accountability they are seen to improve both teacher effectiveness and student outcomes (see Kingdon et al. (2013) for a systematic review and Bold et al. (2013), for a study of the scaling up of contract teachers in Kenyan primary schools).

- **Teacher-training and education schemes**. For example, Teacher Education in Sub-Saharan Africa (TESSA), which provides online teacher training/education and resources to teachers or the rolling out of INSET training under the SSA in India;

- **Teacher community-assistant programs.** For example, the Ghana government's Teacher Community Assistant Initiative (TCAI), aimed at improving literacy and numeracy levels in basic schools with a view to 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.

In other countries, attempts to scale up reform programmes based on impact evaluations have often been hindered by political-economy factors (Acemoglu 2010), or have been aided

by strong political will, or stymied by technical barriers and financial constraints (Bold et al. 2013). Studies investigating these different schemes may have sought to evaluate the technical, financial and political barriers and 'drivers of change' relevant to implementing these reforms.
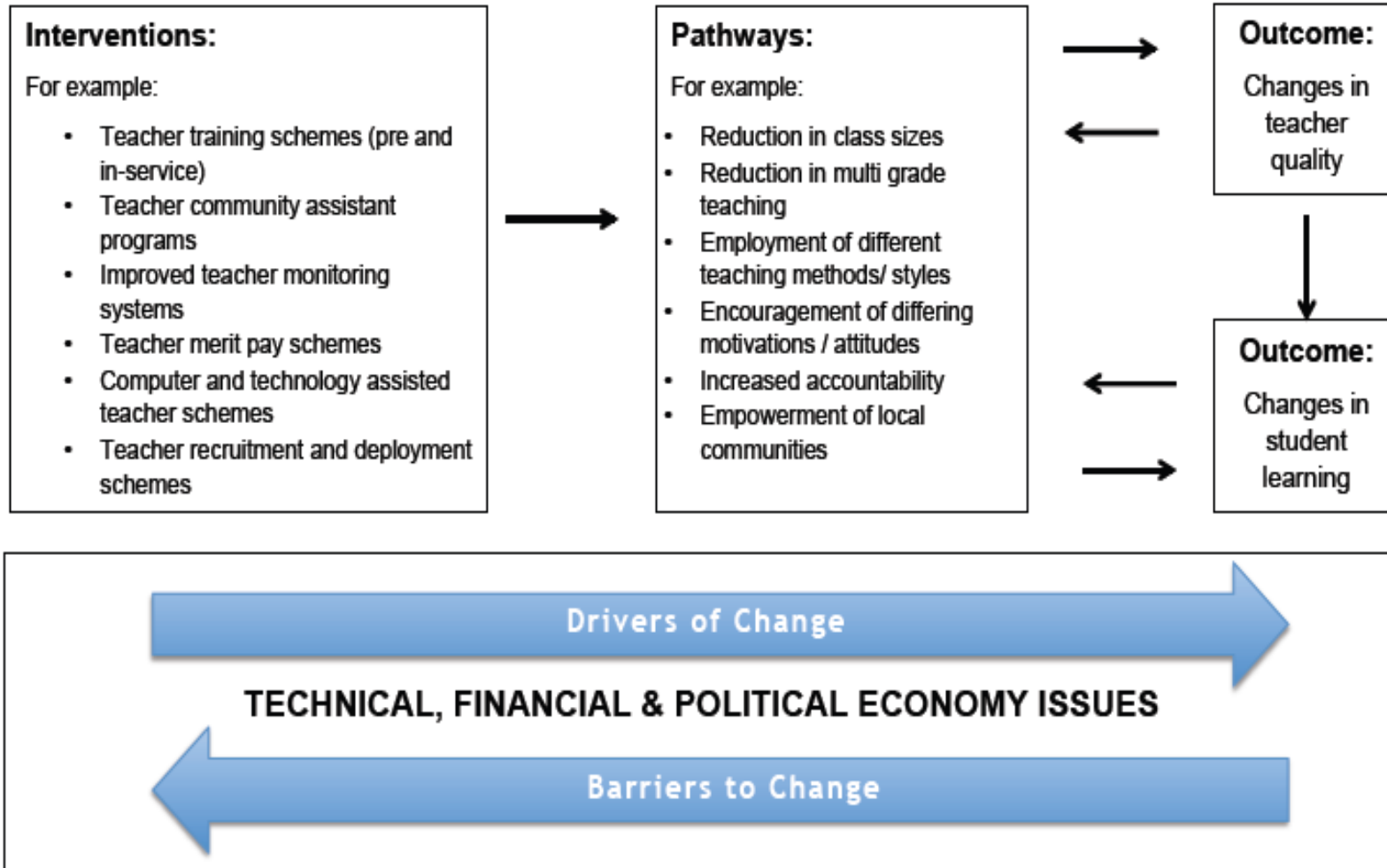
## THEORY OF CHANGE

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 programmes 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 seeks to 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. This is shown in a revised diagram (Figure 4.1).

The left-hand side of this diagram indicates the first step in this relationship, namely, the interventions. For example, these may include (but 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 influence student-learning outcomes, indirectly, through their impact on teacher quality, and/or directly, 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 a 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 will potentially negatively impact student outcomes. The ultimate goal of education policy is the improvement of student learning and any reform that improves teacher quality should also demonstrate a resultant improvement in student learning, in order to be deemed successful. However, sometimes, this impact may not yet have translated into observed improvements in student learning through research and, therefore, policies that show improved teacher quality without necessarily showing corresponding improvements in student learning may actually be showing flaws in the research evaluation, as opposed to the policy itself. Given that student learning is the most critical output, and the fact that the literature tends to focus on this (albeit often because it is more easily, and relatively more accurately, measurable), this review will also place more emphasis on student learning as the ultimate goal for policymakers, as opposed to teacher effectiveness. Improvements in teacher effectiveness can, therefore, be seen within the context of this review as not only an

outcome of interest, but also a channel for the main end-goal: improvements in student learning. 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. In particular, educational reforms are shaped by the interests and incentives faced by different stakeholders (in addition to teachers), as well as by the interests of different formal and informal institutions. There are several factors that may inhibit or promote educational reform (for example: political knowledge of the electorate, the extent to which elites dominate the political arena, the extent of centralization of governance, etc. Some of these factors/stakeholders/institutions lie within the remit of education systems, while others may be in the broader political space and yet still impact educational and teacher-related policies. The theory of change depicted below explicitly allows for the examination of these technical, financial and political-economy issues with a view to identifying situations where certain barriers have been overcome and have allowed certain education-system interventions to be achieved at scale. It should be noted that this initial theory of change was adapted and supplemented once the review of literature had been undertaken to cover comprehensively the key associations and identify-specific barriers to reform and drivers of change within different education systems studied. See diagram 4.1 and the discussion thereof in section 4.3.

**Figure 1.1: Theory of Change**



Interventions:

For example:

- Teacher training schemes (pre and in-service)
- Teacher community assistant programs
- Improved teacher monitoring systems
- Teacher merit pay schemes
- Computer and technology assisted teacher schemes
- Teacher recruitment and deployment schemes

Pathways:

For example:

- Reduction in class sizes
- Reduction in multi grade teaching
- Employment of different teaching methods/ styles
- Encouragement of differing motivations / attitudes
- Increased accountability
- Empowerment of local communities

Outcome:

Changes in teacher quality

Outcome:

Changes in student learning

Drivers of Change

TECHNICAL, FINANCIAL & POLITICAL ECONOMY ISSUES

Barriers to Change

## 1.3 POLICY, PRACTICE AND RESEARCH BACKGROUND

The question posed in this review is unique in that it asks a critical, 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 may be 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, 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 policymakers and practitioners. This has been 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 address this question directly. 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.

## 1.4 AUTHORS, FUNDERS, AND OTHER USERS OF THE REVIEW

In funding this review, DFID is basing its programming and policymaking decisions on 'what we know', rather than on conjecture and, in doing so, hopes to achieve the maximum value for money for each and every programme. The increased emphasis by donors and policymakers on commissioning research that is based on solid evidence and strong research designs means that studies such as this one form the backbone of what we know. Reviews such as this one allow funders to identify the evidence base and to test assumptions and hypotheses and, therefore, make informed policy decisions.

It is expected that this review will be of widespread interest, given the uncontested acceptance of the importance of teachers in contributing to the learning outcomes of children across the world. Policymakers across developing countries have and continue to focus on fiscally manageable and effective means of improving the effectiveness of the teaching workforce, and this review provides an indication of the evidence base that currently exists regarding reforms that have been implemented at scale and their possible impact on both teacher quality and student-learning outcomes. The policy-implications section that is included in the document will add rigor to policymakers' decisions in respect of which interventions have been shown to be effective in previous research. Of particular importance will be the discussion of financial, technical and political barriers/drivers of change that have allowed large-scale reform efforts to be implemented.

The authors aim to publish the report in a peer-reviewed journal. The final report will also be presented to policymakers at DFID and externally, where possible, to audiences at, for example, the UKFIET and CIES conferences.

## 1.5 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 experimental literature that is able to identify causation between teacher-effectiveness interventions and two 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. Several of these are weak proxies of teacher quality. Nevertheless, in the absence of any more effective measures, they are used often in the literature investigating teacher effectiveness (Goldhaber et al. 1999, Burgess et al. 2009, Kingdon and Teal 2010, Aslam and Kingdon 2011). The second outcome is student achievement. This will focus solely on learning outcomes, as measured by test scores.

***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 draws on studies that focus on the relationship between teacher effectiveness and student outcomes. This includes numerical data (for example, correlated studies using statistical analysis) and qualitative studies (for example, drawing on participants' perception that outcomes have improved as a result of participating in an intervention), but which do not establish causation. The outcomes investigated in this question remain the same as those on question 1.

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

From the studies that have been identified as answering questions 1 and 2, we 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 has been designed to ensure that we have included a broad range of interventions aimed at improving teacher effectiveness.

A distinction has been made between RQ1 and RQ2 in order to enable a wider study design that allows us to capture both studies that are able to identify the *impact* (a causal link), as well as the *association* between the reform/intervention and the outcomes of interest.

## HOW HAS THE ISSUE OF SCALE BEEN ADDRESSED?

The question of whether an intervention is at scale or not is very context- and programme-specific. Therefore, using a stringent and quantified definition of 'scale' may have limited our research, because a certain number of schools/teachers/pupils targeted in one country may constitute scale, but, in another context, may be insufficient to be considered to be at scale. A wide range of aspects were, therefore, considered in examining this issue. These included 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.

The issue of scale has been analysed in the final stages of the review process (see the Methods section, below). This has meant that, while the in-depth review focuses solely on scale interventions, by retaining non-scale literature, we are able to provide some evidence of medium-high-quality literature that focuses on the impact/association of non-scale teacher-effectiveness reforms on teacher quality and learning outcomes. While these do not form part of the main evidence presented, a short narrative has been provided to summarise the findings of this literature base, as the authors are of the view that this evidence contains useful information for both the funding and policy communities.

## 2      METHODS USED IN THE REVIEW

### OUTLINE OF CHAPTER

This section discusses the methodology used within the review. All necessary details are explicitly identified in order to allow replication of methods.

### 2.1 IDENTIFYING AND DESCRIBING STUDIES

#### DEFINING RELEVANT STUDIES: INCLUSION AND EXCLUSION CRITERIA

PICOS inclusion and exclusion criteria (Higgins 2011) have been used to determine study eligibility.

**POPULATION/COUNTRY FOCUS**: We have focused on all countries currently listed as low- or lower-middle-income by the World Bank (http://data.worldbank.org/about/country-classifications).

However, this runs the risk that some countries that moved from lower-middle-income to higher-middle-income between 1990 and 2014 have been missed. However, as it is inevitable that countries change over time, the risk of missing one or two potentially relevant countries will be inherent in any country-based exclusion criteria. We have not confined 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 has excluded: high-income countries, transition economies and upper-middle-income countries.

**SETTING**: Only studies set in primary and secondary government/state-run schools were eligible.

**INTERVENTION**: The focus has been on teacher-effectiveness interventions (such as contract-teacher schemes, pre-service training, in-service training, merit pay).

**COMPARISON**: To answer RQ1 on the impact of teacher-effectiveness interventions, studies were required to have a comparison group. For example, treated teachers vs. non-treated teachers; or students taught by treated teachers vs. those taught by non-treated teachers. To answer RQ2, studies needed 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. Where comparison groups do not exist, rigorous methods to control for bias, etc., must have been implemented.

**OUTCOMES**: Studies must report learning outcomes; for example, academic achievement tests and/or teacher quality (time on task, teacher motivation, competence, absence, skills, effort, qualifications, credentials, teacher test scores etc. .

Criteria in addition to PICOS included:

**DATE**: 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 material.

**SCALE**: Provide data on the impact of the intervention 'at scale' (this criterion has only been implemented in the final stages of the review process, just prior to the in-depth-review stage).

## IDENTIFICATION OF POTENTIAL STUDIES: SEARCH STRATEGY

A 'search strategy' was devised and search terms arrived at in collaboration with all team members. An iterative procedure was 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 were agreed among team members and all efforts were made to ensure that the final terms allowed 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 were also searched for. The key search terms were determined by the review question and the inclusion and inclusion criteria, as outlined above. The search strategy involved 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 were only run with this concept. Further searches were then run, as below, with additional concepts.

**Concept 2**: Synonyms that 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 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 above for definition and link to countries).

The search strategy was adjusted according to individual databases and web-based interface capabilities, as required:

- The searches in each database initially began with (CONCEPT 1: interventions/reforms) AND (CONCEPT 5: LMIC). This is to ensure that the search is as broad as possible,.

- Next, three separate searches were run and the outputs from each were saved for screening. The searches were 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)

• Where date settings were available, the searches in the database were restricted to literature between 1990 and 2014. When date settings were not available in the databases, we screened out literature pre-1990 during the title and abstract-screening stage.

• Wild cards, proximity searches and thesaurus terms were used as appropriate. Search notes and search strings for each database were maintained and logged within supporting documentation. Examples of search terms and strings applied to an example database can be found in Appendix 2.2.

## SCREENING STUDIES: APPLYING INCLUSION AND EXCLUSION CRITERIA

### TITLE AND ABSTRACT SCREENING:

Once the studies were identified, they were uploaded to EPPI-reviewer and screened for their relevance to the systematic review. The inclusion criteria were formally stated using PICOS method, to which we added a time parameter. These inclusion and exclusion criteria were applied to title and abstract. Excluded studies were those that failed to satisfy at least one inclusion criterion, or met at least one of the exclusion criteria. Studies that failed to meet the inclusion criteria were coded as such, so that it could be reported how many inclusion criteria each study failed to meet and what these criteria were. If a study scored 'Yes' for the relevance criteria, it was brought forward to the full-text-screening stage.

### FULL-TEXT SCREENING:

**Inclusion/Exclusion criteria:** At the full-text-screening stage, full reports were obtained for those studies that appeared to meet the criteria or where there was insufficient information to decide. This involved reviewing the full text and re-applying the PICOS framework. The included studies were coded by various indicators (region, setting, sample size, etc.), and were taken forward to the final stage, the in-depth review. At this point in the review, we distinguished between interventions of the following types: i) Interventions at scale and ii) interventions not at scale. Only studies that fell into category (1) were carried forward to the synthesis stage, and the studies categorised as not at scale are 'described only'.

## CHARACTERISING INCLUDED STUDIES

The review team organised studies by type. Each study was coded based on the intervention/s being studied, the outcomes being analysed, country and findings. At the final

stage, the studies were then coded based on whether the intervention was at scale or not. All the key-worded studies were added to the larger EPPI-Centre database, for others to access via the website. Firstly, this was done to allow recurring themes to be identified. Additionally, by having the literature base coded in such a functional and constructive manner, it allowed the research term to capitalise on the extensive research base that it has 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 accessed and extracted from the current review.

## IDENTIFYING AND DESCRIBING STUDIES: QUALITY-ASSURANCE PROCESS

The systematic review followed the standard EPPI-Centre procedures for maintaining quality. At the scoping review stage, to ensure consistency in application of the selection criteria, reviewers undertook double screening on a sample of papers to pilot the inclusion/exclusion criteria. The remainder of the screening was carried out by individual reviewers. Where there was uncertainty, reports were marked for discussion and, at the end of the screening process, these reports were considered by two or three reviewers, as required. As a final check, all reports selected for inclusion were checked by the second reviewer, in order to confirm their relevance. At the synthesis stage, data extraction and quality-assessment processes were undertaken by two researchers working independently, in order to achieve a high level of consistency.

## 2.2 IN-DEPTH REVIEW

## MOVING FROM BROAD CHARACTERISATION (MAPPING) TO IN-DEPTH REVIEW

The studies identified as meeting the inclusion criteria were coded in-depth using a detailed DFID (2014) data-extraction tool to determine their eligibility for inclusion in the in-depth review (see appendix 2.4). This tool covered the following dimensions:

- the methodological quality of the study;
- the relevance and appropriateness of the research design;
- the relevance of the focus of the study.

Reviewers' judgments on each dimension informed the overall WoE (that is, trustworthiness) of each study. (See table 3 in appendix 2.4).

Determining the quality of each study was undertaken using a step-by-step approach: The validity, reliability and applicability of quantitative studies were determined by applying a hierarchy-of-evidence model. Studies were judged as high-quality (for example, RCTs) to lower quality (for example, studies employing simple descriptive statistics that do not allow causal interpretations, such as comparison of means). The hierarchy used for evaluating quantitative studies was as follows:

- Systematic review of RCTs.
- An Individual RCT that evaluates effect.
- Systematic review of cohort studies.
- Individual cohort study: (a) Using techniques to control for endogeneity. (b) Comparison of means.

The validity of studies with narrative/qualitative data was decided by choosing those that give relevance to the wider context and judging the extent to which studies employed a methodology that minimises the risk of bias.

Risk of bias was based on Higgins' (2011) criteria, for example: selection bias, selective reporting bias, placement bias, consideration of intervention integrity, consideration of differences within groups, explaining variations in outcomes, among other factors.

All studies were screened on their reliability (for example, the extent to which their findings were reproducible), and whether their findings were applicable.

Quality assurance and consistency were ensured by reviewers assessing studies independently, using an agreed-upon approach.

All studies that did not meet the above criteria were excluded from the in-depth review.

In addition to the above, the following issues were also considered while completing the data-extraction forms (see Appendix) for each of the study included:

**Completeness of reporting:**

This entailed assessing transparency, reporting bias and publication bias. We expected a good-quality study to have a description of the intervention and the participants (children), a 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 was considered of poor quality if it failed to meet one or more of these requirements.

**Feasibility of assumptions:**

If the reviewers remained unconvinced about the assumptions made within the study on which the conclusions were based, the study was classified as of low or moderate quality. While studies that did not specifically articulate their assumptions were not automatically excluded, the assessment of quality was affected by whether or not a study articulated its assumptions clearly or not.

**Appropriateness of methodology:**

Methodology was analysed to ensure trustworthiness, reliability and validity. Assessment of the appropriateness of the methodology depended on whether a study was quantitative or qualitative in nature. These were assessed according to the approaches discussed below.

**Consideration of confounding factors:**

These included (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 sampling stage or at analysis stage. If studies took no consideration of confounding factors at either stage, they were considered of poor quality and were excluded from the in-depth review. Studies that controlled for confounding factors at any one stage were considered of moderate quality and were included in the in-depth review. Studies controlling for these factors at both stages were considered to be of high quality and were used for in-depth review.

**Comprehensive reporting of findings:**

Were the studies' findings apparent and comprehensively reported? For example, if a study initially aimed to measure certain outcomes, but did not report on all of the outcomes, it was judged of poor quality and subsequently excluded.

## ASSESSING QUALITY OF STUDIES AND WEIGHT OF EVIDENCE

### EVALUATING THE QUALITY OF METHODOLOGY OF QUANTITATIVE STUDIES

A second round of quality appraisal was taken to evaluate the methodology of quantitative studies. The following recognised questions were probed from each individual study in order to assess not only the quality of the underlying studies, but also in establishing a WoE provided by these studies for each research question posed in the review.

(i) How was the intervention assigned? That is, was the assignment random or non-random? If random, the study was judged to be of high quality; if non-random, it was judged as being of moderate or poor quality, depending on how the intervention was further assigned.

(a) If randomised, is the counter-factual clearly stated? For example:

(1) Teacher training/education versus untrained teachers?

(2) Contract teacher versus regular teacher?

(3) Classroom assistant versus no assistant?

The following options were available: Yes/No/Partly. If for example, the answer to this was a clearly stated 'yes', the study was considered of high quality; if the response was 'partly', the study was coded as moderate; and when it was not clearly stated, the study was considered of poor quality.

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

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

(ii) What question was being asked in the study? Did it evaluate the 'as is' effect of the intervention, or did it evaluate a conditional effect? A study that controls for the 'as is' effect was considered of moderate quality, while one controlling for the conditional effects was considered of high quality. Both were included in the in-depth review.

(iii) Was the intervention effect homogeneous across different student types? Studies that consider the intervention affect across different student types were considered high-quality; those that did not were considered of moderate quality, and both were included in the final review.

(iv) Was the cohort representative of the population? If not, did the sample have any characteristics that may have affected the external validity of results? If so, the study was considered of poor quality and was, therefore, excluded from the in-depth review.

## EVALUATING THE QUALITY OF METHODOLOGY OF QUALITATIVE STUDIES

1) Was the epistemological approach clearly stated? Yes/No/Partly — studies where it was clearly stated were considered to be of high quality; studies where it was partly stated were rated as of moderate quality; and studies where it was not stated at all were ranked as being of poor quality.

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

3) Was data collection appropriate/repeatable and trustworthy? Yes/No/Partly — studies where it was appropriate/trustworthy were considered of high quality, partly stated of moderate quality and, where it was not stated at all, 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 of moderate quality, and, where it was not stated at all, of poor quality.

Based on the findings of the above, studies were judged to be either of high, moderate or low quality. In order to ensure rigor, judgments relating to the above were made independently and the reviewers discussed the studies where any differences in opinion were observed.

In addition to questions on methodology, DFID's How To Note (Appendix 2.4) was used to evaluate each individual study for cogency, reliability, applicability, etc. In arriving at the final quality assessment of a given study, the ratings given in each of these criteria were aggregated and a study was deemed to be of high quality if it met all of the criteria listed. A

study was judged to be moderate quality if it met the majority (but not all) and low quality if it met just a few of the criteria.

## SYNTHESIS OF EVIDENCE

### DATA EXTRACTION

A preliminary instrument was initially designed to guide the researchers in retrieving the information from each study. These forms were filled in for all studies that made it through the screening stage (see Appendix 2.1). Although preliminary, the instrument indicated the types of key questions we aimed to answer for each study. Many of the open-ended questions were then collated into smaller categories and appropriate tables generated. The forms asked 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 led us to a final set of quality studies that provide robust evidence that either supports, counters or is neutral in respect of the different proposed relationships between interventions and outcomes. A diagram to identify the flow of studies has been included in the review, and it maps out the process and indicates how the final set of studies was arrived at to ensure transparency (Figure 3.1).

### METHODS FOR SYNTHESISING

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

**Review Question 1: What is the evidence on the impact 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 included studies that measure the effects of interventions using experimental and quasi-experimental study designs. Specifically, we included (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 groups, provided they included pre- and post-test measures of the outcome variables of interest to ensure equity between groups on the baseline measure, as well as using appropriate methods to control for selection bias and confounding, such as: statistical matching (for example, propensity score matching, or covariate matching), regression adjustment (for example, difference-in-differences, and single-difference regression analysis, instrumental variables, and Heckman selection models).

**Review Question 2: 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 included studies without random allocation in treatment and comparison groups 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. Additionally, it should factor in the potential non-random matching of contract teachers with 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 state 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, which are unable to take this into account as effectively, have been excluded from this review. Only studies that attempt to 'control' for the wider social and economic context, and provide accurately more accurate level of generalisation, have been retained in answering this second question (see Kingdon et al. 2013). Quantitative studies that do not effectively control for confounding factors or self-selection have also been 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 were considered during the review process for inclusion in the final set, provided they met the stringent quality-assessment criteria postulated by the reviewers.

**Review Question 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?**

This question has been addressed using evidence from the literature base identified in answering questions 1 and 2.

Overall, the results of this in-depth review have been collated to provide a WoE from the overall evidence base, directly informing a synthesis of evidence.

Once all the appropriate literature had been identified and assessed, the research was collated. This synthesis has been presented in the form of structured thematic narrative and summary tables (see Chapter 4).

The protocol indicated that, depending on the nature of the studies identified, and the data included in them, we would consider conducting a meta-analysis of the quantitative studies; however, this was not possible. The final set of studies that were found to be investigating teacher-effectiveness reforms, at scale, did not allow us to undertake calculations of effect size or meta-analysis. The synthesis is, therefore, presented in the form of a rich narrative that describes the key pieces of evidence that help answer the questions posed in this review.

## IDENTIFYING GAPS IN THE EVIDENCE BASE AND GENERATING POLICY-RELEVANT IMPLICATIONS

In addition to a synthesis of the established body of evidence, a main objective of this review has also been to identify key gaps in the literature, and derive policy and practice implications. Studies that do not necessarily meet quality-assurance procedures are still discussed in the review, albeit clearly identified as not forming the main body of evidence in the in-depth review, as they provide useful insights and context.

## 3        IDENTIFYING AND DESCRIBING STUDIES: RESULTS

This section describes the studies that are included in the in-depth review and the findings thereof.

## 3.1 STUDIES INCLUDED FROM SEARCHING AND SCREENING

Figure 3.1 illustrates the filtering process from initial screening to in-depth review. The searches were conducted between September and November 2014. A total of 4,010 citations were obtained, on which title and abstract screening were conducted. As a result of this, 325 citations were brought forward for full-text screening. The reviewers went to great lengths to source each of these documents, and only four were not obtainable. Following on from the full document screening, 36 studies were taken forward to quality-assurance stage. Nine were excluded based on the studies' being assessed as low-quality. Quality assurance was conducted independently by two reviewers on each of the 36 studies, following the stringent criteria as highlighted in the previous section. Once reviewers had independently reviewed the studies for quality, any differences in opinion were discussed and studies classified accordingly. The resultant 27 studies were scrutinised in detail and 15 studies were identified as being at-scale and 12 as non-scale. The findings from these 15 studies are discussed below.

**Figure 0.1: Filtering of papers from searching to map to synthesis**

Two-stage screening
Papers identified where there is not immediate screening; for example, electronic searching

4,010 citations identified

Title and abstract screening

Citations excluded based on:
Language – 3
Time - 2
Country - 929
Document Type - 42
Intervention – 1,988
Participants – 282
Teacher-Quality Outcomes - 323
Student Outcomes - 49
Other - 11

TOTAL – 3,629

381 citations

325 citations identified in total

56 duplicates excluded

Acquisition of reports

4 reports not obtained

321 reports obtained

Full-document screening

Reports excluded
Country – 3
Duplicate – 3
Document Type - 31
Intervention - 88
Participants - 4
Review Question - 165
Outcome - 15
Study Design - 115
TOTAL - 424 (studies may be excluded on more than one criteria)

36 studies included

Systematic Map of 36 studies

Studies excluded based on quality assurance - 9

27 studies included

Studies excluded as 'non-scale' - 12

In-depth review
15 studies

## 3.2 CHARACTERISTICS OF THE INCLUDED STUDIES (SYSTEMATIC MAP OF THE EVIDENCE, INCLUDING NON-SCALE STUDIES)

Table 3.1, below, provides a summary of the characteristics of all 27 studies that have been included, based on the quality-assurance process. This includes all of the studies classified as at-scale, as well as non-scale. While these studies cover a wide range of interventions, a large majority focus on those interventions relating to contract teachers (n=10) and those relating to monetary incentives for teachers (n=5). There is also clearly more evidence of a quantitative nature (n= 26) than that of a qualitative nature (n=1). The geographical spread of the studies covers Africa and Asia, as well as parts of Latin America.

### Table 0.1 Summary map of the evidence (scale + non-scale)

| No. | Year | Authors | Country | Methodology | Outcomes | Scale |
|---|---|---|---|---|---|---|
| **Contract Teachers (n=10)** | | | | | | |
| 1 | 2003 | De Laat and Vegas | Togo | Quantitative: Fixed Effects | Student Achievement | Yes |
| 2 | 2005 | Bourdon et al. | Niger | Quantitative: Propensity Score Matching | Student Achievement | Yes |
| **3** | **2007** | **Bourdon et al.** | **Niger, Mali and Togo** | **Quantitative: Quantile Regression** | **Student Achievement** | **Yes** |
| 4 | 2009 | Duflo et al. | Kenya | Quantitative: RCT | Student Achievement and Teacher Quality | No |
| 5 | 2010 | Atherton and Kingdon | India | Quantitative: School Fixed Effects | Student Achievement | Yes |

| 6 | 2010 | Goyal and Pandey | India | Quantitative: Fixed Effects | Student Achievement and Teacher Quality | Yes |
|---|---|---|---|---|---|---|
| 7 | 2010 | Habib | Pakistan | Qualitative: In-depth Interviews | Teacher Quality | Yes |
| 8 | 2012 | Bold et al. | Kenya | RCT | Student Achievement | Yes |
| 9 | 2013 | Muralidharan and Sundararaman | India | RCT | Student Achievement and Teacher Quality | Yes |
| 10 | 2014 | Duflo et al. | Kenya | Quantitative: RCT | Student Achievement and Teacher Quality | No |
| **Monetary Incentives (n=5)** | | | | | | |
| 11 | 2005 | McEwan and Santibanez | Mexico | Quantitative (with controls) | Student Achievement | Yes |
| 12 | 2005 | Mizala and Romaguera | Chile | Quantitative: Fixed Effects | Student Achievement | Yes |
| 13 | 2010 | Barrera-Osorio and Raju | Pakistan | Quantitative: RDD | Student Achievement | Yes |

| 14 | 2012 | Rau and Contreras | Chile | Quantitative: Use three different estimation methods (matched difference in difference, double robust and panel —FE) | Student Achievement | Yes |
|----|------|-------------------|-------|------------------------------------------------------------------------------------------------------------------------|---------------------|-----|
| 15 | 2014 | Pugatch and Schroeder | Gambia | Quantitative: Difference in difference and Regression Discontinuity Design | Teacher Quality | Yes |
| **Other teacher interventions (n=7)** | | | | | | |
| 16 | 1999 | Tan et al. | Philippines | Quantitative: RCT | Student Achievement | No |
| 17 | 2003 | Banerjee et al. | India | Quantitative: RCT | Student Achievement | No |
| 18 | 2007 | Banerjee et al. | India | Quantitative: RCT | Student Achievement | No |
| 19 | 2009 | Piper | Ethiopia | Mixed Methods | Student Achievement and Teacher Quality | Yes |
| 20 | 2010 | Lassibille | Madagascar | Quantitative: RCT | Student Achievement and Teacher | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | Quality | |
| 21 | 2010 | Muralidharan et al. | India | Quantitative: RCT | Student Achievement and Teacher Quality | No |
| 22 | 2014 | Chang et al. | Indonesia | Quantitative (study 1) and Mixed Methods (study 2) | Student Achievement and Teacher Quality | Yes |
| **Teacher performance pay (n=5)** | | | | | | |
| 23 | 2010 | Glewwe et al. | Kenya | Quantitative: RCT | Student Achievement and Teacher Quality | No |
| 24 | 2011 | Muralidharan | India | Quantitative: RCT | Student Achievement and Teacher Quality | No |
| 25 | 2011 | Muralidharan and Sundararaman | India | Quantitative: RCT | Student Achievement and Teacher Quality | No |
| 26 | 2012 | Duflo et al. | India | Quantitative: RCT | Student Achievement and Teacher Quality | No |
| 27 | 2012 | Behrman et al. | Mexico | Quantitative: RCT | Student Achievement | No |

The results of the in-depth review are discussed in more detail in the following chapter. These results pertain only to studies where the intervention has been deemed to be at-scale. The authors of the review considered the issue of scale as relating to the intervention, as opposed to the evaluation. If an evaluation was of a smaller scale, but is assessing a large-scale reform/intervention, it has been included in the in-depth review.

# 4    IN-DEPTH REVIEW: RESULTS

## 4.1 SELECTING STUDIES FOR THE IN-DEPTH REVIEW

Chapter 3 described the findings of the first stage of the review process and provided an overview of the 27 studies that met the review criteria for inclusion in the map. This chapter describes the second stage of the review process and presents the quality and findings of the 15 studies also meeting the review criteria, but which also evaluate interventions at scale to answer the questions presented in chapters 1 and 2.

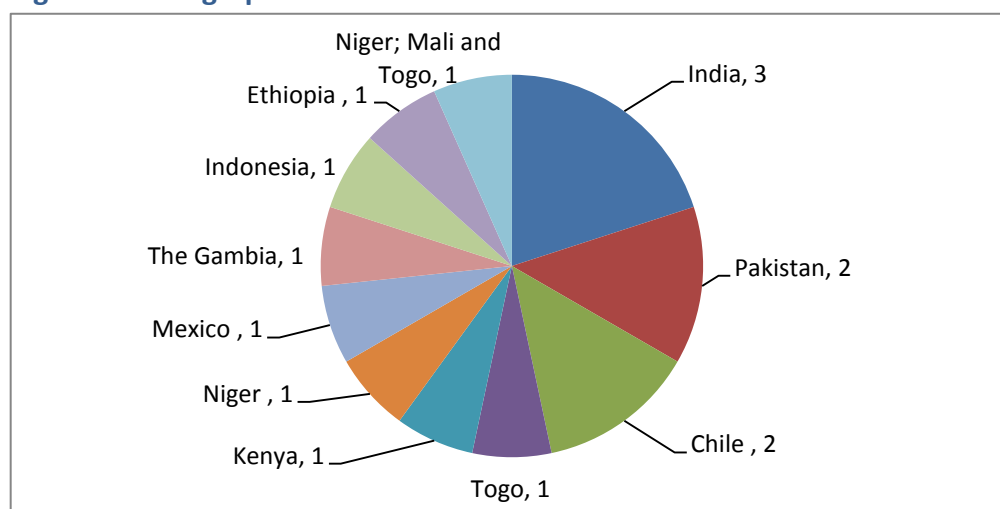## 4.2 FURTHER DETAILS OF STUDIES INCLUDED IN THE IN-DEPTH REVIEW

### OVERVIEW OF STUDIES INCLUDED IN THE IN-DEPTH REVIEW

In assessing the overall strength of evidence, we have used the **quality** (based on assessment of cogency, reliability, methodology, as per DFID's How To Note Template, Appendix 2.6, and discussed above) of individual studies constituting the body of evidence, the **size** of the body of evidence (whether it is large: 30 studies or more; medium: 10-30 studies; or small: fewer than 10 studies), the **context** they cover (global or context-specific) and the **consistency of findings** (a range of studies pointing to identical or similar conclusions versus different studies pointing to different findings), to conclude whether our review shows strong, modest or insufficient evidence of the relationship being studied.

### GEOGRAPHICAL LOCATION/CONTEXT

The majority of studies were conducted in Africa (six: one study on three countries in West Africa and Sub-Saharan Africa; one study on three countries in Sub-Saharan-Africa; and two studies on countries in West Africa), with fewer countries represented from South America (one) and South Asia (two).

### Figure 4.1 Geographical location

## TYPES OF INTERVENTION

The studies included in the in-depth review also cover a broad range of teacher-effectiveness interventions, with the majority focusing on contract teachers or monetary incentives.

### Table 4.1

| Types of intervention: | N=15 |
|---|---|
| Contract-teacher interventions | 9 |
| Teacher-remuneration interventions | 1 |
| Teacher monetary incentives | 4 |
| Teacher-management and deployment | 1 |

## STUDY DESIGN APPROACH

Studies that answer RQ1 on the impact of interventions were solely experimental studies. Studies answering RQ2 on the relationship between educational reforms/interventions use varying methodologies. There were also two studies that use qualitative (or mixed-methods) approaches, appropriate for answering RQ2.

### Table 4.2

| Study design approach | N=15 |
|---|---|
| Quantitative methods | 13 |
| Qualitative approaches | 1 |
| Mixed-methods study | 1 |

## OUTCOMES

The concentration of the evidence base is on measuring student outcomes (n=13), with much of this focused on learning outcomes. Teacher quality was also measured, but only in half the studies (n=8)

### Table 4.3

| Outcomes measure | N=15 |
|---|---|
| Student achievement only | 8 |
| Teacher quality only | 2 |
| Teacher quality and student outcomes | 5 |

DFID's Note on 'Assessing the Strength of Evidence in the Education Sector, 2015' was adapted and used to guide the overall strength of the evidence. This was then used to assess whether the evidence was 'strong', 'modest' or 'insufficient' in respect of the intervention being studied, as follows (Table 4.4).

**Table 4.4: Assessing the strength of evidence**

| | |
|---|---|
| **Strong** | High-quality body of evidence, large or medium in size, generally consistent, and covers several contexts. |
| **Modest** | High- or moderate-quality studies, medium-size evidence body, generally consistent, not covering a wide range of contexts. |
| **Insufficient** | High- or moderate-quality studies, small or medium-sized body, inconsistent, and covers very limited contexts. |

Table 4.5 provides an overview of the results on the set of 15 studies that were reviewed in-depth. In terms of **quality**, of the 15 studies, 11 studies have been ranked as moderate and four studies are ranked high-quality, based on the quality assessment (assessment of quality was based on DFID's How to Note guidance; see Appendix 2.4). In terms of the overall strength of the body of evidence, there is modest evidence of the relationship between contract teachers and teacher quality and student achievement. Similarly, there is modest evidence on the relationship between monetary incentives and student achievement.

In relation to all other interventions and their relationship with the outcomes of interest (for example, monetary incentives and teacher certification and their relationship with teacher quality and learning outcomes), this review finds insufficient evidence. This dearth of evidence would point towards the need for more robust investigations in these particular categories.

Individual summaries of the studies included for the in-depth review are presented in Table 4.6.

## Table 4.5: Summary of evidence reviewed

| Intervention | Outcome Category | Number of Studies (size) | Contexts in which intervention has positive effect | Contexts in which intervention has negative effect | Contexts in which intervention has no effect | Strength of evidence: strong, modest, insufficient) |
|---|---|---|---|---|---|---|
| **Contract Teachers** | Teacher Quality | 4 | 3 | | 1 | Modest |
| **Contract Teachers** | Student Achievement | 7 | 5 | 2 | 2 | Modest |
| **Monetary Incentives** | Teacher Quality | 1 | 1 | | | Insufficient |
| **Monetary Incentives** | Student Achievement | 4 | 2 | | 2 | Modest |
| **Teacher Certification** | Teacher Quality | 1 | | | 1 | Insufficient |
| **Teacher Certification** | Student Achievement | 1 | | | 1 | Insufficient |
| **Teacher Training** | Teacher Quality | 1 | 1 | | | Insufficient |
| **Teacher Training** | Student Achievement | 1 | 1 | | | Insufficient |

*Note: The 'studies' in the table above may cover more than one context and, therefore, in showing the positive/negative/no effect, totals may not necessarily equal the total number of studies.*

*It is also important to note that the use of vote counting means that the synthesis is limited to comparing the number of positive and negative studies (the direction of effect), but does not take into account the size of the sample or the magnitude of effect.*

**Table 4.6: Summary of Included studies**

| Authors | Country | Methodology | Outcomes | Results | Quality of individual study |
|---|---|---|---|---|---|
| **Contract teachers (n=8)** | | | | | |
| **Atherton and Kingdon (2010)** | India | Quantitative: School Fixed Effects | Student Achievement and Teacher Quality | UP contract teachers' positive relationship with student learning. Bihar contract teachers no less effective than regular teachers. UP contract teachers display higher effort via lower absence rates. | Moderate |
| **De Laat and Vegas (2003)** | Togo | Quantitative: Fixed Effects | Student Achievement | Contract teachers underperform regular teachers; authors suggest this may be due to a decline in teacher quality. | Moderate |
| **Muralidharan and Sundararaman (2013)** | India | RCT | Student Achievement and Teacher Quality | Extra contract teachers improves learning outcomes. Contract teachers less likely to be absent and more likely to be observed teaching. Contract teachers no less effective than regular teachers. | High |

| | | | | | |
|---|---|---|---|---|---|
| **Goyal and Pandey (2010)** | India | Quantitative: Fixed Effects | Student Achievement and Teacher Quality | Teacher attendance higher for contract teachers. Also, engagement in teaching higher. Student test scores also shown to have positive relationship with contract-teacher status. Tenure period also matters, with those teachers who were in later tenure periods exerting less effort. | Moderate |
| **Bold et al. (2012)** | Kenya | RCT | Student Achievement | Significant positive effect (0.19 SD) in schools with NGO implementation and zero effect in MOE schools. Assess prospects for scaling up contract-teacher intervention. | High |
| **Bourdon et al. (2005)** | Niger | Quantitative: Propensity Score Matching | Student Achievement | Negative relationship of contract-teacher status and achievement, but authors suggest this may be due to lower experience. | Moderate |
| **Bourdon et al. (2007)** | Niger, Mali and Togo | Quantitative: Quantile Regression | Student Achievement | Contract teachers do relatively better with low-ability students. Positive relationship in Mali, mixed in Togo and negative in Niger. | Moderate |

| | | | | | |
|---|---|---|---|---|---|
| **Habib (2010)** | Pakistan | Qualitative: In-depth interviews | Teacher Quality | Contract-teacher policy was found, overall, to have relatively little impact on teacher absenteeism. The study found that contract teachers were absent with only moderately less frequency than their regular counterparts. | Moderate |
| **Monetary Incentives (n=5)** | | | | | |
| **Mizala and Romaguera (2005)** | Chile | Quantitative: Fixed Effects | Student Achievement | Preliminary estimates show positive effects of SNED on educational outcomes for certain schools. Also show a change in teacher attitudes. Authors state that the cumulative effect of different SNED applications is important. | Moderate |
| **McEwan and Santibanez (2005)** | Mexico | Quantitative (with controls) | Student Achievement | Stronger incentives do not necessarily relate to improvements in student achievement. | Moderate |

| | | | | | |
|---|---|---|---|---|---|
| **Pugatch and Schroeder (2014)** | The Gambia | Quantitative: Diff in diff and RDD | Teacher Quality | Diff in diff analysis shows an increase in percentage of qualified teachers. RDD shows only improvement in pupil: qualified-teacher ratio. Examine whether this is a substitution or scale effect. Seems to be improving the spread of current qualified teachers as we increase the number of newly qualified teachers, the latter being preferable. | Moderate |
| **Rau and Contreras (2012)** | Chile | Quantitative: Use three different estimation methods (matched diff in diff, double robust and panel — FE) | Student Achievement | Find a positive relationship between the SNED programme and student test scores. Provides evidence for differentiated pay structure as a means of improving test scores. However, they show that these types of tournaments are only effective for a certain subset of schools and, therefore, more research is required into different types of designs and incentive mechanisms. | Moderate |
| **Barrera-Osorio and Raju (2010)** | Pakistan | Quantitative: RDD | Student Achievement | Future teacher-bonus awards are not shown to induce learning gains for marginal bonus non-qualifiers. Apart from the pressure from below to maintain a minimum level of learning for the programme, participating schools do not face any effective incentive to continuously raise learning. | Moderate |

**Teacher training (n=1)**

| | | | | | |
|---|---|---|---|---|---|
| **Piper (2009)** | Ethiopia | Mixed Methods | Student Achievement and Teacher Quality | Student achievement is shown to have a positive relationship with the program. While the program did not appear to change pedagogical methods used by trained teachers, it did appear to result in improved decisions in relation to pedagogical choices made by trained teachers. | High |

**Teacher certification (n=1)**

| | | | | | |
|---|---|---|---|---|---|
| **Chang et al. (2014)** | Indonesia | Quantitative (study 1) & Mixed Methods (study 2) | Student Achievement and Teacher Quality | RCT evaluates effects of certification on teacher productivity and finds that professional certification improves teachers' well-being, but does not necessarily make them 'better' in terms of student outcomes. Study 2 — looks inside 'black box' of teaching practice and finds that, while there was no difference between certified and uncertified teachers' subject matter and pedagogy, teacher knowledge stood out as having a strong association with student learning. Findings confirm that certification does not have an impact on teacher practices and behaviour. | High |

## 4.3 SYNTHESIS OF EVIDENCE

This section synthesises the evidence from the 15 studies included in the in-depth review. The synthesis is organised by the type of intervention examined and the type of outcome measured. We have synthesised evidence from studies answering RQ1, followed by RQ2, before providing a narrative synthesis to answer RQ3. It should be noted that the final set of 15 studies that investigate teacher-effectiveness reforms at scale do not allow us to undertake calculation of effect size or meta-analysis thereof. Therefore, the synthesis will be presented in the form of a rich narrative that describes key pieces of evidence that help answer the questions posed in the review.

Figure 4.1 summarises the main interventions and the key channels identified through a review of the evidence in this section. This figure has evolved from the theory of change (Figure 1.1), where we discussed the possible relationships between the various possible programmes/interventions/reforms to improve teacher effectiveness and the channels through which they could potentially impact learning outcomes. The diagram reinforces the relationships as illustrated in the initial theory of change; it also provides specific examples that have emerged from the literature through this stringent process. It highlights the strength of evidence for each of the reforms/interventions that have been identified in the literature in this review.

**Figure 4.1: Evolved Theory of Change**



| INTERVENTION | CONTRACT TEACHERS | MONETARY INCENTIVES | TEACHER CERTIFICATION | TEACHER TRAINING |
|---|---|---|---|---|
| **PATHWAYS** | Changes in accountability<br><br>Changes in incentives<br><br>Lower social distance<br><br>Changes in motivation<br><br>Reduction in class size<br><br>Reduction in multi-grade teaching<br><br>More effective recruitment | Encourage differing attitudes, depending on extrinsic/intrinsic teacher motivation<br><br>Gaming<br><br>Teaching to the test<br><br>Increase supply to teaching profession<br><br>Improved deployment in certain regions/ subjects | Improve the status of the profession<br><br>Ensure minimal standards | Encourage differing styles and methods of teaching<br><br>Expand teacher knowledge<br><br>Encourage student-centered pedagogy |
| **EVIDENCE** | Modest / Modest | Insufficient / Modest | Insufficient / Insufficient | Insufficient / Insufficient |
| **OUTCOMES** | Changes in Teacher Quality / Changes in Student Learning | Changes in Teacher Quality / Changes in Student Learning | Changes in Teacher Quality / Changes in Student Learning | Changes in Teacher Quality / Changes in Student Learning |

## CONTRACT TEACHERS

---

### Summary of findings

This review has identified eight studies that examine contract-teacher policies. Of these, two studies have been rated as of high quality and the remainder as moderate. The studies cover a number of country contexts, including India, Pakistan, Kenya, Niger, Mali and Togo. Of these, there are only two studies that are able to show 'impact', with one based in India and another in Kenya.

The evidence on contract-teacher reforms appears to suggest that, in most instances reviewed (five contexts out of seven studies), contract teachers' students do not perform any less well than those of regular teachers, and sometimes perform better. In all instances, student performance is measured in terms of test scores.

In terms of teacher quality, again the evidence appears to indicate that contract teachers are 'better' than their counterparts (in three out of four instances) where teacher quality is largely measured as teacher effort (proxied by absence rates) or teaching activity in school.

A notable finding, however, is that, while contract teachers exert greater effort compared to regular teachers, this appears to be more the case in their first contract period than in subsequent contract periods. Moreover, absolute effort levels continue to be low for both teacher types in certain contexts.

Some of the key pathways identified from the research have been the effect of the contract on teacher accountability (improving it), the reduction of social distance between the teacher and the student (with contract teachers being mainly recruited from the local community) and reductions in pupil-teacher-ratios, as well as the need for multi-grade teaching.

The situations where contract-teacher policies have been less effective have been when the threat of a contract not being renewed has not been credible, and where teacher unions have played mitigating roles in the implementation of the intervention.

The role of teacher unions and the extent to which the contract-teacher reform was implemented in a centralised, rather than a decentralised, manner have been identified as key barriers to the effective implementation of the reform, while the involvement and empowerment of parents and local communities emerges as a key pathway driving this reform effort in certain contexts. There is also evidence of contract teachers organising themselves to demand permanent contracts. Together with the evidence on decreasing returns over time, this could have implications for the sustainability of this reform.

---

### CONTEXT

In the face of teacher shortages and heavily constrained education budgets, policymakers across the developing world have turned to the use of contract teachers to help mitigate the impact this may have on the educational outcomes of millions of children across the globe (Kingdon et al. 2013). While the nature of these interventions differs according to context, the main driving force behind the implementation of this widespread intervention has remained the same, namely that low-cost teaching personnel hired on fixed-term contracts may help address teacher shortages in a cost-effective manner. However, the fundamental issue remains: namely, what impact has this potentially cost-effective solution had on the quality of teaching, as well as on student learning.

For example, Muralidharan and Sundararaman (2008) identified four main similarities, many of which can be applied to contract-teacher policies across different countries. Firstly, contract teachers tend to be appointed on renewable contracts with no guarantee of renewal. Secondly, they are often less qualified than regular teachers (this is not always the case in different contexts, but levels of training and qualifications do appear to differ between regular and contract teachers in most contexts). Thirdly, typically, contract teachers are paid substantially less than regular teachers. Finally, they tend to be recruited from the area where the school is located. Theoretically, one cannot determine the effectiveness of contract teachers as, on the one hand, the unfavourable conditions of non-renewable contracts can be deemed unfair and/or de-motivating, and the potentially lower levels of training may lead one to expect a lower quality of teaching thereof. However, on the other hand, further employment prospects of contract teachers are highly dependent on performance and, therefore, the contract status could be presumed to have a positive incentive effect. Additionally, the hiring of contract teachers could potentially have a positive effect on the education of children, as many of these teachers are hired from the local community and, therefore, may be presumed to be more accountable and socially less distant from their students.

## SYNTHESIS

In this review, eight studies examined for the impact of contract teachers on student learning and teacher quality were judged to be mainly of moderate quality. Of these, two studies attempt to answer Research Question 1 on impact (Bold et al. 2012, Muralidharan and Sundararaman 2013). The remaining six studies examine the relationship between contract-teacher interventions and outcomes of interest at scale answering Question 2 (Atherton and Kingdon 2010, De Laat and Vegas 2003, Goyal and Pandey 2010, Bourdon et al. 2005, Bourdon et al. 2007 and Habib 2010).

**Review Question 1: What is the evidence on the impact of reforms/interventions on education systems, at scale, to increase teacher effectiveness on: the quality of teaching and on learning outcomes in low- and middle-income countries?**

There are two studies addressing RQ1 within this intervention. Both studies show positive effects of contract-teacher interventions on student achievement and one shows positive effects on teacher quality.

In terms of experimental studies that investigate the impact of contract-teacher interventions on the outcomes of interest, the study by Muralidharan and Sundararaman (2013) presents evidence from the 'as is' expansion of the implementation of an existing contract-teacher reform in a state-wide programme in the Indian state of Andhra Pradesh. This reform reduced pupil-teacher ratios by allowing schools to hire an additional contract teacher. These teachers were hired based primarily on qualifications, followed by whether they were from the same village and, finally, on their teaching experience. However, whether or not this was adhered to in practice was also examined in the study. The study looked at a sample of 200 state-run rural schools (with 100 schools providing contract teachers). The study examines both student outcomes and measures of teacher quality (namely, teacher-absence rates). The authors find that students in schools with an extra contract teacher hired to teach directly to the students (rather than as an assistant) perform significantly better in both Mathematics and language - Telugu - (0.16 and 0.15 standard deviations, respectively), as compared to students taught by regular teachers. The results suggest that the gains are broadly distributed across all students. However, they do find that pupils in schools located in more remote areas receive more benefit from this extra contract teacher. In relation to teacher quality, the authors also examine the absence of regular, as compared to contract, teachers and find that contract teachers are significantly less likely to display absence than regular government teachers (18% as compared to 27%).

The only other study that examines the impact of this type of intervention is Bold et al. (2012), which examines the contract-teacher intervention that was based on hiring more than 18,000 teachers on fixed-term contracts, which was implemented in all eight Kenyan provinces, both by an NGO and by the Kenyan Government. The study uses RCT methodology to assign schools randomly into treatment and control groups, to identify the impact of the intervention. The authors of the study find that, while implementation by the NGO resulted in a positive effect on Mathematics and English scores (0.19 of a standard deviation), there was no effect when the same intervention was implemented by the Government (see discussion under RQ3).

Both papers discuss, in some detail, the political-economy factors that underlie their findings and, in particular, provide crucial evidence on scaling up of interventions and the need for future research in this regard. These findings are discussed at length when summarising the evidence on RQ3, below.

**Review Question 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?**

There are six non-experimental studies that investigate the relationship between contract-teacher interventions and the outcomes of interest. Overall, the findings are mixed.

There are two studies in India that use empirical approaches that allow us to identify the relationship between this large-scale reform and student learning/teacher quality. Atherton and Kingdon (2010) and Goyal and Pandey (2010) both evaluate this intervention in India. Each finds positive effects of the intervention on both student-learning outcomes and teacher effort.

Atherton and Kingdon (2010) use fixed-effects methodology on a sample of approximately 4,000 children in 160 rural schools in Uttar Pradesh and Bihar states in India to study the relationship between contract-teacher interventions and student learning of grade-2 and 4 students in Language (Hindi)and Mathematics. The authors find that pupils of contract teachers in Bihar score 0.069 standard deviations higher than those taught by regular teachers in the same schools. The effect is substantially higher in UP (0.208 standard deviations) — possibly due to contract teachers' facing higher accountability pressures with annually renewable contracts — than in Bihar, where teachers have jobs for life and, consequently, lower accountability pressures. The contract-teacher 'effect' remains even after the authors have taken into account the lower absence rates of contract teachers in their modelling, as well as other dimensions of teacher effort. Additionally, the authors cite lower social distance between students and contract teachers, as compared to students and regular teachers, as an additional reason behind the positive findings reported for the impact of contract teachers on student learning.

Similarly, Goyal and Pandey (2010), using data from schools across Uttar Pradesh and Madhya Pradesh, examine the relationship between the contract-teacher intervention in these states and teacher effort and student learning. Teacher performance is examined across three dimensions: 1) teacher attendance, 2) teacher engagement in teaching and 3) student test scores. All teachers teaching grades 1-5 form part of the sample used in this study, while students in grades 2, 3 and 4 are tested to measure student outcomes. The authors find that contract teachers have higher average attendance and activity levels compared to their regular counterparts in both states. Another notable finding is that teachers appear to exert higher effort in their first contract period than they do in subsequent contract periods. Moreover, the authors state that absolute effort levels continue to be low for both teacher types. The authors also find that, in both states, teacher activity is significantly and positively correlated with scores in both Mathematics and Language (Hindi), suggesting that the higher effort exerted by contract teachers is associated with the higher achievement scores of their students.

Habib (2010) adopts an interpretive qualitative-research design to identify whether levels of contract-teacher effort are higher than those of their regular counterparts in Pakistan. Using data arrived at through in-depth interviews from a sample of 16 regular, 16 contract and eight Head teachers from eight state schools in one district in Pakistan (Lahore), the author finds that absenteeism among contract teachers is only moderately lower than that among regular teachers. The author identifies certain features of the contract-teacher policy — fewer leave options, greater authority accorded to the school Head to check teacher absence, merit-based hiring and the threat of non-renewable contracts — that were deemed

by the interviewees to have helped reduce absence rates. However, the author discusses how, by failing to address other key perceived causes of teacher absence (transportation problems, low salaries, etc.), the contract-teacher policy could not be fully successful. And while, overall, the contract-teacher policy was found to have little impact on teacher absence rates, the policy did appear to have resulted in 'frequent resignation of contract teachers' (p. vi). This could be partly explained by the design features and implementation issues of the policy; the policy required contract teachers to be well qualified, but paid them relatively low salaries, and this, the author suggests, could have resulted in low motivation and greater turnover among teachers on contract in the country. The author argues that, by design, the contract-teacher policy only included one design feature that is typical of such incentive structures — lower pay —and this, in turn, could possibly explain the failure of the policy in the country.

There are three studies that investigate the contract-teacher intervention across the following set of countries: Niger, Mali and Togo. Bourdon et al. (2007) analyse the relationship between contract-teacher reforms and educational quality by estimating quantile-treatment effects using data from Niger, Togo and Mali. The authors use data from the PASEC on 2nd- and 5th-grade students and their achievement levels in Mathematics and French. PASEC surveys were carried out in Niger and Togo during the academic year 2000/01, and in Mali in 2001/02. The sample consists of all primary-school teachers available through the databases of the respective ministries of Education. Overall, the authors of the study find that, while specific characteristics can determine the success of the intervention, contract teachers are found to improve the outcomes of low-ability children in lower grades more than for higher-ability children in higher grades, suggesting that contract teachers may be better at teaching children performing at the lower end of the ability scale than the more able and advanced students. In terms of country-specific results, the authors found positive effects of the intervention in Mali, mixed effects in Togo and negative effects in Niger. These findings, the authors stated, are consistent with the ways in which the contract-teacher scheme was implemented across the three countries. The study has found positive effects of contract teachers on student learning in Mali, and this is consistent with close monitoring and hiring of these teachers within the local community (and, to some extent, this is also the case in Togo), while, in Niger, where the evidence is negative, contract teachers were hired through a more centralised manner, with no local monitoring. Bourdon et al. (2005) use a Propensity Score Matching method to identify the relationship between the contract-teacher intervention and learning outcomes, using PASEC data from Niger (2000/01). This study shows that the performance of students taught by contract teachers is no worse than that of students taught by regular teachers. In particular, the study finds, that while there is no achievement advantage of students in grade 5 taught by regular teachers (as compared to contract teachers), there appears to be a 'sizeable' (increase in performance of 6 percentage points, corresponding to almost 40% of a standard deviation and to 15% of the mean score across all 2nd-grade classes for Mathematics and French) advantage among students of regular teachers in grade 2. However, as soon as job experience is appropriately controlled for, this advantage disappears.

Vegas and De Laat (2003) used a fixed-effects technique to identify the relationship between contract-teacher interventions and student outcomes in Togo. The authors used PASEC (2000/01) data from Togo on 5th-grade students tested in Mathematics and French. Using a sample of 233 schools, the authors found that students of contract teachers underperform those of regular teachers (the coefficient on the contract-teacher variable in the regressions is large, negative and highly significant). This, the authors suggested, could be due to a decline in teacher quality, stemming from a decline in the quality of teacher entrants. The authors suggest that the positive effect of teacher education and outcomes may mean that the negative effects on student learning arising from a lowering in the quality of entrants into the teaching profession (due to a switch in contractual regimes) can be mitigated in a cost-effective manner through improvements in teacher education. However, in other contexts (such as India), positive contract-teacher effects have been found, despite the similarly lower level of CV characteristics (qualifications, experience, etc.). These have been explained through improvements in teacher effort and the incentives brought in with changes to the nature of the contracts.

**Review Question 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*?

The eight studies provide evidence to suggest a strong role of teacher unions in hindering this reform effort in certain contexts. Another potential barrier identified within the studies is the extent to which this reform was implemented in a centralised, as opposed to a de-centralised, manner. The possible empowerment of parents and communities in given contexts is seen as a driver of change.

One key barrier to the effective implementation of the contract-teacher policy has been identified as the role of teacher unions in undermining the effectiveness of these reforms. For example, as stated by Goyal and Pandey (2010), the process of teacher recruitment in India is highly politicised. Similarly, evidence from many of the states has shown that contract teachers are also organising themselves and placing pressure on state governments to regularise them. If this continues, and the contract-teacher scheme becomes a pathway to regular appointment, the teaching profession may end up with a labour force for which performance incentives are as weak as those for current regular teachers, but with a far larger number of unqualified and untrained teachers.

In this regard, the study by Bold et al. (2010) is highly significant, as it allows some of the key political-economy issues to come to the fore in respect of the contract-teacher interventions. This is because, in a unique experiment, the authors of the study evaluated the impact of contract-teacher interventions when implemented by an NGO, as compared to the government. In doing so, the authors showed that small-scale interventions of an RCT nature have provoked a number of criticisms with regards to the generalisability of their findings, and this has included concerns about external validity, general equilibrium effects, and the neglect of political economy in much of the evaluation literature (Acemoglu 2010, Deaton 2010, Heckman 1991, Rodrik 2009). As noted above, the authors found totally

different results of the intervention when it was implemented by the NGO (positive), as compared to when it was implemented by the government (zero). These differences, the authors stated, rest on the role played by teachers' unions, which 'waged an intense political and legal battle that successfully altered the contract-teacher program in subsequent years, in ways that may have undermined some of its incentive effects' (p. 4). The Ministry initially hired 18,000 contract teachers, which equated to nearly one teacher per school nationwide. Initially, these teachers were hired on non-renewable contracts for a two-year fixed-term period. However, the Ministry succumbed to intense political pressure by allowing the contract teachers both to become unionised, as well as subsequently hiring all 18,000 teachers at the end of their contract periods. The Government's ambitious plan to employ 18,000 contract teachers nationwide posed a significant threat to the Kenyan National Union of Teachers and, in this regard, this large-scale policy intervention provoked political-economy reactions from groups feeling threatened by the reform. This, ultimately, undermined the effectiveness of these reforms. The strong political resistance posed by the teacher unions in Kenya not only posed significant implementation challenges, but also eroded teacher accountability by forcing the Government to absorb all contract teachers into civil-service jobs. This, the authors argue, resulted in a large number of 'politically potent' teachers. Habib (2010) also suggests that this was one of the reasons for the failure of the contract-teacher policy in Pakistan. While the Government's reform effort was aimed precisely at reducing the political clout of teachers and preventing hiring based on political favours through the initiation of these large-scale reforms, in reality, there was intense pressure on the Government to regularise these teachers' contracts. Additionally, teachers highlighted the use of political pressure and connections by politically connected teachers that form a hindrance to achievement of the aims of any government reform intended to improve teacher effectiveness.

Another potential barrier identified within the studies is the extent to which this reform was implemented in a centralised, as opposed to a de-centralised, manner. Bourdon et al. (2007) suggest that the differentiated results across the countries studied can be explained by the way in which the contract-teacher scheme was implemented across the three countries, as discussed above. Many of the contract-teacher papers have suggested that one of the pathways through which this type of policy can affect outcomes is through the involvement and empowerment of parents and local communities. While contract-teacher interventions such as those implemented in Indian contexts have this similarity (that is, recruitment of teachers from the local community as contract teachers) with community-teacher reforms (for example, those in some of the West African contexts), it must be noted that both these types of interventions have different design features and objectives. Hence, community-teacher reforms and contract-teacher reforms may have different impacts, as well as distinct sustainability and viability characteristics. As stated by Bourdon et al. (2007), one important caveat is that relying heavily on the cooperation of these groups has the potential to reinforce existing inequalities. Therefore, entirely relying on poor communities to pay and monitor their teachers can lead to unacceptably regressive education policy. Therefore, there is a need to encourage local initiatives and autonomy, while also ensuring a pro-poor distribution of educational expenditure. Moreover, as noted by Muralidharan and

Sundaraman (2013), previous research on decentralisation has highlighted a potential concern that locally hired teachers may become the subject of abuse by local elites and that these jobs could become the object of patronage and result in high absence rates of those on whom they are bestowed. This paper indicates that this was not the case in this setting, as the decentralisation of hiring and empowerment of local school committees to hire said contract teachers led to lower absence rates being displayed by these teachers, significantly improved outcomes of the children they teach, and all this was done in a more cost-effective manner than civil-service hiring.

## MONETARY INCENTIVES

### Summary of findings

Evidence on this particular intervention is limited, with only five studies identified as focusing on monetary incentives for teachers, none of which considers impact. The studies cover a range of interventions: group incentives at school level, based on student performance in national assessments (Chile); salary increases, based on an assessment of teacher characteristics, as well as student test scores (Mexico); a salary premium given as a hardship allowance to recruit teachers to rural areas (Gambia); and a school-level subsidy based on student test scores, as well as a group-based teacher bonus (Pakistan).

The evidence on reforms aimed at providing monetary incentives to teachers is mixed, with two studies identifying positive associations of the incentive on student outcomes, and two studies identifying negative associations. In respect of whether monetary incentives improve teaching quality, there is evidence of this in only one study, albeit positive, and teaching quality is measured in terms of teacher qualifications.

The pathways through which monetary-incentive interventions can work mainly relate to providing teachers with stronger incentives. However, the multi-dimensional and complex role of teachers cannot be fully reflected in the outcomes of their students and, therefore, these interventions are often of more limited value. This may be particularly problematic where the schemes face challenges of gaming (i.e. where it is possible to abuse the system) corruption and/or teaching to the test.

Two of the five studies examine the SNED in Chile. This intervention rewards schools based on their pupils' performance. Both studies show positive effects of the programme for a subset of schools (in particular, those closest to the cut-off point for winning the award). The third study examines the Carrera Magisterial programme in Mexico, which was a teacher-wage reform that included wage increases for those teachers whose students performed well. The study does not find evidence of a positive impact of these monetary incentives on student achievement. The fourth study, which examines FAS in Pakistan, assesses the effectiveness of conditional cash subsidies to low-cost private schools. The paper takes the dual approach of examining whether positive incentives (group bonuses) or negative incentives (removal from the programme) can induce improvements in learning outcomes. The authors find that only the maintenance of minimum levels of learning to remain in the

programme are met through this set of incentives. The final study in this category examines the Gambian Hardship Allowance, which aimed to improve the provision of teachers in remote rural locations through salary incentives. There appears to be a suggestion emerging from this latter piece of evidence that, while the reform was generally 'successful', it did not reach the most remote parts of the country.

The role of teacher unions is highlighted as an important political-economy factor in some studies discussing these reforms, and the support or resistance they put forward appears to matter. It would seem that, where an intervention is well designed, effectively implemented, introduced gently in collaboration with stakeholders and showing results in the interim, it can be successful even if faced with resistance in the initial stages.

## CONTEXT

Teachers' monetary incentives have been the subject of controversy for programme-makers and researchers alike. Advocates of such programmes put forward the argument that teachers need stronger incentives, which reward them based on their pupils' attainment, rather than on observable characteristics such as their own educational attainment, training and tenure, which have been shown to be very weak in demonstrating how effective a teacher actually is. On the opposite end of the argument are those who argue, firstly, that a teacher's role is complex and multi-dimensional and, therefore, his or her effectiveness cannot be fully reflected in the outcomes of his or her students. Additionally, some would argue that teaching, as such, is an intrinsically motivated profession and, therefore, such extrinsic financial motivations are of limited value. Such incentive schemes also face corruption challenges, such as gaming the system or teaching to the test.

## SYNTHESIS

**Review Question 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?**

There are five studies that look at a type of monetary incentive of some sort, at scale. All five papers examine the relationship between monetary incentives to teachers and the outcomes of interest, although do not provide causal evidence . There are four studies investigating the relationship between monetary incentives and student achievement, of which two show a positive effect and two show no effect. One study investigates the relationship between this incentive and teacher quality, and finds a positive relationship between the two.

Of the five studies, two examine the only scaled-up teacher-incentive programme in the world, namely the SNED, in Chile, which covers 90% of Chile's schools. This programme is a teacher-effectiveness intervention that uses a monetary incentive to reward schools based on their pupils' performance; this monetary reward is normally distributed equally among all

teachers within the winning schools. The first paper assessing this programme is by Rau and Contreras (2012), which uses data from national tests over 10 years (1989-99). The authors use matching and double-robust methods and panel-data estimation and find a significant positive effect of the programme on Mathematics and Language (Spanish) test scores on students in the 4th, 8th and 10th grades. The results are robust to different model specifications and vary with a standard variation of 0.14 to 0.25 for Mathematics and from 0.09 to 0.23 for Language scores. In this paper, the authors provide support for educational policies that provide differentiation in the salary structure of teachers. The authors do note, however, that these results are only valid for the subset of schools they are able to consider. The second study to evaluate the same programme is by Mizala and Romaguera (2005), which presents preliminary evidence from the effect of the SNED programme on student outcomes. The paper uses a general fixed-effects model to show positive effects of the programme on educational outcomes for those schools that were the closest to the cut-off point for winning the award. The study also found that there was a change in teachers' attitudes, with teachers being more open to performance evaluation across all schools.

Another study from Latin America, by McEwan and Santibanez (2005), examines whether the monetary incentives provided to teachers through the Carrera Magisterial programme in Mexico (initiated in 1993) resulted in the improvement of student test scores. They also examine both positive and negative routes through which this may have been affected, namely, through increased teacher effort or gaming/abusing the system by teachers. Before the reform, teacher pay was determined based on CV characteristics: that is, education and years of experience. The reform allowed teachers and Heads to become eligible for significant permanent wage increases if they 'performed well' during a year-long assessment that emphasised student learning and school performance. Whist entrance into the programme was voluntary, according to the authors, a vast majority of Mexico's eligible teachers and school Heads participated in this programme. The authors used data from years 9-11 of the programme (1999-2002) and examined a sample of teachers from grades 3-6. The paper did not find robust evidence that teachers facing stronger monetary incentives improved student achievement in the year in which they were assessed.

The study by Barrera-Osorio and Raju (2010) evaluates a monetary-incentive intervention in FAS in the Punjab province of Pakistan. This FAS programme provides conditional cash subsidies to low-cost private schools and, in return for this assistance, the school has to satisfy certain conditions, including minimum levels of student outcomes. Additionally, schools are eligible for substantial group-based bonuses at school and teacher levels, based on their test scores. This paper examines whether stick incentives (removal from the programme) or carrot incentives (group teacher incentives) can induce learning gains in programme schools. Using data on student achievement from 2007-10, the authors find that the threat of programme exit for marginal first-time failures does tend to induce large improvements in student outcomes. However, the estimates do not show that the prospect of teacher bonuses induces learning gains for marginal bonus non-qualifiers. Therefore, the authors state that, on the whole, over and above the pressure from below to maintain minimum levels of learning for programme participation, the incentives to promote

continual student learning are not effective. It should be noted that there are limits to the generalisability of these results, as they apply to a very specific subset of schools that meet a set of minimum conditions pertaining to facilities and learning outcomes.

Pugatch and Schroeder (2014) provide evidence from the Gambian Hardship Allowance, initiated in 2005 and designed to relocate teachers to remote locations through provision of a salary premium of 30-40%. The incentive focused on primary-school teachers and aimed to attract qualified teachers to the poorest and most remote parts of the country. The size of the allowance was based on distance from the capital. The outcome of interest is teacher quality, measured in terms of teacher qualifications, and the authors adopted a difference-in-difference and regression-discontinuity design to identify the relationship between the intervention and the outcome of interest. The authors examine the issue of whether an increase in qualified teachers in hardship areas was met by new entrants or by the relocation of existing teachers. They find that the hardship allowance increased the number of qualified teachers in hardship areas by 10 percentage points. However, while the main goal of recruiting qualified teachers to rural areas has been achieved, it was less successful in respect of the most remotely located schools. The authors find that these improvements were due to the improved spread of current, qualified teachers, as well as improving the numbers of newly qualified teachers into the teacher labour market, while also acknowledging that the latter would be a preferable route for policy success.

**Review Question 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**?

These five studies provide limited evidence pertaining to RQ 3. The role of teacher unions emerges as an important factor affecting this reform effort.

One key aspect that arises in the implementation of monetary incentives is the role of teacher unions. Rau and Contreras (2012) highlight that, in many countries, such as those in Latin America, where teacher unions are very important, pay structures that recognise levels of productivity should, in theory, be very efficient. This is because the incentives created through pay-for-productivity schemes may lead to increased effort on the part of the teachers, thereby improving the quality of education and leading to improved student test scores. Their paper provides evidence showing how such mechanisms can increase student test scores. Similarly, Mizala and Romaguera (2005) argue that, as a result of the SNED intervention, there appears to have been a positive change in attitudes among teachers towards monetary incentives, which could explain why the teachers' union recently accepted a proposal to enhance the variable part of salaries that is linked to performance. Several surveys, the authors noted, show this change in teachers' traditional resistance to evaluation systems (p.141). These studies, therefore, highlight the potential drivers for change that can be effective in large-scale reform efforts. In this instance, it would seem that a well-designed and effectively implemented intervention, which was gently introduced through collaboration with stakeholders and which showed results, proved successful in overcoming initial resistance to reform.

## TEACHER CERTIFICATION

### Summary of findings

The evidence on this particular intervention is limited to only one study, in Indonesia. The study, comprising two pieces of research, helps to answer RQ1 and RQ2 (and, consequently, is able to identify impact and assess relationships) and suggests that teacher certification alone is not sufficient (at least in the context studied) to improve student outcomes (measured in terms of test scores) or teacher quality (measured in terms of teacher practices and behaviour in class).

Teacher certification can set minimum quality standards, as well as providing recognition to teachers that meet those requirements. This endorsement can also be a pathway to improving not only the reputation of the profession, but also potentially improving the calibre of candidates choosing to enter the profession. The reforms in Indonesia aimed to improve the quality of teaching, firstly by using the professional allowance as a means of attracting better-qualified entrants into teaching and, secondly, by improving the skills and competencies of those already within the profession by giving them the opportunity to acquire further qualifications. And, finally, the reforms aimed to improve teacher effort through recognition and increasing teachers' income as a means of motivating them towards being more productive. However, as this certification was based on factors other than merit, any positive impacts were muted. Certification must differentiate more effective teachers from less effective teachers, and must be based on proven competencies of an effective teacher to demonstrate results. Certification policies should also be linked to accountability reforms with sanctions, such as withdrawal of certification, forming part of the process.

The study authors also highlight bureaucracy, corruption and the lack of political will as some of the key factors that prevented successful implementation of the reforms within this context.

## CONTEXT

Teacher certification can set minimum quality standards, as well as provide recognition to those teachers that meet those requirements. This endorsement can also potentially improve the reputation of the profession and increase the calibre of candidates choosing to enter it. However, if this certificate is not true to form, and does not differentiate more effective teachers from less effective ones, its impact on educational quality will be minimal. In order, however, to be truly effective, certification policies should also be linked to accountability reforms, with sanctions such as withdrawal of certification forming part of the process.

## SYNTHESIS

There is only one study that examines this type of intervention and it allows investigation of all three research questions.

**Review Question 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?**

Chang et al. (2014) investigate widespread teacher-effectiveness reforms in Indonesia. The evidence presented shows no effect of teacher certification on either student learning or on teacher-quality measures. This book focuses on the comprehensive landmark Teacher Law reform of 2005, which encompassed several key features, including broader aspects of teacher-management development, which aimed at reforming the Indonesian education system. In particular, the reforms were aimed at improving the status of teachers by putting into place a massive scheme of academic qualifications (a four-year degree) and formal certification, combined with dramatic increases in teacher salary. The certification process is intended to improve the quality of teaching through: 1) 'the attraction channel'; that is, by using the professional allowance as a means to improve the attractiveness of the teaching profession and to encourage better-qualified entrants; 2) the 'upgrading channel'; that is, by giving those teachers who do not qualify for certification normally the opportunity to do so by acquiring a four-year degree and thereby improve their skills and competencies to improve teaching and student learning; and 3) the 'behavioural channel', aimed at improving teachers' recognition and doubling of income to motivate them and make them more productive. The reform itself placed eligibility criteria (four-year university degree, high rank in the civil service or a very senior teaching post) for certification. This meant that all teachers in the system would eventually have minimum levels of defined competencies. This certification was aimed at improving teachers' welfare and increasing their status and recognition. Additionally, the programme resulted in the doubling of the teachers' wage bill. An important feature of the reform was the consequent doubling of teacher incomes, which was *permanent* and *not conditional on subsequent performance*, except for a requirement to teach a certain number of hours per week.

Chang et al. (2014) extract evidence for the synthesis based on two studies. The first, using evidence from an RCT to evaluate the effects of certification and the corresponding increase in income on teacher productivity helps answer RQ1. The study finds that professional certification improved teachers' well-being, but does not necessarily make them 'better' in terms of the student outcomes they produce. This may have been due to the fact that early rounds of certification were based on seniority, rather than merit, and, therefore, the process of certification started with a low level of competence. Additionally, the study found that the process of academic upgrading, while it may have improved the quality of candidates enrolling in teaching, has not automatically translated into substantial steps forward in terms of better performance in the classroom.

**Review Question 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?**

Study 2 within Chang et al. (2014) looks inside the 'black box' of teaching practice by undertaking a video study of 8th-grade Mathematics classrooms that were participants in the TIMSS. This study, therefore, helps answer RQ2 (see above). This study, based on 200 teachers, finds that, while there was no difference between certified (according to the reform) and uncertified teachers' subject matter and pedagogy, teacher knowledge stood out as having a strong association with student learning. The findings confirm that certification alone does not have an impact on teacher practices and behaviour. There are striking differences in the teaching practices of teachers with greater subject and pedagogical knowledge compared to those with less, and with teachers holding Mathematical Education degrees (compared to those with pure Mathematics degrees), and both these groups tended to use practices that are associated with improved student-learning outcomes (pp.150-151).

**Review Question 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?**

In terms of evidence on RQ3, the book discusses evidence from a political-economy analysis (PEA) and provides substantial evidence of financial and political-economy factors that may have distorted the impact of the Indonesian teacher reforms. The authors highlight bureaucracy, corruption and the lack of political will as some of the key factors that prevented successful implementation of the reforms.

The authors of the study note that many teachers entered the certification route through nepotism or outright corruption, rather than merit, thereby potentially muting any positive impacts. Secondly, the authors highlight the fact that the bureaucratic environment can be so deeply entrenched within the system that it is difficult for people who have been part of this system to have the ability or the initiative to become agents of change. This reform involved an automatic doubling of teacher salaries based on certification. However, as mentioned above, because the certification was based on factors other than merit, the authors highlight a key policy pointer: namely, that, where countries are considering instituting teacher-pay increases as part of reforms, they must be based on the proven competencies required of an effective teacher. This must be in addition to ensuring that those who are ineffective in performing their teaching duties do not remain in the teaching profession. Moreover, politicians and senior leaders need to be drivers of change in ensuring that high standards are implemented, that competency is accurately and effectively assessed and implemented, and that teachers who do not meet the required standards may be dismissed. As with all large-scale reform efforts, the teacher reform in Indonesia also required a significant financial outlay from the Government. One of the challenges faced as a

consequence of this reform in Indonesia is that it is attracting more and higher-quality candidates into teaching at the expense of other fields. Therefore, although the reform has had little impact on the quality of current teachers or student outcomes, it may produce a better cadre of teachers in the future, provided minimum competency standards are enforced at all stages of teacher management and development, the growth of poor-quality teacher-education institutions is curtailed and, finally, if qualified graduates have merit-based access to jobs on completion of their degrees (p. 185). The authors argued that another key barrier to successful implementation is the existence of significant inefficiencies associated with teacher hiring and deployment, which need to be addressed to ensure the reforms actually achieve their desired results. This highlights shortcomings in even the most comprehensive reform programmes, where factors outside of the scope of the reforms still have the ability to hinder the desired outcomes.

## TEACHER TRAINING

### Summary of findings

Evidence on teacher-training reforms at scale is limited to one study, in Ethiopia. The evidence helps answer RQ2 and reports positive evidence for both improved student outcomes (test scores) and teacher quality (greater knowledge and improved pedagogical practices).

The study examines an in-service teacher-training programme and finds that the training programme was particularly effective in improving test scores (especially for girls). Trained teachers did appear to be making more effective choices in respect of pedagogical methods used, but they did not appear to be more cognizant of identifying relevant strategies for teaching, nor did they appear to adopt a more reflective teaching method. Additionally, the range of pedagogical practices adopted by trained teachers was not any wider than that used by non-trained teachers. This would suggest that it is not the methodologies that teachers use per se, but how they use them that ultimately impacts student learning.

## CONTEXT

Low teacher quality is a problem faced by many countries the world over. Policymakers have, therefore, looked at potential solutions to address this issue and many have turned to professional development as a means to improve the learning of children. This professional development can take the form of both pre-service and in-service training, with the latter being seen as a more efficient means of impacting a larger proportion of teachers in a shorter period of time. The effectiveness of the training, however, depends mainly on the quality of the training programme and whether it not only updates the trainees' content knowledge, but also whether it assists them in imparting that knowledge to their students in the most effective manner.

**Review Question 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?**

There is only one study that appears to evaluate a large-scale training programme and has met the stringent requirements to be considered for in-depth review. This study is in Ethiopia and, because it uses non-experimental methods, it helps answer RQ2. Piper (2009) evaluates one part of a large-scale programme that included in-service training, pre-service training, and capacity building for the decentralised education system. In his thesis, Piper specifically evaluates the cluster-based teacher-education programme — the ITPD — part of the BESO II/BEP programme. This particular intervention aimed to use student-centred strategies to improve teacher knowledge and skills, and thereby improve their pedagogy, with the ultimate aim of improved educational outcomes. The author uses a mixed-methods design to evaluate programme effectiveness. In one part of the study, the author uses multi-level difference-in-difference methodology to show that the programme improved student learning of 4th-grade students by a standard deviation of between 0.2 and 0.4. Using national data sets (from 2000 and 2004), the author shows that the programme was especially effective for girls' achievement, and for those with unqualified and experienced teachers. In the qualitative part of the study, the author sampled 10 case-study urban schools and collected interview and classroom-observation data on 4th-grade Mathematics teachers. The author argued that the increase in levels of student achievement witnessed in the quantitative analysis could not be explained by a higher incidence of student-centred pedagogy among the trained teachers. While trained teachers displayed greater knowledge, they did not appear to adopt more student-centred pedagogy as compared to untrained teachers. The training did not appear to change the range of pedagogical methods used; however, the trained teachers appeared to make more effective choices in respect of which methods to use and apply during teaching. This would suggest that it is not the methodologies that teachers use per se, but how they use them that ultimately impact student learning.

## SUMMARISING KEY FEATURES OF THE INTERVENTIONS IN THE INCLUDED STUDIES

Table 4.7, below, summarises the key features of the reforms that have been discussed in the 15 studies synthesised above.

**Table 4.7 summarises the reforms discussed within each of the studies reviewed in this section.**

| Studies | Country | Reform Features | States/provinces analysed |
|---|---|---|---|
| **Contract Teacher Interventions** | | | |
| **Atherton and Kingdon (2010),Goyal and Pandey (2010), Muralidharan and Sundararaman (2013)** | India | Contract-teacher reforms — implemented across India as part of the primary-school reforms over the last two decades. Nationally, contract-teacher salaries tend to be 35% of regular-teacher-pay rates, and this is likely to have fallen more after the increase in regular teachers' salaries, following the Sixth Pay Commission increases. Levels vary across states, but, generally, contract/para teachers are hired on fixed-term (typically annually renewable) contracts, typically don't have pre-service training or very stringent qualifications requirements, and are hired from within the community. They are also usually appointed in remote schools that serve disadvantaged children. | States of Uttar Pradesh, Bihar, Andhra Pradesh and Madhya Pradesh. |
| **De Laat and Vegas (2003), Bourdon et al. (2007)** | Togo | Contract-teacher reforms — initiated in Togo in the 1980s and 1990s to cope both with slowing economic growth and the resultant decline in Government resources. The Government responded by freezing public-sector wages. There was also significant overcrowding in schools and a resultant overhaul of teacher-hiring policies, with the Government shifting hiring responsibilities away from relatively expensive civil-servant functionaries to *contractuels* or *auxiliaires*. In Togo, the reform was initially driven by local communities who employed their own private teachers when the Government failed to provide the required number of staff. Most | Not clear, random sample of 233 schools based on PASEC data. |

| | | contract teachers attend regular teacher-training institutes and tend to have the same qualifications as regular teachers, but receive, on average, 40% of civil-servant teacher wages (depending on qualifications), have fewer promotion rights, lower pensions, etc. The authors stated that only 45% of current primary-school teachers in Togo are regular teachers, and 55% of them are contract teachers. Bourdon et al. (2007) also explain how the reform was triggered by initiatives taken by local communities who engaged their own community teachers on contract when the state failed to provide them with teachers. Contract teachers are meant to have at least Junior Secondary education; may have from no training to up to three years of training; often have no career plan; and their contracts are supposed to be permanent. However, contract teachers are in more 'insecure' jobs than their regular counterparts. Therefore, they are considered part of the contract-teacher labour force, rather than the regular-teacher labour force. | |
|---|---|---|---|
| **Bourdon et al. (2005, 2007),** | Niger | The contract-teacher policy was put in place in 1998, and no more civil-servant teachers were hired at primary level thereafter. A significant number (2,800) of teachers were hired on contract after 1998, compared to an annual average of 520 between 1990 and 1998. Arguably, the unsustainable level of salaries paid to regular teachers is what triggered this reform, and contract teachers were hired on one-third of the salary of regular teachers. The contracts becomes permanent after four years. | Not clear |

| Bourdon et al. (2007) | Mali | In Mali, the Government was recruiting contract teachers as early as 1991, but numbers became high regionally only towards the end of the decade. In Mali, the contract-teacher reform also stemmed from initial community initiatives to hire local teachers when the state failed to meet communities' teaching needs. Post-1998, regular and contract-teacher hiring were roughly equivalent. Contract teachers made up 30% of teaching staff in primary schools in 2000 and almost 65% by 2004. Contract teachers earn significantly less than their regular counterparts (initial rates were about 25% of what regular teachers earn for contract and 15% for community teachers, respectively). Contract teachers are expected to have at least 11-12 years of education/training, are provided with three months of professional, pre-service training, and have permanent contracts. As in other contexts, described above, the fact that they have more insecure jobs (while the contracts are permanent — that is, not for a fixed term — they do not guarantee tenure) than regular teachers qualifies them to be considered part of the contract-teachers scheme for the purposes of this review. | Not clear |
| Bold et al. (2012) | Kenya | The government of Kenya in 2009 initiated a reform to provide funds to employ teachers on contract outside the Teacher Service Commission (TSC) system, which had formerly been responsible for the hiring of all civil-service teachers in Kenya. In a bid to 'scale up' the hiring of contract teachers, the Government subsequently, in October 2010, hired 18,000 teachers on contract: | All eight Kenyan provinces |

| | | nearly one teacher per school. Initially, these teachers were hired on non-renewable two-year contracts and paid US$135/month. However, in 2011 the Ministry succumbed to union pressure and subsequently integrated these teachers into the civil service. Critically, the scale-up of the project in similar NGO schools demonstrated effects on a comparable scale, but had no effect in government schools. It was not the going to scale that changed the impact, but, rather, operating in a government system with more complex political-economy realities. | |
| --- | --- | --- | --- |
| **Habib (2010)** | Pakistan | The contract-teacher policy was introduced as one of seven strategic initiatives under the Punjab Education Sector Reform Programme (PESRP, 2001-04) by the Federal Ministry of Education. The contractual hiring of teachers was one of the seven initiatives and was meant to be aided by the National Devolution Plan 2001, which was intended to give a greater level of autonomy to provincial and district governments in implementing educational services. The contract-teacher policy introduced a sanction for dismissal as a result of unsatisfactory attendance and fewer leave options for teachers on contract. It was also meant to allow for merit-based hiring (test scores on BA exams, work experience and training, preference given to female teachers). According to the policy, from 2002 onwards, all new state-school teachers hired in Punjab were given five-year contracts, renewable dependent on performance criteria, including attendance. No new regular civil- | Punjab province |

service-teacher hiring was meant to take place under this reform. In 2003, 13,000 new contract teachers were hired and 16,000 new contract posts were approved in 2004-05. A minimum qualification requirement of a Bachelors degree was set, and teachers were required to have a Primary Teaching Certificate (PTC); in terms of educational qualifications, contract teachers were better qualified than their regular counterparts. Contract teachers were hired on lower salaries and did not have a clear career-progression path. The teachers were hired on five-year site-specific contracts, renewable based on performance.

**Monetary Incentives**

| | | | |
|---|---|---|---|
| **Mizala and Romaguera (2004), Rau and Contreras (2012)** | Chile | Since 1996, the government of Chile has adopted a monetary-based productivity bonus called the SNED, which allows the Government to link teachers' salaries to their performance, across government schools. The SNED rewards teachers and aims to improve their motivation. A rank-order tournament, this performance-based incentive is aimed at all municipal and privately subsidised schools in the country, which enrol 90% of the students. The incentive was awarded based on pupils' results on standardised tests. Schools with similar characteristics were grouped into homogenous groups and the competition took place within each distinct group. Students are, therefore, ranked within appropriately defined comparison sets, so that teachers are only competing with other teachers who work in similar schools and, therefore, potentially | National level |

| | | teach similar students. The incentive is allocated at school level (and is, therefore, a group incentive) based on students' performance on a national assessment, and the rewards are allocated equally among the teachers in the school. For the 1996-97 SNED competition, the winning teachers were allocated about US$370, approximately 40% of a teacher's monthly income, equivalent to a 3.33% salary increase. | |
|---|---|---|---|
| **McEwan and Santibanez (2005)** | Mexico | The Carrera Magisterial programme was a teacher-incentive programme initiated in January 1993 as one component of a large-scale educational-reform programme, known as the National Agreement for the Modernization of Basic Education. Prior to this reform, teacher and school-administrator salaries were determined by levels of education and experience. Post-reform, teachers and school Heads became eligible for significant salary increases if they performed well in a year-long assessment process, which included assessment of the teachers' levels of education, years of experience, professional development, peer review, teacher/principal knowledge demonstrated in a test score, and their students' test scores (with points awarded for each factor and given a different weighting, adding up to a total score out of a possible 100). The process of engaging in this was voluntary, but a large majority of schools took part. Overall, the Carrera Magesterial allows five levels of promotion: A, B, C, D and E, with each representing a successively larger bonus and each individual starting from A. Once promoted, there is no | National level |

possibility of demotion and the wage bonus is received in all subsequent years. In recent years, according to the authors, teachers in band A have received 24.5% of their wage pay as a bonus, hence, the bonuses represent quite significant amounts.

| **Pugatch and Schroeder (2014)** | The Gambia | In 2005, the Gambian Government introduced an incentive policy aimed at attracting teachers in lower basic grades in state-run rural schools. The Hardship Programme, intended to provide an additional allowance to teachers in lower basic grades in the three (out of the six total) regions that are furthest from the capital and most economically disadvantaged. Schools were classified as being in hardship if they were located more than 3km from a main road. The reform programme allocated the following incentives: allowances of 30%, 35% and 40% of salary, depending on how far the region was from the capital. Both qualified and unqualified teachers received this salary premium. The premium is large relative to the average teacher's salary of US$67/month (before hardship allowance). | National level |

| | | | |
|---|---|---|---|
| **Barrera-Osorio and Raju (2010)** | Pakistan | Initiated in 2005, the FAS programme aimed to provide publicly funded conditional cash subsidies to low-cost private schools, with the objective of offering good-quality schooling opportunities to the disadvantaged. The programme was funded through the Punjab Education Foundation (PEF), a publicly funded, semi-autonomous organisation established in 1991, which serves as the main institutional conduit for Public Private Partnership (PPP) programmes in Punjab province. In exchange for receiving the subsidy, the school has to waive tuition fees for all students and ensure all school students achieve a minimum pass rate in a curriculum-based test (the Quality Assurance Test). Programme schools meeting these subsidy conditions are also eligible for group-based teacher bonuses and school bonuses, based on rankings in QAT scores. As of June 2010, the FAS programme had reached 798,000 students in 1,779 schools in 29 of the 36 districts in Punjab, making it one of the largest PPP initiatives in the developing world. | Punjab province |
| **Teacher Training** | | | |
| **Piper (2009)** | Ethiopia | The Basic Education System Overhaul (BESO I), Basic Education Strategic Objective (BESO II) and Basic Education Program (BEP) were teacher-training programmes initiated at different phases in Ethiopia, but, essentially, they used cluster-based teacher training and professional-development approaches to enhance the capacity of teachers. The BESO I programme began as a pilot in two regions of the country in | National level |

| | | 1995, and expanded to all schools in the regions by 2002. The BESO II (later renamed BEP) innovation began in 2002 and continued until 2007. This innovation developed teacher self-instructional kits, handbooks and specific support materials for women teachers, aimed at promoting student-centred learning, pedagogical methods, classroom management of large classes, etc. This training program adopted both cluster and cascade teacher-training models. This US$30m USAID-funded programme was instituted in all 11 of Ethiopia's regions and city administrations. It included three main interventions: pre-service training, in-service teacher professional development (ITPD), and capacity building for the decentralised education system. This thesis evaluates the ITPD aspects of the BESO II/BEP. | |
|---|---|---|---|
| **Certification** | | | |
| **Chang et al. (2014)** | Indonesia | In 2005, the Indonesian Government approved a comprehensive Teacher and Lecturer Law that was aimed at transforming the teacher-management process in the country. This came to be known as 'Teacher Law 2005 and covered all aspects of teacher management, including development of: a) the competencies required of teachers in four main areas (pedagogy, personal, social and professional); b) their incorporation into national teaching standards; c) the role of various ministerial units and agencies in supporting teachers in reaching these stipulated competencies; d) the teacher-certification process and the qualifications needed by teachers; and e) the conditions under which | National level |

teachers could receive special/professional allowances. The Law also raised important points regarding issues such as continuous professional development and promotion and salary increments, and, as such, was seen to provide a comprehensive package of reforms for improving national education.

## 5 IMPLICATIONS

## 5.1 SUMMARY OF RESULTS OF SYNTHESIS

We sought high-quality evidence on reforms/interventions in education systems aimed at improving teacher effectiveness, at scale. The review sought evidence that helped answer each of the following questions:

*RQ1. 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?*

*RQ2. 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?*

*RQ3. 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*?

This review of 15 studies has found evidence on four key interventions in the literature: contract teachers, monetary incentives, teacher certification and teacher training. A summary of the findings and Weight of Evidence pertaining to each of these is discussed below.

### CONTRACT TEACHERS

Four studies examine the question of whether the contract-teacher intervention in a particular country context is related to the 'quality' of teaching. In three of these contexts (including one that shows impact), the authors have found a positive effect of contract teachers on teacher quality (variously measured), while, in one context, no effect is found. In examining contract teachers' association with student achievement, of the seven studies that do so, in five contexts, we noted a positive effect, while, in two contexts, the authors found a negative effect and, in two, there appears to be no effect. Therefore, the authors of the review found modest evidence in relation to this reform and the outcomes of interest. The studies cover a number of country contexts and appear to suggest that, in most instances reviewed, contract teachers' students do not perform any less well than those of regular teachers, and sometimes perform better. In terms of teacher quality, again, the evidence appears to indicate that contract teachers are 'better' when compared to their counterparts. There is some evidence to suggest that not only are absolute effort levels low for both teacher types, but, in certain instances, there appear to be diminishing returns for contract-teacher policies. Moreover, with contract teachers organising themselves into effective and powerful bargaining groups, in several instances, the characteristics of the contract mutate into a 'regular' format, with low accountability and similarly low incentives to permanent government-teacher jobs, resulting in no difference in effect between the two

contract types. Together with the evidence on decreasing returns over time, this could have implications for the sustainability of this reform.

## MONETARY INCENTIVES

Five studies examine the relationship between monetary incentives and the two outcomes of interest. The interventions include group incentives at school and teacher levels, relocation and hardship allowances, as well as salary increases. The WoE in respect of teacher quality is found to be insufficient, with only one study showing a positive effect of the intervention on teaching quality, as measured in terms of teaching qualifications. There is modest evidence in relation to student achievement. Of the four studies, a positive effect is found in two contexts and no effect in two contexts. Monetary incentives could be argued to be the most efficient manner in which to increase teacher productivity. However, the research studies examined in this review highlight that one major obstacle to this reform effort are stakeholders such as teachers' unions. However, in certain contexts, unions can be an important driver of change, as shown by the SNED intervention in the Chilean context. This is of particular relevance in light of the initial resistance that was faced in the early stages of the implementation of the intervention.

## TEACHER CERTIFICATION

Teacher certification has been a policy of choice in many settings to implement minimum quality standards, and by recognising teachers who meet those requirements improving the status of the teaching profession and, therefore, potentially, the calibre of new entrants into the teaching labour market. This review has identified two pieces of research within one study that show no effect of teacher certification on teacher quality, and show no effect of this intervention on student achievement. The study looks at a teacher-certification programme in Indonesia and examines whether this intervention improves student outcomes and/or teacher practices or behaviours in class. The authors find that, because the certification was based on factors other than merit, any potential positive impacts were muted. The authors also highlight the key policy pointer that, where countries are considering instituting pay raises for teachers as part of other reforms (this reform included an automatic doubling of salaries), it must be based on proven competencies required of an effective teacher in order for these policies to demonstrate results. The authors do state that more recent implementation of this policy has been less influenced by corruption and nepotism and, therefore, future evaluations should show more promising results. In light of the fact that there is only one study in this category, it can be concluded that there is insufficient evidence about this particular intervention. Even the most comprehensive of reforms, when not accompanied by continual incentives and mired in bureaucracy and nepotism, may not be sustainable and can fail to show adequate rewards.

## TEACHER TRAINING

The review has identified one study examining the relationship between teacher-training reforms and teacher quality and student achievement. The study looks at an in-service training programme aimed at promoting student-centred learning, improved pedagogical methods and better classroom management of larger classes. The author of the study being reviewed concluded that the programme improved student test scores (particularly for girls). Trained teachers also showed more effective pedagogical choices and, while the methodologies they used did not change per se, how they chose to use them ultimately impacted on student learning.

## 5.2 STRENGTHS AND LIMITATIONS OF THIS SYSTEMATIC REVIEW

One of the key strengths of the research approach taken in this review is the adoption of a broad view of evidence to be included and the heterogeneity of the types of publications and research designs that were ultimately incorporated. As a result of this, the amount of information that could be included was maximised and the narrative-synthesis approach allowed incorporation of a diverse range of studies, adopting varied methodologies and research designs in difference disciplines that could be considered for review.

Another key strength of this review was the adoption of a comprehensive and specific definition of both 'reforms' that could potentially be allowed for inclusion, and of the outcome variables: namely, teacher quality and student outcomes. More flexibility was used in relation to the definition of 'scale', so as not to constrict the usefulness of the studies thereby identified. Additionally, the fact that the scale criterion was only applied in the latter stages allowed for pertinent information from the non-scale evidence to be filtered through to the final report. This has also created an additional resource for users of the review who may be interested in identifying and updating the list of studies that have not formed part of the in-depth review in this instance (due to not meeting the scale requirement), but may nevertheless assist in answering other pertinent questions.

While the approach taken has several strengths, it comes with a set of related limitations. As with all reviews, one of the limitations is the extent to which such differing studies, adopting various methodological approaches, can be appropriately compared. However, while this may, to a certain extent, restrict the ease with which general conclusions can be drawn, the rigorous discussion of these included studies and their particular contexts allow the presentation of a very extensive range of views. And, while strict quality control and process systems were implemented in the conduct of this review, as with all reviews, the element of subjectivity of the reviewers can always be a limitation.

Additionally, as the previous chapter has highlighted, there is only a limited number of studies that examine the true impact of teacher-effectiveness reforms, at scale, on the outcomes of interest. And, while there are several studies that utilise stringent techniques, many of these tend to be in a cross-sectional, rather than panel-data context, all of which lack the quality of data on which rigorous research in this field has been based. One

important caveat to note is that even experimental techniques have their drawbacks and, by allowing the second research question in this review, it is intended to present a far more useful output than if only experimental evidence were used.

Finally, while this systematic review has implemented very stringent search processes and subsequent attempts to access varied literature from local and international sources, the final evidence base may have resulted in some relevant research not being included, due either to the fact that it may not be publicly available, or that it is published in a language that has not been used for searching in this review.

## 5.3 IMPLICATIONS FOR POLICY, PRACTICE AND RESEARCH

This systematic review identified 15 studies that met the inclusion and quality criteria. And, while the authors adopted a broad definition of what constitutes teacher-effectiveness reforms, the final set of studies included for the in-depth review was limited, not only in the range of interventions identified (contract teachers, monetary incentives, teacher certification and teacher training), but also in the final WoE. There was a relatively broad geographical base from which the overall evidence is drawn. However, the results remain very context-specific. Unfortunately, the evidence covered in this systematic review does not give us sufficiently detailed information to uncover whether it is specific design features and/or contextual factors (and how much of each) that have driven the effect/relationship that is observed in the study. Given the variation in the design of reforms and contexts, it is, therefore, important for future research to consider both in detail, and subsequently to attempt to disentangle the independent and joint effects of each. For example, teachers vary greatly across different contexts, and the human resources available to different governments are also widely varying. In some instances, the existence of a private sector for teacher training could result in the surplus of training graduates available to take up contractual positions, while, in others, this may not be the case. All of these differing contextual features will ultimately impact the effectiveness of interventions, and future research should endeavour to uncover more of these factors. Furthermore, the intervention itself may vary in design or implementation, depending on surrounding factors. Therefore, you could have an identical intervention implemented in different contexts that results in different outcomes, or opposing interventions that have identical outcomes due to different contexts. Unless one fully investigates all of these factors, even the most rigorous impact evaluations will not provide meaningful directions for future policy.

Many of the studies (8/15) focused solely on one particular type of intervention: contract teachers. Moreover, most of the studies helped answer RQ2 (that is, the relationship between the intervention and outcomes), rather than helping to identify causal relationships (RQ1). Most importantly, there was limited evidence on RQ3 within the studies and any that did exist was merely in passing (with the exception of a few studies). This is of concern if the research cannot identify what the drivers of change or hindrances are to a policy's being effective (or not), as it limits the usefulness of that evidence for future policymaking.

There is a clear paucity of high-quality research to address all three questions in the developing world, particularly when one considers the issue of scale. The stringent and detailed review undertaken here has confirmed this. This could, in part, be due to the fact that not enough resources have to date been allocated to fund quality education research (as compared, for instance, with other sectors, such as health and agriculture). Part of the problem also lies in how 'scale' is defined. To date, any research that has focused on scale tends to define it in a one-dimensional way; for example, simply by looking at the number of schools reached by an intervention. Future research should focus on improving this traditional definition of scale to one that is multi-dimensional in nature and allows for research in this field to be brought together and synthesised. The research team struggled particularly with definitional issues, as many of the dimensions of scale are less easily measured than others. For example, it is much more challenging to measure conceptual change than the presence or absence of materials or activities (Coburn, 2003).

The key policy message emerging from this review is the need for focused and specific research to help answer the critical questions raised in this review, and, in particular, RQ3. There is a clear need for research focusing on how educational systems operate. If anything, this review has emphasised the need for this kind of research, which comprehensively uncovers the working of systems to identify the underlying ingredients of successful reform programmes that can be taken to scale to improve student learning.

The evidence reviewed here shows that the relationship between teacher-effectiveness reforms and the outcomes of interest do, in several contexts, manifest themselves in a manner similar to that suggested in the Theoretical Framework. However, in what way and to what extent these processes result in change varies from context to context. This review suggests that there is an urgent need to devise policies and encourage research in this field in a more concerted manner. In particular, as a generalization, it can be said that there is robust evidence. However, in many instances, this evidence is not available to drive policy design and implementation. And, while this is, in and of itself, not necessarily a problem, it does need to be highlighted, because these reforms, which may actually be highly effective, do not appear in bodies of evidence such as ours because they are not the subject of rigorous evaluation. Therefore, this review is intended to encourage not only more research-driven policies, but, equally importantly, where policies and reforms have been implemented, research into why and how these reforms may or may not have worked, to feed into effective policymaking for the future. Research and policy initiatives are integral parts of the educational mechanism that should be moving in tandem, each contributing to the other. Currently, it would seem that policy is being driven by decision-makers' ideologies, voter considerations and political expediency, while research, by default, tends to be driven by reliance on donor interests and data availability. In this regard, Bold et al. (2012) can be used as an example of good practice, wherein researchers and the government align initiatives to implement and evaluate an intervention, with contrasting results. Studies such as this should be conducted in different contexts to broaden our understanding of teacher-effectiveness reforms and their potential effects.

A pertinent suggestion resulting from the authors' perusal of the literature base is the need for researchers, going forward, to consider adopting data-collection methods that allow for more nuanced analyses. The evidence has highlighted the independent existence of studies such as those by Muralidharan and Sundararaman (2013), which adopt elaborate and convincing experimental methods that help answer RQ1 in a given context. Contrastingly, there are other studies that adopt different methodologies, such as that by Habib (2010), which tries to put forward explanations for certain phenomena, rather than aiming to identify causal relationships. However, the research world is increasingly recognising the importance of high-quality mixed-methods approaches, such as those adopted by Piper (2009), wherein attempts are made to identify not only causal relationships, but also to understand the underlying conditions and factors supporting or hindering them. In this regard, the authors of this review would strongly urge the research community to consider adopting more sequential research designs that involve undertaking inexpensive qualitative research beforehand, followed by more expensive evaluations, such as RCT. In this way, the triangulation of findings will present a more extensive evidence base on which to inform future policymaking and research initiatives.

## 6 REFERENCES

### 6.1 STUDIES INCLUDED IN MAP AND SYNTHESIS

*Atherton, P. & Kingdon, G. (2010). The relative effectiveness and costs of contract and regular teachers in India. CSAE Working Paper, University of Oxford, WPS/2010-15.

Banerjee, A., Cole, S., Duflo, E. & Linden, L. (2003). Improving the Quality of Education in India: Evidence from Three Randomized Experiments, NBER Working Paper Series, http://users.nber.org/~confer/2003/chs03/duflo.pdf

Banerjee, A., Cole, S., Duflo, E. & Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *Quarterly Journal of Economics* 122 (3): 1,235-1,264.

*Barrera-Osorio, F. & Raju, D. (2010). Short-run learning dynamics under a test-based accountability system: Evidence from Pakistan. Policy Research Working Paper, World Bank, WPS 5465.

Behrmen, J., Parker, S., Todd, P. & Wolpin, K. (2011). Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools. Working Paper, University of Pennsylvania.

*Bold, T., Kimenyi, M. & Mwabu, G. (2012). Interventions & Institutions Experimental Evidence on Scaling up Education Reforms in Kenya. Working Paper. Institute for International Economic Studies, Stockholm University and Goethe University Frankfurt,

*Bourdon, J., Froelich, M. & Michaelowa, K. (2005*). Broadening Access to Primary Education: Contract Teacher Programs and Their Impact on Education Outcomes in Africa: An Econometric Evaluation for Niger*. Conference Paper for the Annual Conference of the Research Committee Development Economics of the German Economic Association on Pro-Poor Growth, 8-9 July 2005.

*Bourdon, J., Froelich, M. & Michaelowa, K (2007). Teacher Shortages, Teacher Contracts and Their Impact on Education in Africa, Discussion Paper, Paper No. 2,844, Institute for the Study of Labour (IZA).

*Chang, M., Shaeffer, S., Al-Samarrai, S., Ragatz, B., de Ree, J. & Stevenson, R. (2014). Teacher reform in Indonesia: The role of politics and evidence in policy making. World Bank. Washington D.C.

*De Laat, J. & Vegas, E. (2003). Do differences in teacher contracts affect student performance? Evidence from Togo. Working Paper. World Bank. Washington D.C.

Duflo, E., Dupas, P. & Kremer, M. (2009). Additional resources versus organizational changes in education: Experimental evidence from Kenya. Abdul Latif Jameel Poverty Action Lab (JPAL). Working Paper. Massachusetts Institute of Technology (MIT).

Duflo, E., Dupas, P. & Kremer, M. (2014). School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools, *Journal of Public Economics*  123 (2015) 92–110

Duflo, E., Hanna, R. & Ryan, S. (2008). Incentives Work: Getting Teachers to Come to School, *American Economic Review*, 102 (4): 1,241-1,278.

Glewwe, P., Ilias, N. & Kremer, M. (2010). Teacher Incentives. *American Economic Journal: Applied Economics* 2 (July): 205-227.

*Goyal, S. & Pandey, P. (2011). Contract Teachers in India. *Education Economics* 21 (3): 464-484.

*Habib, M. (2010). *The impact of 2002 national teacher contract policy reform on teacher absenteeism in Lahore, Pakistan*. Unpublished PhD thesis, George Washington University.

Lassibille, G., Tan, J., Jesse, C. & Van Nguyen, T. (2010). Managing for Results in Primary Education in Madagascar: Evaluating the Impact of Selected Workflow Interventions. *World Bank Economic Review*, Advance Access, August 2010.

*McEwan, P.J. & Santibanez, L. (2005). Teacher and Principal Incentives in Mexico in *Incentives to improve teaching: Lessons from Latin America*, Vegas, E (ed). The World Bank. Washington D.C.

*Mizala, A. & Romaguera, P. (2005). Teachers' Salary Structure and Incentives in Chile in *Incentives to Improve Teaching: Lessons from Latin America*, Vegas, E (ed). The World Bank. Washington D.C.

Muralidharan, K. & Sundaraman, V. (2010). The Impact of Diagnostic Feedback to Teachers on Student Learning: Experimental Evidence from India. *The Economic Journal*, August. Volume 120, Issue 546, pages F187–F203,

Muralidharan, K. & Sundaraman, V. (2011). 'Long-Term Effects of Teacher Performance Pay: Experimental Evidence from India', Working Paper.

Muralidharan, K. & Sundaraman, V. (2011). Teacher Performance Pay: Experimental Evidence from India, *The Journal of Political Economy*, 119 (1): 39-77.

* Muralidharan, K. & Sundaraman, V. (2013). Contract Teachers: Experimental Evidence from India, National Bureau of Economic Research Working Paper Series, NBER Working Paper No.19440.

*Piper, B. (2010). *Student-centred pedagogy's causal mechanisms: An explanatory mixed-methods analysis of the impact of in-service teacher professional development in Ethiopia*. Unpublished PhD thesis, Harvard Graduate School of Education.

*Pugatch, T. & Schroeder, E. (2014). Incentives for teacher relocation: Evidence from the Gambian hardship allowance, *Economics of Education Review*, 41: 120-136.

*Rau, T. & Contreras, D. (2012). Tournament Incentives for teachers: Evidence from a scaled-up intervention in Chile, *Economic Development and Cultural Change*, 61 (1): 219-246.

Tan, J., Lane, J. & Lassibille, G. (1999). Student Outcomes in Philippine Elementary Schools: An Evaluation of Four Experiments, *World Bank Economic Review*, 13 (3): 493-508.

## 6.2 OTHER REFERENCES USED IN THE TEXT OF THE TECHNICAL REPORT

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

Aslam, M. & Kingdon, G. (2011). What can teachers do to raise student achievement?, *Economics of Education Review*, 30 (3): 559-574.

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

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. http://www.reform.uk/wp-content/uploads/2014/10/The_good_news_from_Pakistan_final.pdf

Béteille, T. (2009). Absenteeism, transfers and patronage: The political economy of teacherlabor markets in India. Unpublished PhD. Dissertation, Stanford University.

Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A., & Sandefur, J. (2013). *Scaling Up What Works: Experimental Evidence on External Validity in Kenyan Education*. January. Centre for Global Development. Washington D.C.

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

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

Coburn, C. (2003). Rethinking Scale: Moving Beyond Numbers to Deep and Lasting Change. *Educational Researcher* 32 (6): 3-12.

DFID (2013). *Improving Learning, Expanding Opportunities*. Department for International Development Education Position Paper, July.

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

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

Glewwe, P., Hanushek, E., Humpage, S. & Ravina, R. (2013). School Resources and Educational Outcomes in Developing Countries: A Review of the Evidence from 1990 to 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. & Woessmann, L. (2011). Overview of the symposium for performance pay for teachers, *Economics of Education Review* 30: 391-393.

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. Reforms to Increase Teacher Effectiveness in Developing Countries 17.

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

Kunje, D. (2002). The Malawi integrated in-service teacher education programme: An experiment with mixed-mode training. *International Journal of Educational Development* 22: 305-320.

Lewin, K. (2002). The costs of supply and demand for teacher education: dilemmas for development. *International Journal of Educational Development* 22 (2002): 221–242.

Stacki, S. (2002). *Women Teachers in India: Teacher Training through a Gender Lens*. UNICEF. New York, NY, USA.

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.

UNICEF (2012). *Global Initiative on Out of School Children: Ghana Country Study*. UNESCO. Paris: Institute of Statistics.

UWEZO http://www.uwezo.net/ accessed 14 March 2014.

**Table A1.1 : PICOST inclusion/exclusion criteria for defining studies (quantitative and qualitative)**

|  | Included | Excluded |
|---|---|---|
| Population | Lesser-developed countries<br><br>Middle-income countries<br><br>Primary- and secondary-school children in government schools<br><br>General schooling | High-income countries<br><br>Transition economies<br><br>Upper-middle-income countries<br><br>Tertiary schooling<br><br>Private- or aided-school children<br><br>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 | RQ1 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.<br><br>RQ2: If there is no comparison group, studies may be relevant for RQ2. | |
| Outcome | Academic achievement tests (learning outcomes)<br><br>Teacher quality (time on task, teacher motivation, competence, absence, skills, effort, qualifications, credentials, teacher test scores, etc.) | Self-reported happiness, measures of well-being<br><br>Non-cognitive scores<br><br>School enrolment, attendance, completion, transition |

Concepts and Search Terms

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

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),  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) programme(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) programme(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 e.g. for Bangladesh and Bangladeshi]

The search strings and strategy used to construct them within each database are detailed below.

### Table A2.2.1 Search Strings

| EBSCO Host | |
|---|---|
| Database | Search Strategy |
| ERIC | Concept 1 and concept 5 searches are run using the strings below, with date limitation for 1990 to 2014, and limiting results to "primary education" and "secondary education". They are then combined using "AND". This yields >1,000 hits.<br><br>Concept 1<br><br>("academic" OR "child*" OR "classroom" OR "education" OR "learning" OR "pupil*" OR "school*" OR "student*" OR "teach*") N3 ("reform*" OR "intervention*" OR "incentive*" OR "programme*" OR "scheme*" OR "initiative*")<br><br>Concept 5<br><br>"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 "Lesoth*" 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*"<br><br>ERIC also contains a full thesaurus of subject terms (a separate field "keywords KW" of author- |

supplied terms is disregarded). After a manual search of the thesaurus for relevant terms, the following string is made and included to expand concept 1.

DE ("Academic Achievement" OR "Achievement Gains" OR "Educational Attainment" OR "Educational Attainment" OR "School-Based Management" OR "Summative Evaluation" OR "Questioning Techniques" OR "Reflective Teaching" OR "Time on Task" OR "Academic Achievement" OR "Access to Education" OR "Teacher Education" OR "Inservice Teacher Education" OR "Resource Centres" OR "Inservice Teacher Education" OR "Professional Continuing Education" OR "Merit Pay" OR "Computer Uses in Education" OR "Technology Uses in Education" OR "Virtual Classrooms" OR "Charter Schools" OR "Outcome Measures" OR "Competency Based Teacher Education" OR "Performance Based Assessment" OR "Time on Task" OR "Active Learning" OR "Educational Methods" OR "Class Activities" OR "Class Organization" OR "Classroom Communication" OR "Classroom Environment" OR "Incentives" OR "Lesson Plans" OR "Literacy" OR "Motivation" OR "Teacher Motivation" OR "Numeracy" OR "Community Education" OR "Educational Administration" OR "Educational Technology" OR "Skill Development" OR "Skill Centres" OR "Interpersonal Communication" OR "Positive Reinforcement" OR "Problem Solving" OR "Teacher Characteristics" OR "Teacher Effectiveness" OR "School Effectiveness" OR "Instructional Effectiveness" OR "Outcomes of Education")

This yields >800 hits.

Separate strings for concepts 2, 3, 4 are then run within the title/abstract/subject fields as follows.

Concept 2

teach* AND (improve* OR method* OR quality OR resource* OR skill OR style OR strateg* OR practice* OR effective* OR observ* OR absent* OR attend* OR accountab* OR competen* OR knowledge*)

Concept 3

(academic OR child* OR classroom OR cognitive OR education OR grade OR learning OR pupil* OR schola* OR student*) N3 (achievement* OR attainment* OR assessment* OR attendance* OR evaluation* OR enrolment* OR performance* OR progress OR skill* OR test* OR test-score* OR mark* OR result* OR retention OR outcome*)

| | |
|---|---|
| | Concept 4

(block* OR resist* OR enabl* OR oppos* OR advanc* OR partner* OR implement* OR advance* OR authoriz* OR assist* OR encourag* OR reinforce* OR support* OR sanction*) N3 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)

The final search string is then run using the following structure (C1 AND C5) AND (C2 OR C3 OR C4). This yields 700 hits.

Publication types in each string are further restricted to "Reports — ALL", "Doctoral Dissertations", "Journal Articles", and "Books".

The final number of hits is 651. |
| ECONLIT | The database "Econlit with Full Text" is used to search for literature from the period 2000-14. Concept 1 and concept 5 searches are run in title, subject and abstract field, with date limitations applied as below:

Concept 1

("academic" OR "child*" OR "classroom" OR "education" OR "learning" OR "pupil*" OR "school*" OR "student*" OR "teach*") N3 ("reform*" OR "intervention*" OR "incentive*" OR "program*" OR "scheme*" OR "initiative*")

Concept 5

"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 "Lesoth*" 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*"

The following three strings are then run individually in title, abstract and subject fields, with date limitations applied:

Concept 2

teach* AND (improve* OR method* OR quality OR resource* OR skill OR style OR strateg* OR practice* OR effective* OR observ* OR absent* OR attend* OR accountab* OR competen* OR knowledge*)

Concept 3

(academic OR child* OR classroom OR cognitive OR education OR grade OR learning OR pupil* OR schola* OR student*) N3 (achievement* OR attainment* OR assessment* OR attendance* OR evaluation* OR enrolment* OR performance* OR progress OR skill* OR test* OR test score* OR mark* OR result* OR retention OR outcome*)

Note: N3 was used instead of AND to cut hits down to a manageable number.

Concept 4

(block* OR resist* OR enabl* OR oppos* OR advanc* OR partner* OR implement* OR advance* OR authoriz* OR assist* OR encourag* OR reinforce* OR support* OR sanction*) N3 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)

All five strings are combined using (C1 AND C5) AND (C2 OR C3 OR C4). This yields 327 hits. The database does not contain a thesaurus.

For literature in the period 1990-99, the ECONLIT (not full text) database was searched using the

| | |
|---|---|
| | concept 1 search string only. This yielded an additional 25 hits. |
| | Total hits from this database: 352 |
| ERC | For this database, we only had access to literature from 2000 onwards. The strings for concept 1 and 5 are run in title, subject and abstract fields. Several restrictions are applied: date restrictions, English-language papers only, and articles/proceedings/case studies only. Next, strings are run for concepts 2, 3 and 4 in title, subject and abstract fields. The same restrictions are applied. These strings are identical to those used in ECONLIT (above). All five search strings are additionally appended using the following to thesaurus string: |
| | AND (DE "ELEMENTARY education*" or "PRIMARY education*") |
| | Finally, all search strings are combined using (C1 AND C5) AND (C2 OR C3 OR C4). |
| | This yielded 87 hits |
| | Note: This search was completed in April 2014. It was not possible to revisit this search in September 2014 to include secondary education and literature from 1990 onwards, as the team no longer had access to the database. |
| TRC | The strings for concepts 1 and 5 are run in title, subject and abstract fields, applying date restrictions, and restricting to cover peer-reviewed pieces only. Next, strings are run for concepts 2, 3 and 4 in title, subject and abstract fields. These strings are identical to those used in ERC and ECONLIT (above). All search strings are combined using (C1 AND C5) AND (C2 OR C3 OR C4). |
| | This yields 241 hits. The database does not contain a thesaurus. |
| | |

| | |
|---|---|
| eBook Collection | This database does not contain abstracts or a thesaurus, so title and subject searches are run on concept 1, with appropriate date restrictions (1990-14). Concept 1 searches were only run to keep the search as broad as possible. The string is as below: |
| | Concept 1 |
| | ("academic" OR "child*" OR "classroom" OR "education" OR "learning" OR "pupil*" OR "school*" OR "student*" OR "teach*") N3 ("reform*" OR "intervention*" OR "incentive*" OR "program*" OR "scheme*" OR "initiative*") |
| | This yields 21 hits. |
| British Education Index | Title, abstract and subject searches are run on concepts 1 and 5, using the strings noted below. Dates are restricted to cover only January 1990 to September 2014. Outputs are limited to English-language papers only. Educational level is limited to either "elementary school", "middle school" or "secondary school". |
| | Concept 1 |
| | ("academic" OR "child*" OR "classroom" OR "education" OR "learning" OR "pupil*" OR "school*" OR "student*" OR "teach*") N3 ("reform*" OR "intervention*" OR "incentive*" OR "program*" OR "scheme*" OR "initiative*") |
| | Concept 5 |
| | "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 "Lesoth*" 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\*"<br><br>Strings for concepts 1 and 5 are then combined using "AND". This yields 195 hits.<br><br>This database does not contain a thesaurus. |
| Psycinfo | A manual search of the thesaurus is run to locate relevant descriptor terms. This yields the following string:<br><br>DE ("Primary School Students" OR "High School Education" OR "Classroom Behaviour" OR "Classroom Environment" OR "Classroom Management" OR "Classrooms" OR "Incentives" OR "Literacy" OR "Education" OR "Educational Administration" OR "Educational Attainment Level" OR "Educational Audiovisual Aids" OR "Educational Incentives" OR "Educational Measurement" OR "Educational Personnel" OR "Educational Programs" OR "Educational Quality" OR "Educational Reform" OR "Educational Standards" OR "School Administrators" OR "School Attendance" OR "School Based Intervention" OR "School Enrolment" OR "School Environment" OR "School Learning" OR "Student Engagement" OR "Teacher Characteristics" OR "Teacher Education" OR "Teacher Effectiveness" OR "Teacher Effectiveness Evaluation" OR "Teacher Recruitment" OR "Teacher Stunt Interaction" OR "Teachers" OR "Teaching Methods")<br><br>This is combined with concepts 1 and 5 as (1 or DE) AND 5. Results are restricted to 1990 onwards and English language only. As there are a large number of irrelevant hits, strings are run with the N1, rather than N3 proximity term, as below. This search yields 2,543 hits.<br><br>Concept 1<br><br>("academic" OR "child\*" OR "classroom" OR "education" OR "learning" OR "pupil\*" OR "school\*" OR "student\*" OR "teach\*") N1 ("reform\*" OR "intervention\*" OR "incentive\*" OR "program\*" OR "scheme\*" OR "initiative\*")<br><br>Concept 5 |

"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 "Lesoth*" 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*"

To cut hits down further, concepts 2, 3 and 4 are also run as separate strings, with date restrictions applied, N1 proximity term, and results limited to English language only. These strings are as below:

Concept 2

teach* N1 (improve* OR method* OR quality OR resource* OR skill OR style OR strateg* OR practice* OR effective* OR observ* OR absent* OR attend* OR accountab* OR competen* OR knowledge*)

Concept 3

(academic OR child* OR classroom OR cognitive OR education OR grade OR learning OR pupil* OR schola* OR student*) N1 (achievement* OR attainment* OR assessment* OR attendance* OR evaluation* OR enrolment* OR performance* OR progress OR skill* OR test* OR test-score* OR mark* OR result* OR retention OR outcome*)

Concept 4

(block* OR resist* OR enabl* OR oppos* OR advanc* OR partner* OR implement* OR advance* OR authoriz* OR assist* OR encourag* OR reinforce* OR support* OR sanction*) N1 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)

The final string is then run as:

((1 OR DE) AND 5) AND (2 OR 3 OR 4)

This yields 1,101 hits. These are further restricted by removing papers with the subjects "AIDS", "AIDS prevention", "health education", "health knowledge", "HIV", "American Indians", "special education", "immigration", "higher education", "human sex differences" and "colleges". Papers whose methodology is "non-clinical case study", "mathematical model", "literature review", "meta analysis" and "scientific simulation" are also removed. Hits from the following classifications are also removed.

Social processes & social issues:

 educational/vocational counseling & student services

Promotion & maintenance of health & wellness:

 special & remedial education

 health & mental health treatment & prevention

 developmental psychology

behavior disorders & anti-social behavior

culture & ethnology

psychosocial & personality development

social psychology

community & social services

drug & alcohol rehabilitation

| | immunological disorders |
|---|---|
| | industrial & organizational psychology |
| | social structure & organization |
| | substance abuse & addiction |
| | human factors engineering |
| | linguistics & language & speech |
| | vision & hearing & sensory disorders |
| | child-rearing & child care |
| | criminal behavior & juvenile delinquency |
| | developmental disorders & autism |
| | group & family therapy |
| | health psychology & medicine |
| | medical treatment of physical illness |
| | mental retardation |
| | neurological disorders & brain damage |
| | personality traits & processes |
| | professional psychological & health personnel issues |
| | psychological disorders |
| | The final search yields 642 hits. |

**Proquest**

| Database | Search Strategy |
|---|---|
| ASSIA | Relevant terms are culled from the thesaurus, and added to the C1 search. Initial searches for C1 and C5 yield over 16,000 papers. Date restrictions are applied, and the thesaurus is used to limit results only to those papers that examine primary or secondary schools. |

| | |
|---|---|
| | The final string yields 89 hits:<br><br>(TI,SU,AB((academic OR child* OR classroom OR education OR learning OR pupil* OR school* OR student* OR teach*) W/3 (reform* OR intervention* OR incentive* OR programme* OR scheme* OR initiative*)) OR SU.EXACT("Classrooms" OR "In-service training" OR "Wage incentives" OR "Information and communication technologies" OR "Literacy" OR "Numeracy" OR "Pedagogy" OR "Problem solving" OR "School environment" OR "School-community relationship" OR "Teacher recruitment" OR "Teacher training" OR "Student participation" OR "Teacher-student relationship" OR "Teachers")) AND YR(>=1990) AND SU.EXACT("Primary schools" OR "Elementary education" OR "high schools" OR "continuation high schools" OR "junior high schools" OR "secondary schools" OR "secondary education" OR "junior secondary schools") AND (TI,AB,SU(Afghan* OR Armen* OR Bangladesh* OR Benin* OR Bhutan* OR "Burkina Faso*" OR Burundi* 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 Ghana* OR Guinea* OR Guatemala* OR Haiti* OR Honduras* OR Guyana* OR India* OR Indonesia* OR Kenya* OR Kiribati* OR Kyrgyz* OR Lao* OR Kosovo* OR Lesotho* OR Liberia* OR Madagascar* OR Malawi* OR Mali* OR "Marshall Islands" OR Mauritania* OR Micronesia* OR Moldova* OR Mongolia* OR Mozambi* OR Moro* OR Nepal* OR Nicaragua* OR Niger* OR Myanmar OR Pakistan* OR "Papua New Guinea*" OR Paraguay* OR Philippine* 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 "East Timor" OR Togo* OR Tonga* OR Ukrain* OR Palestine* OR "West Bank" OR Gaza* OR Turkmenistan* OR Tuvalu* OR Uganda* OR Uzbek* OR Vanuatu* OR Vietnam* OR Yemen* OR Zambia* OR Zimbabwe*)) |
| IBSS | This database utilises a thesaurus field, using the field code SU. The concept 1 search is broadened with relevant terms from the thesaurus. Results are limited by year, and to papers examining primary and secondary schools only.<br><br>(TI,SU,AB((academic OR child* OR classroom OR education OR learning OR pupil* OR school* OR student* OR teach*) W/3 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)) OR SU.EXACT("Classrooms" OR "In-service training" OR "Wage incentives" OR "Information and communication technologies" OR "Literacy" OR "Numeracy" OR "Pedagogy" OR "Problem solving" OR "School environment" OR "School-community relationship" OR "Teacher recruitment" OR "Teacher training" OR "Student participation" OR "Teacher-student relationship" OR "Teachers")) AND YR(>=1990) AND SU.EXACT("Primary schools" OR "Elementary education" OR |

"Secondary education" OR "secondary school*" OR "high school*")


A combined search with concepts 1 and 5 yields 237 hits. This has been cut by combining with remaining concepts. This final search string is:


(((TI,SU,AB((academic OR child* OR classroom OR education OR learning OR pupil* OR school* OR student* OR teach*) W/3 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)) OR SU.EXACT("Classrooms" OR "In-service training" OR "Wage incentives" OR "Information and communication technologies" OR "Literacy" OR "Numeracy" OR "Pedagogy" OR "Problem solving" OR "School environment" OR "School-community relationship" OR "Teacher recruitment" OR "Teacher training" OR "Student participation" OR "Teacher-student relationship" OR "Teachers")) AND YR(>=1990) AND SU.EXACT("Primary schools" OR "Elementary education" OR "Secondary education" OR "secondary school*" OR "high school*")) AND (TI,AB,SU(Afghana* OR carmen* OR Bangladesh* OR Benin* OR Bhutan* OR "Burkina Faso*" OR Burundi* 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 Ghana* OR Guinean* OR Guatemala* OR Haiti* OR honduras* OR Guyana* OR India* OR Indonesia* OR Kenya* OR Kiribati* OR Kyrgyz* OR Lao* OR kosovo* OR Lesotho* OR Liberia* OR Madagascar* OR Malawi* OR Mali* OR "Marshall Islands" OR Mauritania* OR Micronesia* OR Moldova* OR Mongolia* OR Mozambi* OR Moro* OR Nepal* OR Nicaragua* OR Niger* OR Myanmar OR Pakistan* OR "Papua New Guinean*" OR Paraguay* OR Philippine* 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 Palestine* OR "West Bank" OR Gaza* OR Turkmenistan* OR Tuvalu* OR Uganda* OR Uzbek* OR Vanuatu* OR Vietnam* OR Yemen* OR Zambia* OR Zimbabwe*) AND YR(>=1990))) AND ((TI,AB,SU(teach* W/3 (improve* OR method* OR quality OR resource* OR skill OR style OR strategy* OR practice* OR effective* OR observe* OR absent* OR attend* OR accountable* OR competent* OR knowledge*)) AND YR(>=1990)) OR (TI,AB (academic OR child* OR classroom OR cognitive OR education OR grade OR learning OR pupil* OR scholar* OR student*) W/3 (achievement* OR attainment* OR assessment* OR attendance* OR evaluation* OR enrolment* OR performance* OR progress OR skill* OR test* OR test-score* OR mark* OR result* OR retention OR outcome*) AND YR(>=1990)) OR (SU (academic OR child* OR classroom OR cognitive OR education OR grade OR learning OR pupil* OR scholar* OR student*) W/3 (achievement* OR attainment* OR assessment* OR attendance* OR evaluation* OR enrolment* OR performance* OR progress OR skill* OR test* OR test-score* OR mark* OR result* OR retention OR outcome*) AND YR(>=1990)) OR (TI,AB,SU((block* OR resist* OR enable* OR oppose* OR advanc* OR partner* OR implement* OR advance* OR authorize* OR assist* OR encourage* OR reinforce* OR support* OR sanction*) W/3 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)) AND YR(>=1990)))

| | |
|---|---|
| | This yields 69 hits. |
| Proquest Dissertations: GLobal | A combined search of concepts 1 and 5 yielded 3,139 results. This search was then combined with C2, C3, and C4. The final search string is:<br><br>((TI,SU,AB((academic OR child* OR classroom OR education OR learning OR pupil* OR school* OR student* OR teach*) W/1 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)) AND YR(>=1990)) AND (TI,AB,SU(Afghana* OR carmen* OR Bangladesh* OR Benin* OR Bhutan* OR "Burkina Faso*" OR burundi* 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 ghana* OR guinean* OR guatemala* OR Haiti* OR honduras* OR guyana* OR India* OR Indonesia* OR Kenya* OR Kiribati* OR Kyrgyz* OR Lao* OR kosovo* OR lesotho* OR Liberia* OR madagascar* OR Malawi* OR Mali* OR "Marshall Islands" OR Mauritania* OR Micronesia* OR Moldova* OR mongolia* OR Mozambi* OR moro* OR Nepal* OR Nicaragua* OR Niger* OR Myanmar OR Pakistan* OR "Papua New guinean*" OR Paraguay* OR philippine* 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 palestine* OR "West Bank" OR Gaza* OR Turkmenistan* OR Tuvalu* OR Uganda* OR Uzbek* OR Vanuatu* OR Vietnam* OR Yemen* OR Zambia* OR Zimbabwe*) AND YR(>=1990))) AND (((TI,AB ((academic OR child* OR classroom OR cognitive OR education OR grade OR learning OR pupil* OR scholar* OR student*) W/1 (achievement* OR attainment* OR assessment* OR attendance* OR evaluation* OR enrolment* OR performance* OR progress OR skill* OR test* OR test-score* OR mark* OR result* OR retention OR outcome*)) AND YR(>=1990)) AND (SU ((academic OR child* OR classroom OR cognitive OR education OR grade OR learning OR pupil* OR scholar* OR student*) W/1 (achievement* OR attainment* OR assessment* OR attendance* OR evaluation* OR enrolment* OR performance* OR progress OR skill* OR test* OR test-score* OR mark* OR result* OR retention OR outcome*)) AND YR(>=1990))) OR (TI,AB,SU((teach*) W/1 (improve* OR method* OR quality OR resource* OR skill OR style OR strategy* OR practice* OR effective* OR observe* OR absent* OR attend* OR accountable* OR competent* OR knowledge*)) AND YR(>=1990)) OR (TI,AB,SU((block* OR resist* OR enable* OR oppose* OR advance* OR partner* OR implement* OR advanc* OR authorize* OR assist* OR encourage* OR reinforce* OR support* OR sanction*) W/1 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)) AND YR(>=1990)))<br><br>This yielded 486 hits. The results from this string are further cut by using the menu of options to exclude papers with non-relevant subjects and index terms. Expressed in string form, the restrictions are: |

| | |
|---|---|
| | Subject: (teacher education AND secondary education AND primary education AND school administration)

Note: For concept 3 Title and abstract searches have been combined with subject searches by using the operator "AND" instead of "OR", as this yielded a manageable number of hits overall.

The final number of hits from this database = 266 |
| Australian Education Index | Thesaurus terms are referred to as "subject heading, all" in Proquest. In addition to the standard search, a manual search of the thesaurus is run to locate descriptor terms. A hand search of the thesaurus yields 32 relevant subject terms, which are searched for using the following string (yielding over 66,000 hits):

SU.EXACT.EXPLODE ("Cognitive skills" OR "Literacy" OR "Teacher competencies" OR "Performance based assessment" OR "Praise" OR "Summative evaluation" OR "Educational resources" OR "Teacher effectiveness" OR "Motivation" OR "Student motivation" OR "Incentives" OR "Numeracy" OR "Classrooms" OR "Problem solving" OR "Time on task" OR "Teacher motivation" OR "Lesson plans" OR "Academic achievement" OR "Performance contracts" OR "School effectiveness" OR "Educational facilities improvement" OR "Teaching process" OR "Information and communications technology" OR "Classroom techniques" OR "Motivation techniques" OR "Communication skills" OR "Classroom environment" OR "School organisation" OR "Professional development" OR "Accountability" OR "Active learning" OR "Teaching skills")

This string is used to expand concept 1 (using OR), and then with combined with concept 5, yielding 2,143 hits:

(TI,SU,AB((academic OR child* OR classroom OR education OR learning OR pupil* OR school* OR student* OR teach*) W/3 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)) OR SU.EXACT.EXPLODE("Cognitive skills" OR "Literacy" OR "Teacher competencies" OR "Performance based assessment" OR "Praise" OR "Summative evaluation" OR "Educational resources" OR "Teacher effectiveness" OR "Motivation" OR "Student motivation" OR "Incentives" OR "Numeracy" OR "Classrooms" OR "Problem solving" OR "Time on task" OR "Teacher motivation" OR "Lesson plans" OR "Academic achievement" OR "Performance contracts" OR |

"School effectiveness" OR "Educational facilities improvement" OR "Teaching process" OR "Information and communications technology" OR "Classroom techniques" OR "Motivation techniques" OR "Communication skills" OR "Classroom environment" OR "School organisation" OR "Professional development" OR "Accountability" OR "Active learning" OR "Teaching skills")) AND TI,AB,SU(Afghana* OR carmen* OR Bangladesh* OR Benin* OR Bhutan* OR "Burkina Faso*" OR burundi* 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 ghana* OR guinean* OR guatemala* OR Haiti* OR honduras* OR guyana* OR India* OR Indonesia* OR Kenya* OR Kiribati* OR Kyrgyz* OR Lao* OR kosovo* OR lesotho* OR Liberia* OR madagascar* OR Malawi* OR Mali* OR "Marshall Islands" OR Mauritania* OR Micronesia* OR Moldova* OR mongolia* OR Mozambi* OR moro* OR Nepal* OR Nicaragua* OR Niger* OR Myanmar OR Pakistan* OR "Papua New guinean*" OR Paraguay* OR philippine* 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 palestine* OR "West Bank" OR Gaza* OR Turkmenistan* OR Tuvalu* OR Uganda* OR Uzbek* OR Vanuatu* OR Vietnam* OR Yemen* OR Zambia* OR Zimbabwe*) AND YR(>=1990)

This string is then combined with each of concepts 2, 3 and 4, in turn, over separate searches. These yield 730, 580 and 180 hits, respectively. It is not possible to combine these searches further, as the required string length overflows the capacity of the database. Therefore, each string's results are restricted individually using the menu of options:

Concept 2:

TI,AB,SU(teach* W/3 (improve* OR method* OR quality OR resource* OR skill OR style OR strateg* OR practice* OR effective* OR observ* OR absent* OR attend* OR accountab* OR competen* OR knowledge*)) AND YR(>=1990)

Restricted to remove papers on the subjects "English (second language)", "postsecondary education", "university teaching", "universities", adult education", "international students", "university students", and "higher education"; 375 results remain.

Concept 3:

TI,AB ((academic OR child* OR classroom OR cognitive OR education OR grade OR learning OR pupil* OR schola* OR student*) W/3 (achievement* OR attainment* OR assessment* OR attendance* OR evaluation* OR enrolment* OR performance* OR progress OR skill* OR test* OR

test-score* OR mark* OR result* OR retention OR outcome*)) AND YR(>=1990)

Does not search in subject for concept 3, as the database is unable to interpret this search. Only Title and abstract fields are applied. The resulting 580 hits are restricted to remove papers on the subjects "English (second language)", "postsecondary education", "university students", "universities", adult education", "international students", "university teaching", "graduate students" and "higher education"; 284 hits remain.

Concept 4:

TI,AB,SU((block* OR resist* OR enable* OR oppose* OR advance* OR partner* OR implement* OR advanc* OR authorize* OR assist* OR encourage* OR reinforce* OR support* OR sanction*) W/3 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*)) AND YR(>=1990))

Restricted to remove papers on the subjects "English (second language)", "postsecondary education", "university teaching", "universities", "vocational education and training", "international students", "university students", "university administration" and "higher education"; 101 results remain.

All items in each of the three search strings are selected for export, which automatically removes duplicates across strings; 638 hits result.

| Web of Knowledge |
| --- |

| Web of Knowledge | Initial Search<br><br>Searches for concepts 1 and 5 were run using the topic field for each database. Searches were restricted by language (English) and by document type (Article OR Book Chapter). Here, searches within the databases "Social Sciences Index" and "Conference Proceedings Citation Index – Social Sciences & Humanities" (CPCI-SSH) are run together. The initial search strings are as follows<br><br>Concept 1:<br><br>(TS=((academic OR child* OR classroom OR education OR learning OR pupil* OR school* OR student* OR teach*) NEAR/2 (reform* OR intervention* OR incentive* OR program* OR scheme* |

OR initiative*))) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article OR Book Chapter)

Indexes=SSCI, CPCI-SSH Timespan=1990-2014

Concept 5:

(TS=("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 "Lesoth*" 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*")) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article OR Book Chapter)

Indexes=SSCI, CPCI-SSH Timespan=1990-2014

The initial searches for concept 1 and concept 5 were refined by Web of Science category terms.

(WC=(PUBLIC ADMINISTRATION OR EDUCATION EDUCATIONAL RESEARCH OR EDUCATION SPECIAL OR SOCIOLOGY OR ECONOMICS OR SOCIAL ISSUES OR EDUCATION SCIENTIFIC DISCIPLINES OR PSYCHOLOGY OR SOCIAL SCIENCES INTERDISCIPLINARY OR PSYCHOLOGY APPLIED OR PSYCHOLOGY MULTIDISCIPLINARY OR DEMOGRAPHY OR STATISTICS PROBABILITY OR SOCIAL SCIENCES MATHEMATICAL METHODS OR POLITICAL SCIENCE OR PSYCHOLOGY EDUCATIONAL)) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article OR Book Chapter)

Indexes=SSCI, CPCI-SSH Timespan=1990-2014

This yielded 1,004 hits. Separate searches within the topic field were run for the remaining

concepts as follows:

Concept 2

(TS=(teach* NEAR/2 (improve* OR method* OR quality OR resource* OR skill OR style OR strateg* OR practice* OR effective* OR observ* OR absent* OR attend* OR accountab* OR competen* OR knowledge*))) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article OR Book Chapter)

Indexes=SSCI, CPCI-SSH Timespan=1990-2014


Concept 3

(TS=((academic OR child* OR classroom OR cognitive OR education OR grade OR learning OR pupil* OR schola* OR student*) NEAR/2 (achievement* OR attainment* OR assessment* OR attendance* OR evaluation* OR enrolment* OR performance* OR progress OR skill* OR test* OR test-score* OR mark* OR result* OR retention OR
outcome*))) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article OR Book Chapter)

Indexes=SSCI, CPCI-SSH Timespan=1990-2014


Concept 4

(TS=((block* OR resist* OR enabl* OR oppos* OR advanc* OR partner* OR implement* OR advance* OR authoriz* OR assist* OR encourag* OR reinforce* OR support* OR sanction*) NEAR/2 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative*))) AND LANGUAGE:(English) AND DOCUMENT TYPES: (Article OR Book Chapter)

Indexes=SSCI, CPCI-SSH Timespan=1990-2014


The final search string was generated by combining the concepts using the following structure: (C1 AND C5) AND (C2 OR C3 OR C4).

Note: The proximity search "NEAR/2" was used as this yielded the most manageable number of hits.

This yielded 408 hits.

Other

| Database | Search Strategy |
| --- | --- |
| JSTOR | Since JSTOR is a non-bibliographic database and mainly a journal platform, it is not as well indexed as other databases available via the Proquest and EBSCOhost platforms. The database cannot cope with sophisticated search strategies that involve combining multiple concepts. Search strings have very limited character restrictions, allow for only four wild cards at a time. With all of these limitations, it was found that even a basic search using ALL concept-1 terms was not possible.<br><br>In view of this, the decision was made to hand-search JSTOR, in order to be able to work within its limited functionality. This decision is reasonable, since many of the journals archived within JSTOR are also available via the databases being searched, via Proquest and EBSCOhost platforms. Therefore, an additional hand search of the JSTOR website further ensures that relevant literature is not missed.<br><br>Multiple search strings were run for all the terms from concept 1. Only Item Title searches have been run. Since JSTOR only contains abstracts for 10% of its journal articles, these have not been run. Full-text searches are far too broad, as the terms searched for appear anywhere in the article. This yielded an unmanageable number of hits. Searches were restricted to research in English, to content from within and outside JSTOR, and to Economics, Education, Social Sciences, Population Studies, Development Studies, Sociology, Psychology and Public Policy disciplines. No date restrictions were applied, and screening was limited to work from 1990 onwards.<br><br>Search results were sorted by "relevance" and the first 100 titles were manually screened. If a relevant hit was found, this was manually uploaded to EPPI Reviewer.<br><br>The following search strings were run using the Basic Search Form. Field codes "ti", "la" and "disc" mark out title, language and disciplinary restrictions. Due to space constraints, only the strings using "academic" as the prefix are presented. Similar strings were constructed for the rest of the terms in concept 1 and involve replacing "academic" with "child", "classroom", "education", "learning", "pupil", "student", "school", "teacher" and "teacher training".<br><br>((ti:("academic")) AND ti:("reform&")) AND la:(eng OR en) AND disc:(populationstudies-discipline OR economics-discipline OR education-discipline OR social sciences-discipline)<br><br>((ti:("academic")) AND ti:("reform&")) AND la:(eng OR en) AND disc:(developmentstudies- |

discipline OR psychology-discipline OR sociology-discipline OR publicpolicy-discipline)

((ti:("academic")) AND ti:("program&")) AND la:(eng OR en) AND disc:(populationstudies-discipline OR economics-discipline OR education-discipline OR social sciences-discipline)

((ti:("academic")) AND ti:("program&")) AND la:(eng OR en) AND disc:(developmentstudies-discipline OR psychology-discipline OR sociology-discipline OR publicpolicy-discipline)

((ti:("academic")) AND ti:("scheme&")) AND la:(eng OR en) AND disc:(populationstudies-discipline OR economics-discipline OR education-discipline OR social sciences-discipline)

((ti:("academic")) AND ti:("scheme&")) AND la:(eng OR en) AND disc:(developmentstudies-discipline OR psychology-discipline OR sociology-discipline OR publicpolicy-discipline)

((ti:("academic")) AND ti:("scheme&")) AND la:(eng OR en) AND disc:(populationstudies-discipline OR economics-discipline OR education-discipline OR social sciences-discipline)

((ti:("academic")) AND ti:("intervention&")) AND la:(eng OR en) AND disc:(developmentstudies-discipline OR psychology-discipline OR sociology-discipline OR publicpolicy-discipline)

((ti:("academic")) AND ti:("intervention&")) AND la:(eng OR en) AND disc:(populationstudies-discipline OR economics-discipline OR education-discipline OR social sciences-discipline)

((ti:("academic")) AND ti:("initiative&")) AND la:(eng OR en) AND disc:(developmentstudies-discipline OR psychology-discipline OR sociology-discipline OR publicpolicy-discipline)

((ti:("academic")) AND ti:("initiative&")) AND la:(eng OR en) AND disc:(populationstudies-discipline OR economics-discipline OR education-discipline OR social sciences-discipline)

((ti:("academic")) AND ti:("incentive&")) AND la:(eng OR en) AND disc:(developmentstudies-discipline OR psychology-discipline OR sociology-discipline OR publicpolicy-discipline)

((ti:("academic")) AND ti:("incentive&")) AND la:(eng OR en) AND disc:(populationstudies-discipline OR economics-discipline OR education-discipline OR social sciences-discipline)

| | |
|---|---|
| Science Direct | The following search was inputted into the Expert Search Form. This yielded 380 hits.<br><br>pub-date > 1989 and tak((academic OR child OR classroom OR education OR learning OR pupil OR school OR student OR teach*) PRE/3 (reform* OR intervention* OR incentive* OR program* OR scheme* OR initiative* )) AND tak(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*)[All Sources(Economics, Econometrics and Finance, Psychology, Social Sciences)]. |
| AJOL / AsiaJOL | We ran separate search strings. Within AJOL, search strings were run twice each, once in title, once in keywords. Within Asiajol, Searches were run in the category "all fields". Separate searches were also run in title and subject fields, as these yielded different hits.

All journals are searched in these databases. No language or journal restrictions are applied. Date restrictions are applied in the year field (1990-2014 for AsiaJOL, 2003-14 for AJOL (note: literature is only available from 2003 onwards for AJOL).

No RIS files for outputting are supported. Search results were sorted by relevance, and, in cases where >101 hits were returned, the first 100 hits were manually screened and uploaded to EPPI Reviewer.

"academic reform*" OR "academic program*" OR "academic scheme*" OR "academic intervention*" OR "academic incentive*" OR "academic initiative*"

"child reform*" OR "child program*" OR "child scheme*" OR "child intervention*" OR "child incentive*" OR "child initiative*"

"classroom reform*" OR "classroom program*" OR "classroom scheme*" OR "classroom intervention*" OR "classroom incentive*" OR "classroom initiative*"

"education reform*" OR "education program*" OR "education scheme*" OR "education intervention*" OR "education incentive*" OR "education initiative*"

"learn* reform*" OR "learn* program*" OR "learn* scheme*" OR "learn* intervention*" OR "learn* incentive*" OR "learn* initiative*"

"pupil reform*" OR "pupil program*" OR "pupil scheme*" OR "pupil intervention*" OR "pupil |

| | |
|---|---|
| | incentive*" OR "pupil initiative*"<br><br>"school reform*" OR "school program*" OR "school scheme*" OR "school intervention*" OR "school incentive*" OR "school initiative*"<br><br>"student reform*" OR "student program*" OR "student scheme*" OR "student intervention*" OR "student incentive*" OR "student initiative*"<br><br>"teach* reform*" OR "teach* program*" OR "teach* scheme*" OR "teach* intervention*" OR "teach* incentive*" OR "teach* initiative*"<br><br>"teacher training reform*" OR "teacher training program*" OR "teacher training scheme*" OR "teacher training intervention*" OR "teacher training incentive*" OR "teacher training initiative*"<br><br><br>Total number of hits from these databases = 5 |
| LAMJOL | Searches were run within the title field.<br><br>When search strings were constructed using only concept 1, there were no hits — refer to BATCH 1 .<br><br>Concept 1 and Concept 5 were then used to construct search strings and these yielded some hits — refer to BATCH 2.<br><br>Results from BATCH 2 were manually screened and uploaded to EPPI reviewer.<br><br><br>BATCH 1<br><br>academic reform*" OR "academic program*" OR "academic scheme*" OR "academic intervention*" OR "academic incentive*" OR "academic initiative*"<br><br>"child reform*" OR "child program*" OR "child scheme*" OR "child intervention*" OR "child incentive*" OR "child initiative*"<br><br>"classroom reform*" OR "classroom program*" OR "classroom scheme*" OR "classroom intervention*" OR "classroom incentive*" OR "classroom initiative*"<br><br>"education reform*" OR "education program*" OR "education scheme*" OR "education intervention*" OR "education incentive*" OR "education initiative*"<br><br>"learn* reform*" OR "learn* program*" OR "learn* scheme*" OR "learn* intervention*" OR |

"learn* incentive*" OR "learn* initiative*"

"pupil reform*" OR "pupil program*" OR "pupil scheme*" OR "pupil intervention*" OR "pupil incentive*" OR "pupil initiative*"

"school reform*" OR "school program*" OR "school scheme*" OR "school intervention*" OR "school incentive*" OR "school initiative*"

"student reform*" OR "student program*" OR "student scheme*" OR "student intervention*" OR "student incentive*" OR "student initiative*"

"teach* reform*" OR "teach* program*" OR "teach* scheme*" OR "teach* intervention*" OR "teach* incentive*" OR "teach* initiative*"

"teacher training reform*" OR "teacher training program*" OR "teacher training scheme*" OR "teacher training intervention*" OR "teacher training incentive*" OR "teacher training initiative*"

No hits


BATCH 2

(academic* OR child* OR classroom* OR educ* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*)(Afghan* OR Armen* OR Bangladesh* OR Benin* OR Bhutan*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*)(Burkina Faso* OR Burund* OR Bolivia* OR Cambodia*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*)(Cameroon* OR Verde* OR Central African Republic)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*)(Chad* OR Comoros OR Congo* OR Côte d'Ivoire)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*)(Ivory Coast OR Djibouti* OR Eritrea* OR Ethiopia*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*)(Egypt* OR El Salvador* OR Georgia* OR Gambia*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*)(Ghan* OR Guinea* OR Guatemal* OR Haiti* OR Hondu*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Guyan* OR India* OR Indonesia* OR Kenya* OR Kiri*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Kyrgyz* OR Lao* OR Kosov* OR Lesotho*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Madagasca* OR Malawi* OR Mali* OR Marshall Islan*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Mauritania* OR Micronesia* OR Moldova* OR Mongo*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Mozambi* OR Moroc* OR Nepal* OR Nicaragua*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Niger* OR Myanmar OR Pakistan* OR Papua New Guin*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Paraguay* OR Philippin* OR Rwanda* OR Samoa*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Sao Tomé and Principe OR Senegal* OR Sierra Leon*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Solomon Islands OR Somalia* OR Sudan*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Swazi* OR Syria* OR "Sri Lank*" OR Tajik*)

(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative*

| | |
|---|---|
| | OR program* OR scheme*) (Tanzania* OR Timor-Leste OR Togo*)<br><br>(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Tonga* OR Ukrain* OR Palestin* OR West Bank)<br><br>(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Gaza OR Turkmenistan* OR Tuvalu* OR Uganda*)<br><br>(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Uzbek* OR Vanuatu* OR Vietnam*)<br><br>(academic* OR child* OR classroom* OR education* OR learning* OR pupil* OR school* OR student* OR teach* OR teacher training)(reform* OR intervention* OR incentive* OR initiative* OR program* OR scheme*) (Yemen* OR Zambia* OR Zimbabwe*)<br><br><br><br>Total number of hits from Batch 2 (and from this database) = 3 |
| SSRN | Each term within concept 1 is searched for individually as exact-phrase searches within quotation marks within the tiles + abstract +keywords field. These are not combined with other concepts due to the limited capacity of the database to interpret sophisticated search strings. Additionally, using terms from concept 1 only kept the search as broad as possible.<br><br><br><br>Since RIS files were not supported to export citations, hits were manually screened on the website and relevant titles were uploaded to EPPI Reviewer. In cases where hits were >101, the first 100 titles were screened only.<br><br><br><br>The following terms in concept 1 yielded 3 relevant hits: "education reform" and "teacher incentives". |
| NBER | Using truncated searches, strings are run separately, as below, for concept 1 only (to keep the search as broad as possible). Search strings are run within the Full-Text Publications search field. Where possible, abstracts were screened on the website before importing to EPPI Reviewer.<br><br>"academic reform*" OR "academic intervention*" OR "academic incentive*" OR "academic initiative*" OR "academic program*" OR "academic scheme*" |

"child reform*" OR "child intervention*" OR "child incentive*" OR "child initiative*" OR "child program*" OR "child scheme*"

"classroom reform*" OR "classroom intervention*" OR "classroom incentive*" OR "classroom initiative*" OR "classroom program*" OR "classroom scheme*"

"education reform*" OR "education intervention*" OR "education incentive*" OR "education initiative*" OR "education program*" OR "education scheme*"

"learning reform*" OR "learning intervention*" OR "learning incentive*" OR "learning initiative*" OR "learning program*" OR "learning scheme*"

"pupil reform*" OR "pupil intervention*" OR "pupil incentive*" OR "pupil initiative*" OR "pupil program*" OR "pupil scheme*"

"school reform*" OR "school intervention*" OR "school incentive*" OR "school initiative*" OR "school program*" OR "school scheme*"

"student reform*" OR "student intervention*" OR "student incentive*" OR "student initiative*" OR "student program*" OR "student scheme*"

"teaching reform*" OR "teaching intervention*" OR teaching incentive*" OR "teaching initiative*" OR "teaching program*" OR "teaching scheme*"

"teacher reform*" OR "teacher intervention*" OR "teacher incentive*" OR "teacher initiative*" OR "teacher program*" OR "teacher scheme*"

"teacher training reform*" OR "teacher training intervention*" OR "teacher training incentive*" OR "teacher training initiative*" OR "teacher training program*" OR "teacher training scheme*"

"teacher pay reform*" OR "teacher pay intervention*" OR "teacher pay incentive*" OR "teacher pay initiative*" OR "teacher pay program*" OR "teacher pay scheme*"


Number of hits from this database = 19

| | |
|---|---|
| Econ papers | We inserted search strings within the keywords and title fields and set to "search for phrase or word forms". A search of concept 1 alone yielded an unmanageable number of hits. Search strings have been constructed such that they combine concepts. For each set of terms within the parentheses, separate strings are run. These are then combined using the "combine" function at the bottom of the page. We inputted any quotation marks manually, as copy and paste will not replicate them. |

#CONCEPT 1, CONCEPT 2 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, CONCEPT 3 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, CONCEPT 4 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 = 274.

| | |
|---|---|
| World Bank | Separate searches have been run for each individual term within concept 1. Search terms are automatically enclosed within quotation marks. Title searches are run separately in the following sections of the WB databases.<br><br>Policy Research Working Papers<br><br>Policy Research Reports<br><br>World Bank Economic Review<br><br>Development Impact Evaluation (DIME) |

| | |
|---|---|
| | Titles (and, where possible, abstracts) from search hits had to be screened manually on the website and relevant hits were uploaded to EPPI Reviewer.

Note: Some search terms have been excluded, as they yielded too many hits to be processed manually (For example, child interventions, incentives, etc.).

The following terms within concept 1 yielded hits that were relevant and uploaded to EPPI Reviewer:

academic reforms

academic interventions

classroom incentives

pupil/student programs

school interventions

teacher programs

teaching interventions

Total number of hits = 14 |
| IMF | Searches were run using concept 1 only. Due to the limited capacity of this database, it was not possible to combine concept 1 with other concepts and develop more sophisticated search strings. The advantage of using concept 1 is also that the search is kept as broad as possible.

The following search strings were run simultaneously in the title & subject/keyword field, along with a date restriction. These strings yielded 0 hits.

Note: Due to space constraints, only the search string using the term "academic" has been presented. Similar strings were constructed for the rest of the terms in concept 1 and involve replacing "academic" with "child", "classroom", "education", "learning", "pupil", "student", "school", "teacher" and "teacher training". |

| | |
|---|---|
| | Title: "academic reform" OR "academic reforms" OR "academic program" OR "academic programs" OR "academic scheme" OR "academic schemes" OR "academic intervention" OR "academic interventions" OR "academic incentive" OR "academic incentives" OR "academic initiative" OR "academic initiatives". Subject Keyword: "academic reform" OR "academic reforms" OR "academic program" OR "academic programs" OR "academic scheme" OR "academic schemes" OR "academic intervention" OR "academic interventions" OR "academic incentive" OR "academic incentives" OR "academic initiative" OR "academic initiatives" Date: After 1989 — 0 hits<br><br><br>Total number of hits from this database = 0 |
| UNDP | Initially, the following website was searched: http://www.undp.org/content/undp/en/home/librarypage.html. Concept 1 search terms were used only, as the website does not allow the construction of complex search strings. Each term was used individually. To keep the search as broad as possible, phrase-specific searches are not conducted (that is, the term can appear anywhere in the title/abstract/document. The basic search box in the centre of the Research and Publications page is used. This appears to generate results from full-text searches of publications archived on the website. The terms are first inputted as singular, and, in cases where results contains the plural of the term, the UNDP website indicates this and the term is adapted. For example, a search for "academic intervention" yielded no hits, but the website suggested trying "academic interventions", which yielded two hits. In contrast, a search for both "academic reform" and "academic reforms" brought up 0 hits. Since RIS files for export of citations are not supported, hits were manually screened on the website.<br><br>In addition, the following UNDP country website sections were also hand-searched (that is, titles and, where possible, abstracts were screened) for relevant papers:<br><br>http://web.undp.org/developmentstudies/researchpapers.shtml<br><br>http://www.un.org/en/development/desa/working-papers.html<br><br>http://www.africa.undp.org/content/rba/en/home/library/working-papers/ *<br><br>http://www.in.undp.org/content/india/en/home/library/hdr/human-development-working-papers/<br><br>http://www.in.undp.org/content/india/en/home/library/hdr/human-development-discussion- |

| | |
|---|---|
| | papers/<br><br>*Concept 1 search terms were used on this website.<br><br><br>No hits were found relevant across these websites. |
| UNESCO | Each term within concept 1 was searched for individually, as exact-phrase searches within brackets (..). These are not combined with other concepts due to the limited capacity of the database to interpret sophisticated search strings. Additionally, using terms from concept 1 only kept the search as broad as possible. The search strategy is, therefore, one of running each search term individually within the Words from Title search field. The "all words in field option" was selected, date restriction was applied, and only documents in English were searched. All document types were searched.<br><br><br>UNESDOC does not allow output into RIS files, nor does it allow for search results to be emailed; therefore, hits were manually screened before citations were imported into EPPI Reviewer. The following terms yielded relevant hits that were uploaded to EPPI Reviewer.<br><br><br>(child intervention)<br><br>(child initiative)<br><br>(education initiative)<br><br>(school incentive)<br><br>(education reform)<br><br>(education interventions)<br><br>(school program)<br><br>(teaching initiative)<br><br>(student intervention)<br><br>(teacher reform)<br><br>(teacher initiative)<br><br>(teaching incentive) |

| | |
|---|---|
| | This yielded 30 relevant hits. Thesaurus searches were also run for the following phrases: "teacher effectiveness", "learning methods", "teaching and training" — an additional three relevant hits were found from this search. Total number of hits from this database = 33 |
| ILO | We ran searches within the ILO document repository "Labordoc". The "advanced search" function is used. We searched using terms from concept 1, a single term at a time. In most cases, the search fields, "All of the words" within "Any field" are used. This finds the search terms anywhere in the title, author, keywords, abstract and table of contents. When a large number of hits are yielded then the search fields, "Phrases" within "Any Field" are used. Date restrictions are applied to include work from 1990 to 2014. Searches are limited to work in English. LABORDOC does not allow output into RIS files, nor does it allow for search results to be emailed; therefore, hits were manually screened before citations were imported into EPPI Reviewer. When an unmanageable number of hits was returned, the first 100 titles (and, where possible, abstracts and full texts) were screened. The following terms yielded relevant hits that were uploaded to EPPI Reviewer: Education incentive Education initiative Education program Pupil reform School initiative Teacher reform Teaching reform These yielded seven relevant hits. |

| | |
|---|---|
| | Additionally, a manual search of the thesaurus was conducted and the following terms were found relevant. Note: The subject field allows a precise search to be run, if thesaurus terms are used. No relevant hits were found from this search. |
| | Teacher education (subject) |
| | Teaching method (subject) |
| | Educational facilities (subject) |
| | Education policy (only in titles) |
| | Educational innovation (subject) |
| | Total number of hits from this database = 7 |
| CREATE | Separate searches in title and abstract fields were run, with date restrictions. Outputting in RIS files is not supported. Hits have been manually screened and uploaded to EPPI reviewer. |
| | The following strings were run on title |
| | Restricting to 20[00-14] |
| | academic reform\|academic reforms\|academic intervention\|academic interventions\|academic incentive\|academic incentives\|academic initiative\|academic initiatives\|academic program\|academic programs\|academic scheme\|academic schemes\|child reform\|child reforms\|child intervention\|child interventions\|child incentive\|child incentives\|child initiative\|child initiatives\|child program\|child programs\|child scheme\|child schemes\|classroom reform\|classroom reforms\|classroom intervention\|classroom interventions\|classroom incentive\|classroom incentives\|classroom initiative\|classroom initiatives\|classroom program\|classroom programs\|classroom scheme\|classroom schemes\|education reform\|education reforms\|education intervention\|education interventions\|education incentive\|education incentives\|education initiative\|education initiatives\|education program\|education programs\|education scheme\|education schemes\|learning reform\|learning reforms\|learning intervention\|learning interventions\|learning incentive\|learning incentives\|learning initiative\|learning initiatives\|learning program\|learning programs\|learning scheme\|learning schemes\|pupil reform\|pupil reforms\|pupil intervention\|pupil interventions\|pupil incentive\|pupil incentives\|pupil initiative\|pupil initiatives\|pupil program\|pupil programs\|pupil scheme\|pupil schemes\|school reform\|school reforms\|school |

intervention|school interventions|school incentive|school incentives|school initiative|school initiatives|school program|school programs|school scheme|school schemes|student reform|student reforms|student intervention|student interventions|student incentive|student incentives|student initiative|student initiatives|student program|student programs|student scheme|student schemes|teacher reform|teacher reforms|teacher intervention|teacher interventions|teacher incentive|teacher incentives|teacher initiative|teacher initiatives|teacher program|teacher programs|teacher scheme|teacher schemes|teaching reform|teaching reforms|teaching intervention|teaching interventions|teaching incentive|teaching incentives|teaching initiative|teaching initiatives|teaching program|teaching programs|teaching scheme|teaching schemes

Restricting to 19[90-99]

academic reform|academic reforms|academic intervention|academic interventions|academic incentive|academic incentives|academic initiative|academic initiatives|academic program|academic programs|academic scheme|academic schemes|child reform|child reforms|child intervention|child interventions|child incentive|child incentives|child initiative|child initiatives|child program|child programs|child scheme|child schemes|classroom reform|classroom reforms|classroom intervention|classroom interventions|classroom incentive|classroom incentives|classroom initiative|classroom initiatives|classroom program|classroom programs|classroom scheme|classroom schemes|education reform|education reforms|education intervention|education interventions|education incentive|education incentives|education initiative|education initiatives|education program|education programs|education scheme|education schemes|learning reform|learning reforms|learning intervention|learning interventions|learning incentive|learning incentives|learning initiative|learning initiatives|learning program|learning programs|learning scheme|learning schemes|pupil reform|pupil reforms|pupil intervention|pupil interventions|pupil incentive|pupil incentives|pupil initiative|pupil initiatives|pupil program|pupil programs|pupil scheme|pupil schemes|school reform|school reforms|school intervention|school interventions|school incentive|school incentives|school initiative|school initiatives|school program|school programs|school scheme|school schemes|student reform|student reforms|student intervention|student interventions|student incentive|student incentives|student initiative|student initiatives|student program|student programs|student scheme|student schemes|teacher reform|teacher reforms|teacher intervention|teacher interventions|teacher incentive|teacher incentives|teacher initiative|teacher initiatives|teacher program|teacher programs|teacher scheme|teacher schemes|teaching reform|teaching reforms|teaching intervention|teaching interventions|teaching incentive|teaching incentives|teaching initiative|teaching initiatives|teaching program|teaching programs|teaching scheme|teaching schemes

Within abstract, the same search strings as above have been run with the following date

| | |
|---|---|
| | restrictions: |
| | Restricting to 20[00-14] |
| | Restricting to 19[90-99] |
| | |
| | Keyword Search — A manual search of the keywords list provided the following relevant terms: |
| | Accountability |
| | Evaluation |
| | Incentive(s) |
| | Innovation |
| | Motivation |
| | Para-teachers |
| | Training |
| | |
| | Total number of hits from this database found relevant = 7 |
| Index of Conference Proceedings: Available via British Library | Concept 1 search terms are used only as proximity searching, and are not supported. Also, use of concepts 1 and 5 in separate search fields gave confusing results. Further, specific to each string, searches are refined by the subject: economics, education, psychology, social sciences. Where possible, date restrictions are applied. The "advanced search" function is used. No restrictions on Material Type or Search Scope were included. Search string was run by using the default selections: that is, "anywhere" and "contains". Screening has been done on the website itself and relevant hits are uploaded to EPPI Reviewer. Screening has been done mainly on titles, and, where possible, abstracts/full texts have been viewed. The hits have been sorted by relevance, in order to ensure that the most appropriate hits are screened. The first 100 hits are screened when hits >100. |
| | The following search strings were used; |
| | "academic reform" OR "academic reforms" OR "academic program" OR "academic programs" OR "academic scheme" OR "academic scheme" OR "academic intervention" OR "academic interventions" OR "academic incentive" OR "academic incentives" OR "academic initiative" OR |

"academic initiatives" (Refined by Education, Economics, Psychology)

"classroom reform" OR "classroom reforms" OR "classroom program" OR "classroom programs" OR "classroom scheme" OR "classroom scheme" OR "classroom intervention" OR "classroom interventions" OR "classroom incentive" OR "classroom incentives" OR "classroom initiative" OR "classroom initiatives" (Refined by Education, Psychology, Social Sciences)

"education reform" OR "education reforms" OR "education program" OR "education programs" OR "education scheme" OR "education scheme" OR "education intervention" OR "education interventions" OR "education incentive" OR "education incentives" OR "education initiative" OR "education initiatives" (Refined by Education, Economics. Refined also by After 1992).

"pupil reform" OR "pupil reforms" OR "pupil program" OR "pupil programs" OR "pupil scheme" OR "pupil scheme" OR "pupil intervention" OR "pupil interventions" OR "pupil incentive" OR "pupil incentives" OR "pupil initiative" OR "pupil initiatives" (Refined by subject — none of these subjects was relevant)

"student reform" OR "student reforms" OR "student program" OR "student programs" OR "student scheme" OR "student scheme" OR "student intervention" OR "student interventions" OR "student incentive" OR "student incentives" OR "student initiative" OR "student initiatives" (Refined by Education, Economics, Social Sciences.)

"school reform" OR "school reforms" OR "school program" OR "school programs" OR "school scheme" OR "school scheme" OR "school intervention" OR "school interventions" OR "school incentive" OR "school incentives" OR "school initiative" OR "school initiatives" (Refined by Education, Economics. Refined also by after 1991.)

"teacher reform" OR "teacher reforms" OR "teacher program" OR "teacher programs" OR "teacher scheme" OR "teacher scheme" OR "teacher intervention" OR "teacher interventions" OR "teacher incentive" OR "teacher incentives" OR "teacher initiative" OR "teacher initiatives" (Refined by Education, Economics.)

"teaching reform" OR "teaching reforms" OR "teaching program" OR "teaching programs" OR "teaching scheme" OR "teaching scheme" OR "teaching intervention" OR "teaching interventions" OR "teaching incentive" OR "teaching incentives" OR "teaching initiative" OR "teaching initiatives" (Refined by Education. Refined also by after 1995.)

"teacher training reform" OR "teacher training reforms" OR "teacher training program" OR "teacher training programs" OR "teacher training scheme" OR "teacher training scheme" OR "teacher training intervention" OR "teacher training interventions" OR "teacher training incentive" OR "teacher training incentives" OR "teacher training initiative" OR "teacher training initiatives" (Refined by Education.)

(Total number of hits = 6. Four uploaded to EPPI reviewer. One hit was a book and another was

| | |
|---|---|
| | an article to which access was not available. Entries for these are available on EPPI reviewer.) |
| SIGLE | The search form at the topof the home page is used. Even though the search form does not have separate entry fields for title, abstract, subject and keywords, field-code operators, marking out the different fields, are used at the start of the search strings to search within each of these fields. Further, given that it is specifically noted in the help files that subject terms are not filled in for the majority of records, we ran additional keyword searches.

Within each field search, four different searches have been run. These include:

Proximity searches with date restriction AFTER 1989

AND searches with date restriction AFTER 1989

Proximity searches with date restriction AFTER 2000

AND searches with date restriction AFTER 2000

Both "proximity" and "AND" searches were run to ensure the searches were as broad as possible, as the hits with the proximity searches were very few and did not always contain results from within the "AND" searches. Date restrictions "AFTER 1989" and "AFTER 2000" have been applied, because, even though help files indicated that the "AFTER" operator brings up documents with publication dates after that year, the results from "AFTER 1989 "mainly covered literature in the 1990s and the results from "AFTER 2000" mainly covered literature from the 2000s.

Finally, this database does not allow export of citations in the RIS format. The titles (and, where possible, abstracts/full text) have been screened manually on the website.

The search strings are recorded below. Due to space constraints, only the strings used for the title field are presented. The strings for the abstract, subject and keyword searches can be run by replacing the "title:" field-codes with "abstract:", "subject:" and "keyword:", respectively.

title:(("academic" OR "child*" OR "classroom" OR "education" OR "learning" OR "pupil*" OR "school*" OR "student*" OR "teach*") NEAR/3 ("reform*" OR "intervention*" OR "incentive*" |

| | |
|---|---|
| | OR "program*" OR "scheme*" OR "initiative*")) AFTER 1989

title:(("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*")) AFTER 1989

title:(("academic" OR "child*" OR "classroom" OR "education" OR "learning" OR "pupil*" OR "school*" OR "student*" OR "teach*") NEAR/3 ("reform*" OR "intervention*" OR "incentive*" OR "program*" OR "scheme*" OR "initiative*")) AFTER 2000

title:(("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*")) AFTER 2000


Total number of relevant hits found across all search strings = 0 |
| Google Scholar | Google Scholar allows only title and full-text searches. It does not allow abstract or keyword searches. Searching for individual phrases, or groups of phrases, in the full text, yields an overly large body of hits. Proximity searches and wild-card usage are not possible. Further, Google scholar truncates search strings after around 150 characters. So, given the truncation problem, it is not possible to combine concepts in order to reduce hits, as search strings that attempt to do so are cut off.


Therefore, the only search strategy that yields an analysable number of results is to search within titles only, using the "with at least one of the words" field, with search terms from concept 1 only. Dates were restricted to 1990 onwards, and publications to Education, Economics, Psychology, Sociology, Development Studies and Social Sciences titles. Also, since truncation does not allow us to run all of these phrases simultaneously, separate strings have been constructed according to their stems. Each of these strings is run separately for each discipline. Finally, the options "patents" and "citations" are unchecked, as the former refers to legal literature and the latter includes article citations, which are mainly duplicates.


Finally, even though Scholar allows imports into Endnote, multiple imports are not supported, and a single RIS file cannot be generated. Screening has been undertaken on the website itself and relevant hits have been manually uploaded to EPPI Reviewer. In cases where more than 150 hits were yielded, the first 150 titles (and, where possible, abstracts, and full-text scans) were |

| | |
|---|---|
| | screened.<br><br>An example search string is shown below. This is only displayed for the concept 1 terms that are prefixed with "academic", due to space constraints. Similar strings were constructed for the rest of the terms in concept 1 and involve replacing "academic" with "child", "classroom", "education", "learning", "pupil", "student", "school", "teacher" and "teacher training".<br><br>All in title: "academic reform" OR "academic reforms" OR "academic incentive" OR "academic initiative" OR "academic initiatives" OR "academic intervention" OR "academic interventions" OR "academic program" OR "academic programs" OR "academic scheme" OR "academic schemes" OR "academic incentives"<br><br>Total number of relevant hits = 27 |
| JPAL | The JPAL database was manually searched via the following database links:<br><br>http://www.povertyactionlab.org/evaluations<br><br>https://www.povertyactionlab.org/publications<br><br>The first link allows a search of all JPAL's randomised evaluations. Thematic searches are allowed using the theme "Education" There are 132 on-going and completed education-related evaluations available. The search was conducted by restricting the search form to just "Education". This brings up all education-evaluation projects. Where research output is available, (for example, in articles, working papers, reports), these have been screened on the website and relevant hits were uploaded to EPPI reviewer manually.<br><br>The second link allows a search of all JPAL's publications from its randomised evaluations. This was also searched in a similar way as the first link, within the "Education" theme for "Academic" publications. Many of the publications were duplicates of the first link.<br><br>Number of hits from "EVALUATIONS" = 10 hits<br><br>Number of hits from "PUBLICATIONS" = 1 hit |

| | |
|---|---|
| | Total number of hits from this database = 11. |
| 3ie: International Initiative for Impact Evaluation | Within the Title field, an initial search with the following string yielded no hits:<br><br>"academic reform" OR "academic reforms" OR "academic program" OR "academic programs" OR "academic scheme" OR "academic schemes" OR "academic intervention" OR "academic interventions" OR "academic initiative" OR "academic initiatives" OR "academic incentive" OR "academic incentives"<br><br>However, an individual search term such as "academic intervention" yielded relevant hits. Therefore, it was decided to search using each search term in concept 1 individually. To narrow the search, the following restrictions were applied in the drop-downs:<br><br>Sectors:<br><br>Education<br><br>Sub Sectors:<br><br>Education Technology<br><br>Education Inputs<br><br>Girl's Education<br><br>Primary Education<br><br>Secondary Education<br><br>Public/Private Sector Education<br><br>System Reform & Capacity Building<br><br>Countries: All low-income and lower-middle-income countries as classified by World Bank in 2013.<br><br>Year: Between 1990 and 2014. |

| | |
|---|---|
| | Finally, As no RIS/Text file exports of citations are possible, all hits were screened on the website and relevant hits were manually uploaded to EPPI Reviewer.<br><br>Total number of relevant hits = 18 |
| UNICEF | The following website was searched: http://www.unicef.org/publications/<br><br>In view of the website's limited functionality and in order to keep the search as broad as possible, individual terms from concept 1 are used as phrase-specific search terms. Subject is restricted to "Education" and "Economic and social policy". Since publications are available from 1995 onwards, no date restrictions are applied. No regional and document-type restrictions are applied. Search phrases did not appear to differentiate between plural and singular forms, and, therefore, singular searches only were run. Examples: "academic reform", "academic intervention", etc.<br><br>Finally, since no RIS exports for citations are supported, search results were manually screened on the website and relevant hits uploaded.<br><br>Total number of relevant hits = 0 |
| Centre for International al development, University of Harvard (CID) | The database was found to be unsearchable in a systematic manner. No search syntax notes are available. It is not clear how may results searches return, and no citations can be outputted. A manual screening of publication titles has been made directly from the website.<br><br>Number of hits from this database = 0 |

**Table A2.2.2: Search databases used in the review**

| Platform | Database | Details |
|----------|----------|---------|
| **Databases for published papers and reports** | | |
| EBSCO | TRC - Teacher Reference Center | Indexing and 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 | Full-text database of education research |
| | IBSS - International Bibliography of the Social Sciences | Bibliography for social science and interdisciplinary research |
| | ProQuest Dissertations & | Global |

| | | |
|---|---|---|
| | Theses | |
| | AEI | Australian Education Index |
| | PSYCINFO | Psychology journals database |
| | BEI | British Education Indexf |
| | | |
| 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 | 3ie | Impact-evaluation reports (both peer-reviewed and working papers), systematic reviews |

| | | |
|---|---|---|
| Impact Evaluation | | |
| 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 (for example, conferences), and grey literature itself, to be included by team members. | | |
| **Databases for Theses & Other** | | |
| SIGLE | SIGLE | System for Information on Grey Literature in Europe |
| 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 bibliographies/references. |
| --- | --- | --- |

## APPENDIX 2.2: 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 arose that may be interesting/relevant for the readers of this SR:

Please refer to the DFID *How To Note on Assessing the Strength of Evidence*, February 2013, pp.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? | |
| Openness and transparency | Does the study present or link to the raw data it analyses? | |
| | Does the author recognise limitations/weaknesses in their work? | |
| 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? | |
| Validity | Has the study demonstrated measurement validity? | |
| | Is the study internally valid? | |
| | Is the study externally valid? | |
| Reliability | Has the study demonstrated measurement reliability? | |

| | Has the study demonstrated that its selected analytical technique is reliable? | |
|---|---|---|
| Cogency | Does the author 'signpost' the reader throughout? | |
| | Are the conclusions clearly based on the study's results? | |

(Source: DFID, 2013, 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 A.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.)

| | |
|---|---|
| ASER | Annual Status of Education Report |
| DFID | Department for International Development |
| EPPI | Evidence for Policy and Practice Information and Co-ordinating Centre |
| FAS | Foundation Assisted Schools |
| ITPD | In-Service Teacher Professional Development |
| NGO | Non-Governmental Organisation |
| PASEC | Programme on the Analysis of Education Systems |
| PICOS | Population, Intervention, Comparison and Outcomes |
| RCT | Randomised Control Trial |
| RQ | Research Question |
| SNED | National System of School Performance Assessment |
| SSA | Sarva Shiksha Abhyaan |
| TCAI | Teacher Community Assistant Initiative |
| TESSA | Teacher Education in Sub-Saharan Africa |
| TIMMS | Trends in International Mathematics and Science Study |
| WoE | Weight of Evidence |