Routledge
Taylor & Francis Group

Journal of Development Effectiveness

ISSN: 1943-9342 (Print) 1943-9407 (Online) Journal homepage: http://www.tandfonline.com/loi/rjde20

The effectiveness of school-based decision making in improving educational outcomes: a systematic review

Roy Carr-Hill, Caine Rolleston, Rebecca Schendel & Hugh Waddington

To cite this article: Roy Carr-Hill, Caine Rolleston, Rebecca Schendel & Hugh Waddington (2018) The effectiveness of school-based decision making in improving educational outcomes: a systematic review, Journal of Development Effectiveness, 10:1, 61-94, DOI: 10.1080/19439342.2018.1440250

To link to this article: https://doi.org/10.1080/19439342.2018.1440250

JOURNAL OF DEVELOPMENT EFFECTIVENESS, 2018
VOL. 10, NO. 1, 61–94
https://doi.org/10.1080/19439342.2018.1440250

Routledge
Taylor & Francis Group

ARTICLE

The effectiveness of school-based decision making in improving educational outcomes: a systematic review

Roy Carr-Hill[a], Caine Rolleston [a], Rebecca Schendel[a] and Hugh Waddington[b]

[a]Department of Education, Practice and Society, UCL Institute of Education, London, UK; [b]3ie, London, UK

ABSTRACT

The rhetoric around decentralisation suggests school-based management improves education outcomes. Existing reviews on school-based decision-making have tended to focus on proximal outcomes and offer very little information about why school-based decision-making has positive or negative effects in different circumstances. The authors systematically searched for and synthesised evidence from 35 quantitative and qualitative studies evaluating 17 individual interventions on the effectiveness of school-based decision-making on educational outcomes. Devolving decision-making to the level of the school appears to have a somewhat beneficial effect on dropout, repetition and teacher attendance. Effects on test-scores are more robust, being positive in aggregate and for middle-income countries specifically. On the other hand, school-based decision-making reforms appear to be less effective in communities with generally low levels of education, where parents have low status relative to school personnel. The authors conclude that school-based decision-making reforms are less likely to be successful in highly disadvantaged communities.

ARTICLE HISTORY
Received 8 February 2018
Accepted 9 February 2018

KEYWORDS
Decentralisation; education outcomes; systematic review

Background

Education is understood to be a fundamental human right that offers individuals the opportunity to live healthy and meaningful lives. Evidence from around the world also indicates that education is vital for economic and social development, as it contributes to economic growth and poverty reduction, sustains health and well-being and lays the foundations for open and cohesive societies (UNESCO 2014). In recognition of the vital importance of education, governments across the globe have made a substantial effort to expand and improve their education systems, as they strive to meet the Education for All goals, adopted by the international community in 1990. These efforts have borne remarkable results; it is estimated that the number of out-of-school children has halved over the last decade (UNESCO 2014, 53). However, there are still serious barriers to overcome, particularly in terms of access, completion and learning (Krishnaratne, White, and Carpenter 2013). Access to education – particularly for girls, poor children and children in conflict-affected areas – remains a crucial issue.

The devolution of decision-making authority to schools has been widely adopted as a decentralisation model by international agencies, including the World Bank, the US Agency for International Development and the UK Department for International Development (DFID) and by governments, as it is assumed that locating decision-making authority within schools will increase accountability, efficiency and responsiveness to local needs (Gertler, Patrinos, and Rubio-Codina 2008). This devolution includes a wide variety of models and mechanisms, differing in terms of

CONTACT Caine Rolleston c.rolleston@ucl.ac.uk

which decisions are devolved (and how many), to whom decision-making authority is given and how the decentralisation process is implemented (that is, through 'top-down' or 'bottom-up' processes). All models and mechanisms are presumed to increase responsiveness to local needs and accountability by bringing community members into direct contact with schools, and to increase efficiency by making financial decisions more transparent to communities, reducing corruption and incentivising investment in high-quality teachers and materials.

However, there is limited evidence from low-income countries of the general relationship between decentralisation reforms and education outcomes. Much of the literature focuses exclusively on the proximal outcomes of school-based decision-making. This is likely due to the relative ease of measuring such outcomes, as well as the shorter time generally required to identify impact on intermediate outcomes.

Existing systematic reviews on school-based decision-making have also tended to focus on proximal outcomes (for example, Guerrero et al. 2012; on teacher absenteeism; Petrosino et al., 2012; on student enrolment). There are very few that consider the full range of relevant outcomes, such as student learning (for example, Snilstveit et al. 2015). The comprehensive reviews that do exist (Santibanez 2007; World Bank 2007; Bruns, Filmer, and Patrinos 2012; Westhorp, Walker, and Rogers 2014) are not systematic reviews as commonly defined, with clear inclusion criteria, systematic literature searches and transparent appraisal and synthesis of the evidence.[1] Most of the reviews rely on literature that is now nearly ten years out of date and focus almost exclusively on Central America, referencing almost no evidence from other regions. Existing reviews on this topic also tell us very little about why school-based decision-making has positive or negative effects in different circumstances, a gap which this review also aims to address. There is, therefore, a need for a current globally comprehensive systematic review of the impact of school-based decision-making on a wide range of educational outcomes.

In the following sections, we define school-based decision-making interventions and how they are supposed to work. We then present the review objectives and methods, followed by synthesis of evidence on effects and discussion of the mechanisms underpinning positive and negative effects. Finally, we give implications for policy, programmes and research.

School-based decision-making reforms

Decisions about curricula, finance, management and teachers can all be taken at one or more of several administrative levels: centrally at the national or federal state level, by provinces/regions within a country, by districts or by schools. Often described as 'school-based' or 'community based' management, the devolution of decision-making authority to schools includes a wide variety of models and mechanisms. These differ in terms of which decisions are devolved (and how many), to whom decision-making authority is given and how the decentralisation process is implemented (that is, through 'top- down' or 'bottom-up' processes).

'School-based decision-making' can describe models in which decisions are taken by an individual principal or head teacher, by a professional management committee within a school or by a management committee involving local community members. This last model may simply imply an increased role for parents in the management and activities of the school, or it may include provision of training and materials to empower broader community involvement, depending on the model (Krishnaratne, White, and Carpenter 2013).

The devolved decisions can be financial (for example, decisions about how resources should be allocated within a school; decisions about raising funds for particular activities within a school), managerial (for example, human resource decisions, such as the monitoring of teacher performance and the power to hire and fire teachers; decisions relating to the management of school buildings and other infrastructure) or related to the curriculum and/or pedagogy (for example, decisions about how elements of a national curriculum will be taught and assessed within a given school). In order to support the process of decision-making, many models also involve some means

of providing information to community members on the performance of an individual school (or school district) relative to other schools (Barrera-Osorio and Linden 2009). These models and mechanisms are considered to potentially increase accountability and responsiveness to local needs by bringing local community members into more direct contact with schools, and to increase efficiency by making financial decisions more transparent to communities, thereby reducing corruption and incentivising investment in high-quality teachers and materials.

For the purposes of this review, 'school-based decision-making' is defined as including any model in which at least some of the responsibility for making decisions about planning, management and/or the raising or allocation of resources is located within schools and their proximal institutions (for example, community organisations), as opposed to government authorities at the central, regional or district level. The 'intervention' considered within this review, therefore, is *any reform in which decision-making authority is devolved to the level of the school*. Within this broad definition, there are three main mechanisms discussed in the literature: (1) reforms that devolve decision-making around *management* to the school level; (2) reforms that devolve decision-making around *funding* to the school level and (3) reforms that devolve decision-making around *curriculum, pedagogy and other aspects of the classroom environment* to the school level.

School-based decision-making is widely promoted by donors in lower-income countries as a means for improving educational quality and is often taken up enthusiastically by national governments. Both generally articulate the ultimate outcome of school-based decision-making models as being a positive change in student outcomes (including but not restricted to learning outcomes). In addition to learning outcomes (most often measured by standardised tests of cognitive skills), there are many other possible student learning outcomes which may be valued by schools, donors and governments, such as improved student ability to demonstrate psychosocial and 'non-cognitive' skills. Changes in student aspirations, attitudes and behaviours (such as the adoption of safe sex practices) could also be considered important educational outcomes.

However, devolving decision-making to the level of the school does not lead *directly* to such outcomes. Rather, school-based decision-making is likely to impact on outcomes through a number of causal pathways. We developed a conceptual framework depicting our understanding of the causal pathways, contributing factors and underlying processes that could affect the impact of school-based decision-making on educational outcomes (Figure 1), which we used as a 'working hypothesis' (Oliver, Dickson, and Newman 2012, 68) to guide the articulation of our specific review questions and review methodology (as recommended by Anderson et al. 2011).

Reforms that increase accountability and responsiveness to local needs are assumed to lead to positive stakeholder perceptions of (and engagement in) educational provision, which, in turn, is expected to increase enrolment, attendance and retention and to reduce corruption within schools. It is also presumed that increased accountability will encourage schools to make recruitment decisions based on teacher performance, rather than mechanically relying on qualifications or allowing nepotism to interfere. Such personnel practices, in turn, are seen to lead to reduced teacher absenteeism, increased teacher motivation and, ultimately, improvements in the quality of teaching within schools. It is also assumed that local communities will encourage schools to adopt more locally relevant curricula, which can then have a positive impact on the quality of teaching and student opportunities to learn.

At the same time, decentralised funding mechanisms and other reforms aimed at increasing efficiency within schools, particularly when combined with efforts to increase community participation, are presumed to result in more resources being available to schools, another important factor in improving educational quality (Krishnaratne, White, and Carpenter 2013). Increased efficiency is, in turn, assumed to affect the unit costs of educational provision, potentially reducing costs or improving outcomes for a given cost, which may be particularly valued by governments in less well-resourced settings. School-based decision-making mechanisms, therefore, result in many proximal (or intermediate) outcomes, in addition to the final outcomes mentioned above. These

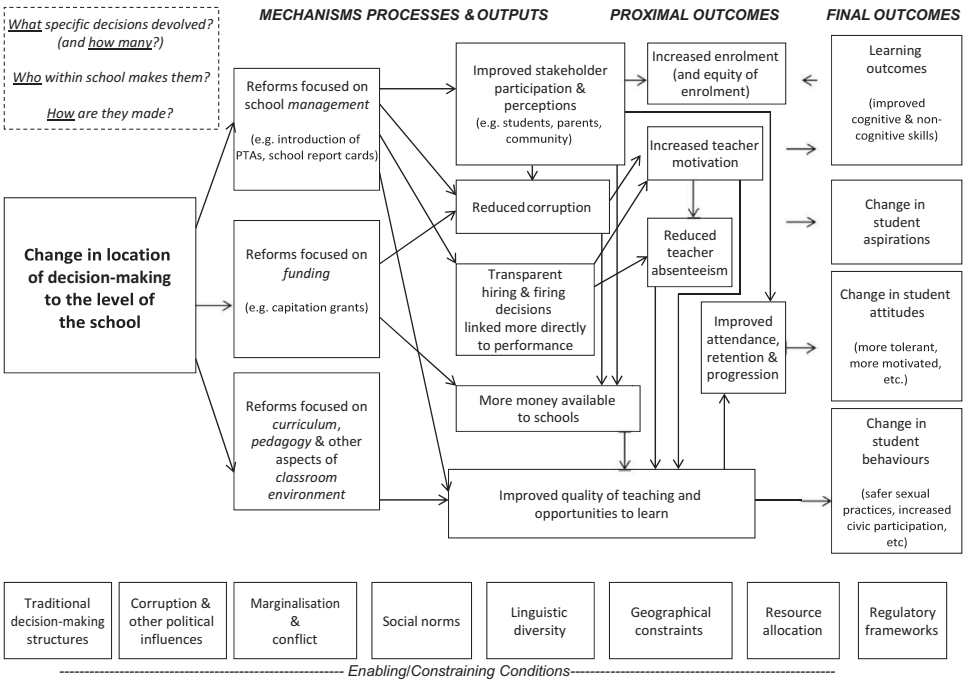

Figure 1. Conceptual framework.
Source: authors.

proximal outcomes include increased enrolment, improved equality of access, improved attendance, improved retention, improved progression and higher quality educational provision.

However, there is growing evidence that decentralisation reforms may have unintended and sometimes negative effects in certain political and economic circumstances (Banerjee et al. 2008; Bardhan and Mookherjee 2000; Bardhan and Mookherjee 2005; Carr-Hill et al. 1999; Condy 1998; Glassman, Naidoo, and Wood 2007; Pherali, Smith, and Vaux 2011; Rocha Menocal and Sharma 2008; Rose 2003; Unterhalter 2012). Decentralising decision-making may lead to 'elite capture' at the local level and/or further corruption within school systems, for example, or may limit educational opportunity for marginalised ethnic groups. There is some consensus in this literature that decentralisation is only likely to have a positive impact on outcomes when (a) there is clear government policy and/or regulations about the powers and role played by different agencies and stakeholders; (b) there are sufficient financial resources available within the system and (c) there is some form of democratic culture (see De Grauwe et al. 2005; Lugaz et al. 2010; Pherali, Smith, and Vaux 2011). Those vested with the authority to make decisions on behalf of the school must also have the capacity and knowledge to make such decisions, or their decisions are unlikely to have a positive impact on outcomes (World Bank 2004). This body of evidence highlights the contingency of the effects of decentralisation, linked to important interactions between formal structures of decision-making and informal structures of power and authority within bureaucracies, communities and schools.

Furthermore, each link in the causal chain rests on certain assumptions which must be met for a change in the location of decision-making to have the desired effect(s). For instance, the assertion that involving parents and community members in the hiring and firing of teachers (an 'accountability' mechanism employed in many contexts) will improve quality of teaching rests on the assumption that (a) parents and community members will be able to identify high-quality teachers who should be retained and/or rewarded, (b) the incentives provided will positively impact student

learning and (c) former more centralised systems were less than optimal with regard to teacher recruitment and accountability, leaving scope for improvement through reform. This is not always achieved. In some contexts, teacher incentive schemes have been found to have a negative impact on overall student learning, if, for instance, they create perverse incentives for teachers to block the enrolment of low-performing students to maintain high average test scores within their classrooms (Glewwe, Ilias, and Kremer 2003). The impact of school-based decision-making models is, therefore, likely to differ depending on a wide variety of implementation factors, relating to the objective of the reform, the decisions that are devolved, the individuals given decision-making authority and the nature of the decision-making process.

Objectives and methods

The review aims to answer two questions:

(1) What are the impacts of school-based decision-making on educational outcomes in low- and middle-income countries (LMICs)?
(2) What are the barriers to and enablers of effective models of school-based decision-making?

The review followed an explicit protocol (see Carr-Hill et al. 2014). Full details of the review approach are provided by Carr-Hill, Rolleston, and Schendel (2016).

To be included in the review, all studies had to: (1) be empirical in nature and focused on primary and secondary schools within LMICs[2]; (2) investigate a change in decision-making authority from a higher level of decision-making authority to the level of the school; (3) provide data on the relationship between school-based decision-making and at least one educational outcome (either proximal, for example, attrition, equality of access, increased enrolment or final, for example, student learning, as captured by test scores, psychosocial and non-cognitive skills, and so forth) and (4) rely on data collected since 1990 be reported in English, French, Spanish or Portuguese.[3] Studies of any follow-up duration and studies with multiple follow-ups were included.

We excluded evidence collected in LMICs located within Central and Eastern Europe (including Turkey) or the former USSR, and studies where the intervention was conceptualised, managed and implemented by an external decision-making agency. Studies of interventions aimed exclusively at improving the functioning of devolved decision-making structures – but not introducing new decision-making authority – were excluded (for example, interventions aimed at strengthening the effectiveness of pre-existing village education committees, such as the report card initiative discussed by Banerjee et al. 2008). We also excluded studies investigating a change in decision-making authority to a level higher than the school (for example, studies of decentralisation to the region or district level). Studies that investigated the effects of privatisation of schooling were excluded on a related basis. Further, studies focusing on decision-making at levels lower than the school were also excluded. These include demand-side interventions (for example, conditional cash transfers) intended to influence decisions made at the household, family or child-level.

We conducted a mixed methods review. To be eligible for review under RQ1, studies needed to be causal in nature, meaning we included: (1) experimental designs using randomised or quasi-randomised assignment and (2) quasi-experimental designs collecting longitudinal data at baseline and endline in intervention and comparison groups, as well as those using cross-sectional endline data only, provided an appropriate method of analysis has been used to control for confounding. Any comparison needed to be contemporaneous – that is, data on a reform group and a non-reform group needed to reflect the same time period. All the included studies needed to analyse data at the level of the child or at the level of the school or community. Studies analysing comparison groups at sub-national or country level were excluded, as were studies in which there was clear evidence of spillovers or contamination to comparison groups from the same communities, and studies in which reporting biases were evident were excluded.[4]

For RQ2, we included studies of any empirical design, so long as they provided evidence on contexts already included in the review, and met the standards of transparency, appropriateness, rigour, validity, reliability and cogency set out by DFID (2014). These included 'process evaluations' and/or project completion reports of any of the school-based decision-making interventions evaluated in reference to the first review question, other empirical studies (employing quantitative, qualitative or mixed methods of analysis) which provided data on either factors found to affect the implementation of one of the school-based decision-making interventions evaluated in reference to the first review question, or conditions or circumstances found to affect the impact of one of the included interventions on the specified outcome(s). Studies reporting stakeholder perceptions of a change in outcomes were excluded, as were studies exclusively reporting on processes or outputs (for example, changes in the frequency of community participation).

Potentially relevant literature was identified through a five-stage search strategy for published literature (for example, journal articles, books, conference papers and institutional grey literature, including reports and process evaluations) and unpublished literature (for example, dissertations, theses and unpublished empirical studies showing null and/or negative results). This comprised: (1) identification of existing systematic reviews in related areas; (2) targeted searches in a wide range of bibliographic databases and websites[5]; (3) hand searches of the eight most relevant journals relating to the topic; (4) citation chasing[6] and (5) contacting experts involved in the research area. Relevant studies were then appraised for robustness of evidence and methodological rigour prior to synthesis.

We identified 2821 titles (135 from systematic reviews, 2141 from databases and 541 from website and hand searches)[7] of which 100 met the review eligibility criteria (Figure 2). A total of 30 of the 100 met the design criteria required for RQ1, but three were removed from RQ1 synthesis, due to high risk of bias.[8] A total of 27 studies were, therefore, included in the review of which 26 studies, investigating the impact of 17 individual interventions, were included in meta-analysis.[9] A total of 23 non-causal studies were identified and critically appraised, of which 9 were included.

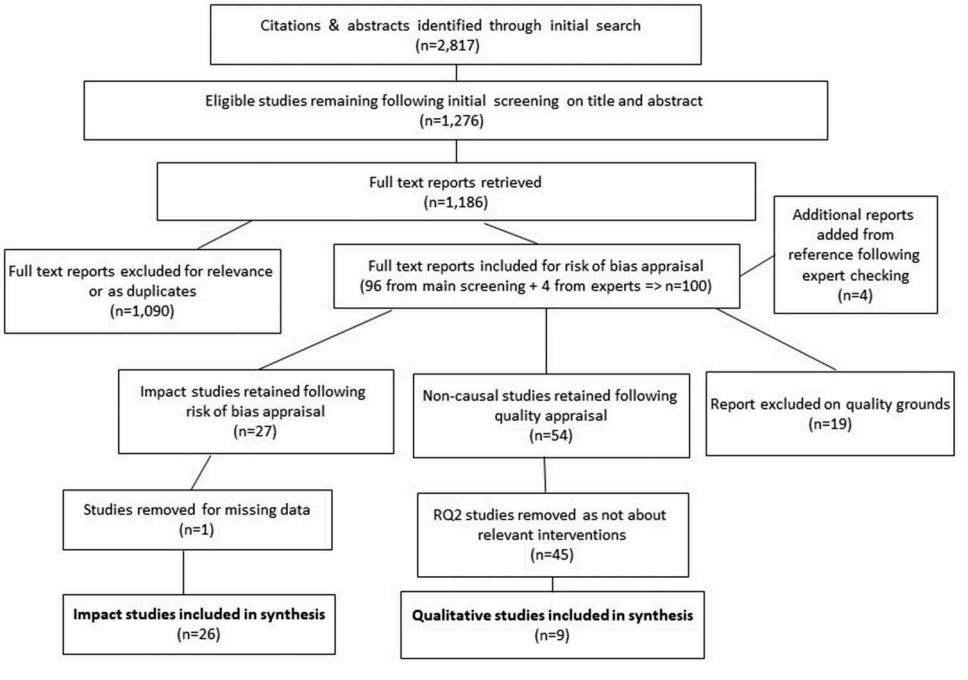

Figure 2. Pipeline of studies.

Table 1. Characteristics of included intervention studies.

Name of intervention	Country, school level	Intervention description	Level of decentralisation	Impact studies (RQ1) and linked studies (RQ2)	Outcomes included in meta-analysis	RQ1 study design (risk of bias)
PDE	Brazil, primary schools	The programme began in 1998 and was administered by the World Bank-financed Fundescola Project. Schools conducted needs assessment and developed plans to address them, which Fundescola project funded. Initially, 401 schools in 9 states took part. In 2001, the programme was rolled out to more than 5600 schools, with total investment over 1 million dollars.	Medium decentralisation	RQ1: Carnoy et al. (2008) RQ2: N/A	None (data not available)	Quasi-experiment (moderate risk of bias)
EDUCO	El Salvador, primary schools	EDUCO, established in 1991, is a national programme that gives communities autonomy over most educational decisions. Under the EDUCO model, community education associations – in which parents are the majority – are responsible for administering and managing the school, including hiring, firing and paying teachers. Community education association members are elected by their peers and receive training on various aspects of school management. Community education association members must be literate and they are elected by their peers. They also receive training prior to assuming their duties.	High decentralisation	RQ1: Jimenez and Sawada (1999); Jimenez and Sawada (2003); Sawada and Ragatz (2005) RQ2: De Umanzor et al. (1997)	Test scores (language, math)	Quasi- experiment (moderate risk of bias)
PROHECO	Honduras, primary schools	The EDUCO programme spawned a number of similar initiatives in Central America, including PROHECO in Honduras. Much like EDUCO, PROHECO schools are managed by parental councils, which have responsibility for a broad range of management duties, including hiring, firing and paying teachers.	High decentralisation	RQ1: Di Gropello and Marshall (2005) RQ2: N/A	Test scores (language, math)	Quasi-experiment (moderate risk of bias)

(*Continued*)

Table 1. (Continued).

Name of intervention	Country, school level	Intervention description	Level of decentralisation	Impact studies (RQ1) and linked studies (RQ2)	Outcomes included in meta-analysis	RQ1 study design (risk of bias)
Autonomous Schools Program	Nicaragua, primary schools	In the early 1990s, the Nicaraguan government established 'consultative councils' in all public schools, in order to stimulate greater participation of teachers and parents in school decisions. Councils consisted of head teachers, teachers, parents and students. In 1993, the consultative councils, at a small sub-sample of public secondary schools, were transformed into School Management Councils in 1993 and given legal status and autonomy over the majority of school decisions. This pilot programme eventually expanded into primary education in 1995. The councils of the newly created autonomous schools, in which parents held the voting majority, had the ability to hire and fire teachers and the responsibility to maintain their infrastructure and academic quality. They also had control over monthly fiscal transfers that paid for teacher salaries, benefits and basic maintenance, and they had the right to charge and retain fees. The Ministry of Education retained control over staff promotion, teacher certification and the national curriculum.	High decentralisation	RQ1: King and Ozler (2005); Parker (2005) RQ2: Fuller and Rivarola (1998); Gershberg and Meade (2005)	Test scores (language, math)	Quasi-experiment (moderate risk of bias)
School-based management	Indonesia, primary schools	School-based management was established in Indonesia in 2003. SBM grants principals, teachers and other local community-based members with autonomy over academic operations of schools. A grant programme accompanied the reform, which provided a per- student amount to all schools that could be disbursed according to local priorities. In 2006, recognising that school committees were largely not realising the autonomy granted to them through the reform, a field experiment was implemented by the World Bank to test four measures aimed at helping committees to fulfil their management roles.	Medium decentralisation	RQ1: Pradhan et al. (2011) RQ2: Bandur (2008); Bjork (2003); Vernez, Karam, and Marshall (2012)	Dropout Repetition Test scores (language, math)	RCT (low risk of bias)

(*Continued*)

Table 1. (Continued).

Name of intervention	Country, school level	Intervention description	Level of decentralisation	Impact studies (RQ1) and linked studies (RQ2)	Outcomes included in meta-analysis	RQ1 study design (risk of bias)
AGEMAD	Madagascar, primary schools	The AGEMAD reform sought to improve the efficiency and effectiveness of the education sector in Madagascar by specifying roles and responsibilities and introducing new monitoring tools at each level of the school management hierarchy. At the school level, the intervention focused on the provision of new administrative tools for teachers (for example, lesson planning forms), the introduction of school report cards and the organisation of school meetings with school staff, parents and members of the community (intended to increase parental and community involvement in monitoring). An RCT was designed to test the impact of three different implementation designs: (1) a cascade model in which district officials were trained to implement the reform through the district; (2) a modified cascade model in which both district and sub-district officials were trained the implement the reform and (3) an intensive model in which district officials, sub-district officials and individual schools were trained directly.	Medium decentralisation	RQ1: Glewwe and Maïga (2011); Lassibille et al. (2010) RQ2: N/A	Test scores (aggregate)	RCT (moderate risk of bias)

(*Continued*)

Table 1. (Continued).

Name of intervention	Country, school level	Intervention description	Level of decentralisation	Impact studies (RQ1) and linked studies (RQ2)	Outcomes included in meta-analysis	RQ1 study design (risk of bias)
Quality Schools Program (PEC)	Mexico, secondary schools	PEC was introduced in 2001 and seeks to increase community participation in school-based decision-making, reducing administrative burden for participating schools and providing technical support to participating schools. The programme is guided by national regulations of the federal government but administered by state departments. The federal government provides match funding to encourage state participation in the funding of PEC. In order for a school to qualify for PEC, school directors, teachers and parents need to identify a school's problems and needs and design a school improvement plan. PEC schools qualify for annual programme grants of up to about $5000 and also receive $2 for each dollar that the school raises from the municipal government or private sector. The grant amount depends on the socio-economic status of the community, the educational needs identified in the school improvement plan and the characteristics of the community population. Communities must spend 80 per cent of their grant in the first four years; funds must be spent on teacher training, interventions for at-risk students, educational materials/teaching equipment or infrastructure. Training is provided to school principals and directors of school-management committees.	Medium decentralisation	RQ1: Bando (2010); Murnane, Willett, and Cardenas (2006); Skoufias and Shapiro (2006) RQ2: Reimers and Cardenas (2007)	Dropout Test scores (language, math)	Quasi-experiment (low risk of bias)

(*Continued*)

Table 1. (Continued).

Name of intervention	Country, school level	Intervention description	Level of decentralisation	Impact studies (RQ1) and linked studies (RQ2)	Outcomes included in meta-analysis	RQ1 study design (risk of bias)
Support to School Management (AGE)	Mexico, primary schools	AGE, a precursor to PEC, was implemented in the late 1990s as part of a broader school reform that aims to improve service delivery and education quality in highly deprived parts of Mexico. AGE provides a small amount of financial support ($500–$700 per year depending on the school size) to parents' associations who have autonomy in using the funds for school improvement. Parents receive training about the role of parent association, the use of school funds and how to participate in a range of activities that involve effective management of the school. The use of funds is restricted and cannot be used to fund salaries.	Medium decentralisation	RQ1: Gertler, Patrinos, and Rubio-Codina (2012) RQ2: N/A	Dropout Repetition	Quasi- experiment (moderate risk of bias)
Program to Strengthen and Invest Directly in Schools (PEC-FIDE)	Mexico, primary schools	PEC-FIDE was a spin-off of PEC, implemented in six Mexican states in 2008. Schools that had participated in PEC were also eligible for PEC-FIDE, but it was not possible for schools to receive funds from both programmes simultaneously. PEC-FIDE was similar to PEC, in that schools received grants in exchange for collaborative school planning and decision-making. The amount of the grant depended on school enrolment but generally averaged around $4500. Funds could be spent on training, interventions for at-risk students, materials/equipment and infrastructure. School councils – comprising head- teachers, teacher representations and parent representatives – were responsible for drafting School Improvement Plans and received training prior to receipt of the grant. Crucially, schools do not opt in to PED-FIDE; they are assigned to the programme by the state government, depending on programme targets.	Medium decentralisation	RQ1: Santibañez, Abreu-Lastra, and O'Donoghue (2014) RQ2: N/A	Dropout Test scores (language, math)	Quasi-experiment (moderate risk of bias)

(*Continued*)

Table 1. (Continued).

Name of intervention	Country, school level	Intervention description	Level of decentralisation	Impact studies (RQ1) and linked studies (RQ2)	Outcomes included in meta-analysis	RQ1 study design (risk of bias)
Third Elementary Education Project (TEEP)	Philippines, primary schools	TEEP, implemented from 2000 to 2006 by the Philippine Department of Education, targeted the most deprived public primary and elementary schools in the Philippines. The act legalising the reform (Republic Act 9155) vested decision-making authority in the office of the school head, not in the broader community. The Act also grants managerial autonomy, not financial freedom nor autonomy over personnel decisions. Under TEEP, schools received cash grants for maintenance and operating expenses, based on the enrolment of the school. Schools were also allowed and encouraged to raise their own funds from their communities. TEEP was a well-resourced programme that combined physical and soft components with institutional reform. The programme invested in physical buildings and textbooks, provided training to teachers and principals, and facilitated partnership between the school and community.	Medium decentralisation	RQ1: Khattri, Ling., and Jha (2010); Yamauchi and Liu (2012) RQ2: N/A	Test scores (math, aggregate)	Quasi- experiment (moderate risk of bias)
School-based management	Philippines, primary schools	Prior to the implementation of TEEP, there was a national law in the Philippines that granted principals autonomy over academic, administrative and financial affairs in their schools. Although the law encouraged the creation of school management committees, there was no mandate to create such committees, so they were only created if individual principals so desired.	Medium decentralisation	RQ1: San Antonio (2008) RQ2: N/A	Test scores (math, aggregate)	Quasi- experiment (moderate risk of bias)

(*Continued*)

Table 1. (Continued).

Name of intervention	Country, school level	Intervention description	Level of decentralisation	Impact studies (RQ1) and linked studies (RQ2)	Outcomes included in meta-analysis	RQ1 study design (risk of bias)
BESRA	Philippines, primary schools	Building on the success of TEEP, in 2006, the Philippine government mainstreamed SBM by including it as an element of the system-wide Basic Education Reform Agenda (BESRA). BESRA was built around five key reform thrusts relating to teacher development, social support for schools, early childhood development, private sector involvement in education and general improvement of educational governance. The SBM component involved the establishment (or capacity building for existing) school governing councils, the preparation of school improvement plans and an increased level of resources managed and controlled at the school level. As part of BESRA, principals and other school staff received training. BESRA was scaled up to schools that were outside the original TEEP catchment area through the use of a partnership model under which non-TEEP schools were partnered with neighbouring TEEP divisions in order to introduce SBM. Unlike TEEP, BESRA did not involve any additional package of investments.	Medium decentralisation	RQ1: World Bank (2013); Yamauchi (2014) RQ2: N/A	Test scores (aggregate)	Quasi- experiment (moderate risk of bias)

(*Continued*)

Table 1. (Continued).

Name of intervention	Country, school level	Intervention description	Level of decentralisation	Impact studies (RQ1) and linked studies (RQ2)	Outcomes included in meta-analysis	RQ1 study design (risk of bias)
Programme for School Improvement	Sri Lanka, primary schools	PSI was designed to increase involvement of the school community including parents, teachers and past pupils in the management of school. The programme emphasised development of a school improvement plan, efficient utilisation of resources and improved cooperation between schools and communities in order to enhance quality of curricular and co-curricular activities. It also prioritised staff training to address the school needs and improve relationship between schools and communities. Under PSI, School Development Committees became responsible for managing schools. A Report Card Programme (SRCP) was implemented simultaneously, on a relatively small scale, in order to inform the school community of their school's performance. Report cards were completed by school personnel and distributed to parents and School Development Committee members.	Medium decentralisation	RQ1: World Bank (2011) RQ2: N/A	Teacher attendance Test scores (language, math)	RCT (moderate risk of bias)
Rural Education Program	Colombia, multiple school levels	The Rural Education Program empowers municipal operating units (comprising local officials and members of the education sector) to assess needs and choose educational interventions for rural communities. Schools in the project are given the authority to implement/monitor their chosen educational intervention and are also provided with a 'basket' of educational goods and teacher training.	Low decentralisation	RQ1: Rodriguez et al. (2010) RQ2: N/A	Dropout Test scores (language, math)	Quasi-experiment (moderate risk of bias)
Whole School Development	Gambia, primary schools	The WSD programme provided training for head teachers, teachers and representatives of students and parents, in addition to a capitation grant. Grants were controlled by school management committees and could only be spent on teaching and learning activities.	Medium decentralisation	RQ1: Blimpo and Evans (2011) RQ2: N/A	Teacher attendance Test scores (language, math)	RCT (low risk of bias)

(*Continued*)

Table 1. (Continued).

Name of intervention	Country, school level	Intervention description	Level of decentralisation	Impact studies (RQ1) and linked studies (RQ2)	Outcomes included in meta-analysis	RQ1 study design (risk of bias)
School-based Management pilot programme	Niger, primary schools	This pilot programme in Niger provided capitation grants to schools. No restrictions were placed on the use of the funds, except that parent associations were given complete authority over their use. Training was provided to committee members prior to disbursement of the grants.	Medium decentralisation	RQ1: Beasley and Huillery (2014) RQ2: N/A	Dropout Teacher attendance Test scores (language, math)	RCT (low risk of bias)
Extra Teacher Program	Kenya, primary schools	Parent teacher associations in Kenya traditionally used money raised through school fees to hire short-term contract teachers. However, when the introduction of Universal Primary Education eliminated fees, PTAs no longer had funding available for teacher recruitment. ETP was designed to reinstate the possibility of contract teacher contracts by providing funds to a random sample of school management committees in Western Kenya. Under the programme, SMCs had the authority to hire and monitor contract teachers. A random subsample of schools in the study were provided additional training for SMC members as a supplementary intervention which was found to reduce the likelihood of reduced effort by non-contract teachers. The programme was subsequently scaled up to the national level.	High decentralisation	RQ1: Bold et al. (2013); Duflo, Dupas, and Kremer (2012) RQ2: N/A	Teacher attendance Test scores (language, math, aggregate)	RCT (low risk of bias)
Evaluation of a participatory report card intervention	Uganda, primary schools	An evaluation was designed to test the relative impact of two kinds of school report card: a standardised report card, designed by the Ministry of Education, and a participatory report card, designed by individual school management committees. Committee members were trained in both treatment groups, but only those in the participatory arm were given the freedom to design their own instrument.	Medium decentralisation	RQ1: Barr et al. (2012) RQ2: N/A	Teacher attendance Test scores (aggregate)	RCT (low risk of bias)

Two studies included to answer RQ2 are multi-country studies (Gunnarsson et al. 2008; Hanushek, Link, and Woessmann 2011).

We created a typology of broad intervention types, based on typologies of school- based management models included by Barrera-Osorio et al. (2009) and Santibanez (2007). The typology was created by coding each study on a range of dimensions, based on elements of our initial conceptual framework, from which we identified three broad intervention types:

(1) *High decentralisation* (4 interventions), comprising all models in which the school (and/or the local community) has decision-making authority over nearly all aspects of school management. Most importantly, the school or school management committee needed to have authority over both financial and personnel decisions (for example, the authority to hire/fire teachers and the authority to pay salaries).
(2) *Medium decentralisation* (13 interventions), in which a school or school management committee needed to have authority over some management decisions. However, schools in this classification would not have authority over personnel decisions.
(3) *Low decentralisation* (1 intervention) includes models in which schools have the power to make curricular decisions and/or decisions about infrastructure and buildings. No schools in this classification have authority over financial decisions.

We calculated standardised mean differences (SMDs) and associated standard errors from studies to compare effects across studies. SMD provides an estimate of the change in outcomes between intervention and control/comparison groups measured in the standard deviation (SD) of the outcome of interest. It is therefore comparable across studies, subject to certain assumptions.[10] Standardised mean differences are scaled naturally so that, for example, an effect size estimate of 0.10 denotes one-tenth of a SD improvement for treatment participants compared to control participants.

We estimated pooled effect sizes using random effects meta-analysis models with inverse variance weights. We explored heterogeneity across studies and within studies, given the variation in samples, interventions, countries and study design methods. Finally, in order to identify the main barriers and enablers that appear to have influenced the impact of the interventions in particular contexts, we examined within-study findings using a framework synthesis approach (Thomas, Harden, and Newman 2012).

Description of included interventions

In total, the 35 included studies investigate the effectiveness of 17 individual interventions (Table 1). Many of the 26 included impact studies (RQ1) involve multiple 'treatment' arms, each reflecting a slightly different variation of school-based decision-making. Of the nine linked studies (RQ2), seven relate to four of the interventions investigated in the impact studies and the remaining two are multi-country studies (Gunnarsson et al. 2008; Hanushek, Link, and Woessmann 2011).

The studies represent a diversity of geographic contexts. The region most heavily represented is Latin America (n = 12), with Mexico (n = 5), El Salvador (n = 3) and Nicaragua (n = 2) being the most common individual countries, and Colombia and Honduras are also represented. This is unsurprising, given that Latin American countries were amongst the first lower income contexts to attempt to decentralise their education systems. Seven of the studies investigate school-based decision-making in sub-Saharan African contexts (Kenya, Madagascar, Gambia, Niger and Uganda). No African country featured in more than two studies. Finally, seven studies analyse South or Southeast Asian contexts, with the Philippines being the most frequent (n = 5). Other Asian countries include Indonesia and Sri Lanka.

The studies are also quite diverse in terms of income classification. Of the 26 impact studies, 8 were based on low-income contexts, 13 in lower middle-income contexts and 5 in upper middle-income contexts. Most of the studies investigated interventions targeted at primary schools (n = 23) or secondary level (n = 1), while two studies considered outcomes at both primary and

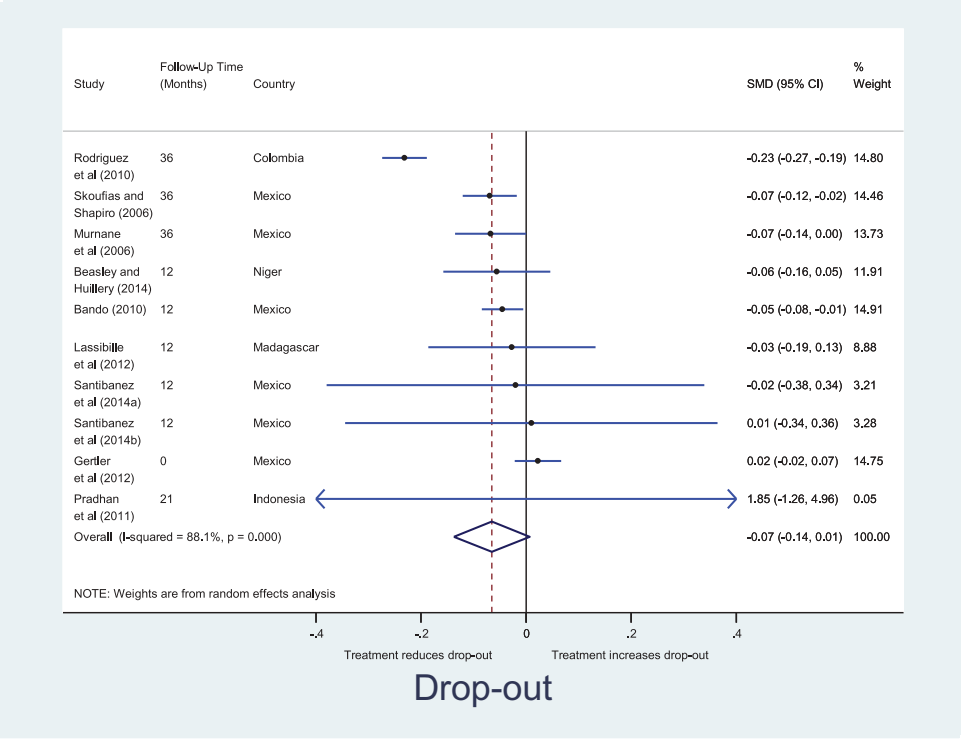

Figure 3. Effects on student dropout.

secondary level. A total of 9 studies (32%) used randomisation to assign participants to groups, while 17 (65%) use quasi-experimental approaches. Although the included studies represent a range of publication dates (from 1999 to 2014), all the studies using random allocation have been published since 2008.

Only six of the studies (23%) were published as articles in academic journals; the majority (N = 16, 62%) are World Bank reports or working papers published by economic think tanks. Three of the included studies were published as chapters in one World Bank publication. One is an unpublished PhD thesis. The implication of this is that about two-thirds of our included studies are reports which may never have been through an external peer review process. The risk of bias assessment indicated that eight studies (27%) could be classified as of low risk of bias overall. All of these studies were assessed as having used randomised assignment appropriately and we were not able to identify any sources of bias relating to factors such as method of allocation, attrition, contamination, motivation bias or biases in analysis reporting. Most other studies (63%), including three randomised control trials (RCTs)s, were classified as having moderate risk of bias, usually due to risks of confounding and/or contamination of comparison groups. As mentioned above, three studies (10%) were assessed as having high risk of bias and were excluded from the meta-analysis.

Results of overall synthesis

We were able to report on the impact of any school-based decision-making reform on six educational outcomes: (1) student dropout and attendance; (2) student repetition; (3) teacher attendance; and (4) student learning, as assessed through (i) language test scores, (ii) math test scores, (iii) aggregate test scores (that is, tests of more than one subject).

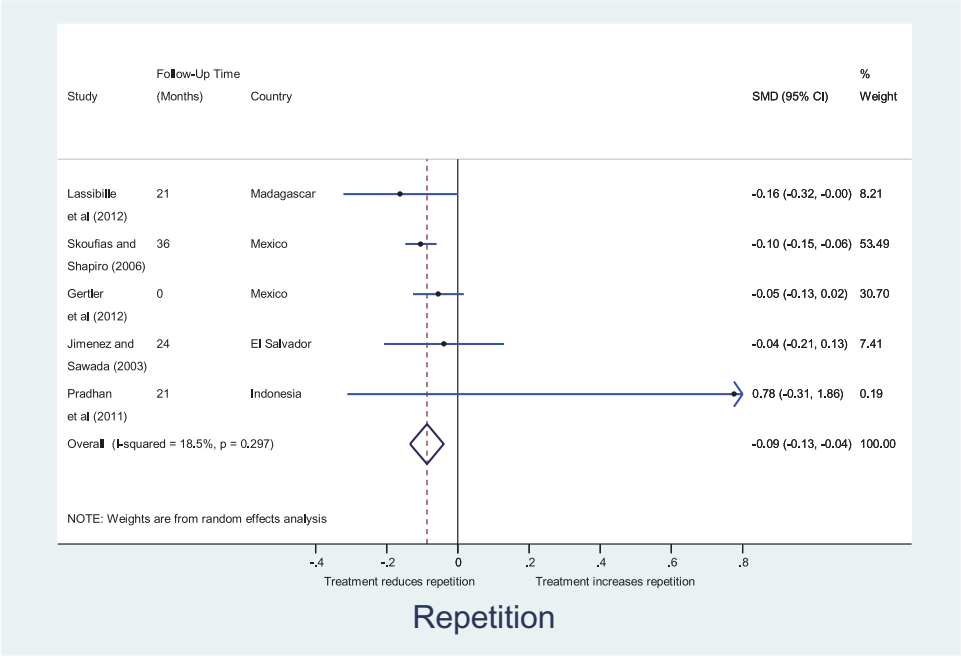

Figure 4. Effects on repetition.

Student drop out and attendance

Seven of the 10 estimates of effects of devolving decision-making to the level of the school on school-level student dropouts are from Latin America. All except two of the 10 estimates are negative and two in Colombia and Mexico are statistically significant (Figure 3). Pooling the findings across studies, we estimate a somewhat beneficial effect on school-level student dropouts – a pooled effect of reducing dropout by 0.07 SDs but not statistically significant at 95 per cent confidence (95% CI = −0.14, 0.01).

However, there is significant heterogeneity in the findings across studies (I-squared = 88%) and evidence in some contexts does suggest statistically significant reductions in dropouts. For example, Rodriguez, Sanchez and Armenta (2010) estimate the biggest reduction in Colombia (−0.23 SMD; 95% CI = −0.27,−0.19).

Six also reported effects on student absenteeism or attendance (Barr et al. 2012; Blimpo and Evans 2011; Di Gropello and Marshall 2005; Jimenez and Sawada 1999; Lassibille et al. 2010; and Sawada and Ragatz 2005). However, none included sufficient data to allow for the calculation of SMDs and subsequent pooling in meta-analysis. Two studies measure absenteeism by collecting data on student attendance on the day of an unannounced visit to a school. Barr et al. (2012) estimate that using a participatory process for developing and using a school report card increased attendance by up to 10 per cent, while Blimpo and Evans (2011) estimate that the Whole School Development intervention reduced student absenteeism by about 5 percentage points from a base of about 23 per cent. However, Jimenez and Sawada (2003) and Sawada and Ragatz (2005), who define absenteeism as the number of days absent in the previous month among students in the 3rd grade, find no difference between EDUCO and traditional schools in overall mean absence. Di Gropello and Marshall (2005), who use a student-reported ordinal measure of attendance, find no evidence that PROHECO schools succeeded in reducing student absences. Lassibille et al. (2010), meanwhile, measure attendance across a given school during the month prior to a visit identify an increase in attendance of approximately 4 percentage points over control schools in schools which

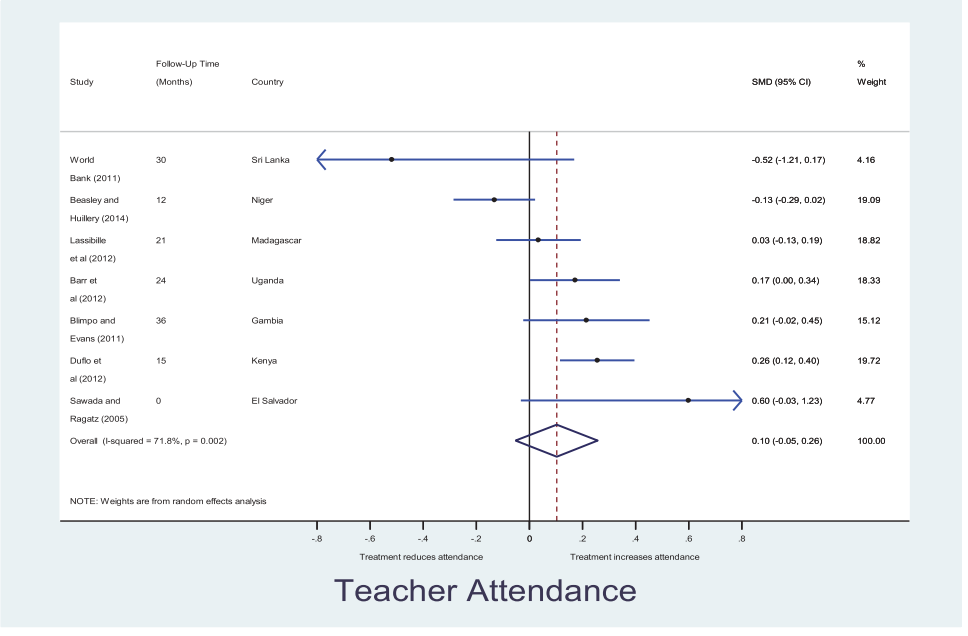

Figure 5. Effects on teacher attendance.

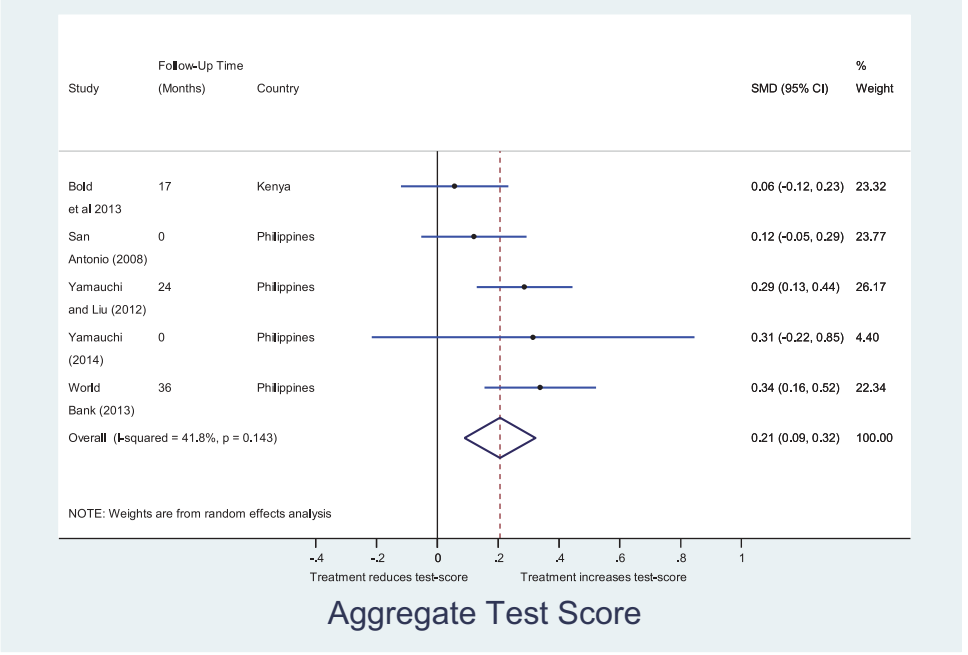

Figure 6. Effects on aggregate test scores.

benefited from interventions at the school level but no effect in districts implementing only the sub-district- and district-level version of the intervention.

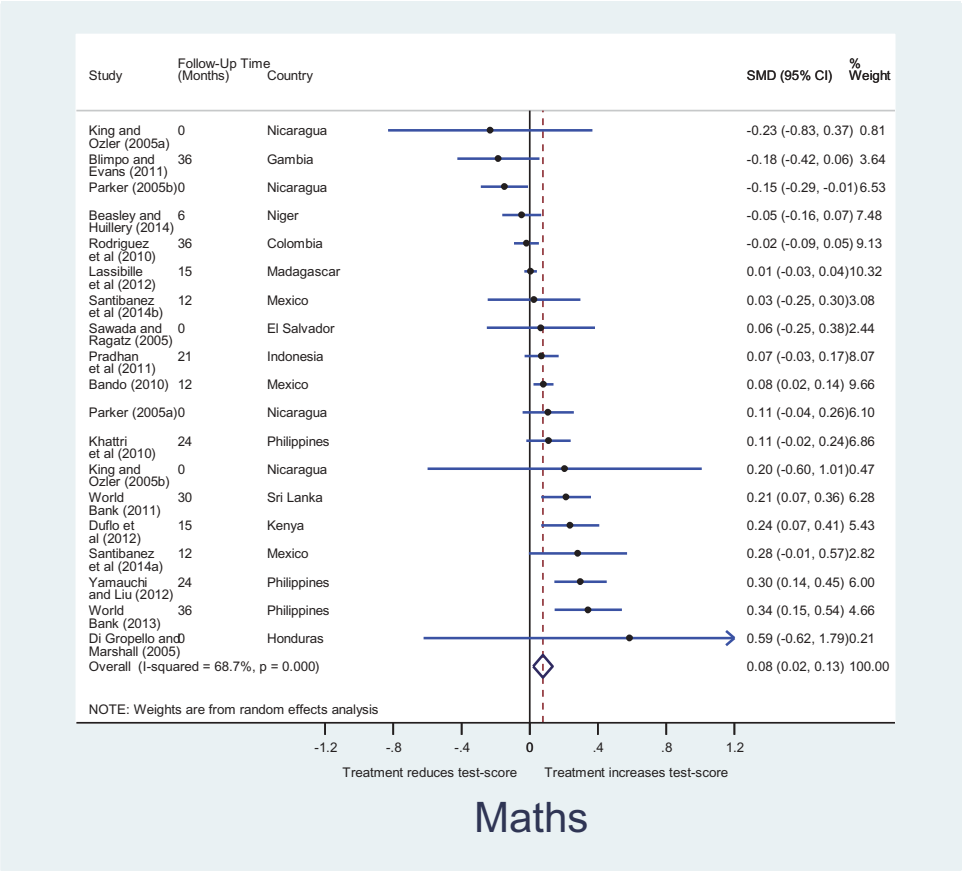

Figure 7. Effects on mathematics test score.

Student repetition, failure and progression

For repetition, the pooled effect of the impact of a school-based decision-making intervention is a reduction in school-level repetition rates of 0.09 SDs (95%CI = −0.13, −0.04) (Figure 4). Three of the five estimates are from Latin America, one is from Madagascar and one from Indonesia. All but one of the individual study estimates are negative, while only two in Madagascar and Mexico are significant at the 95 per cent level. Analysis does not suggest heterogeneity is significant across studies (I-squared = 18%), suggesting the findings are consistent across contexts.

Five studies also investigated impacts on student failure rates all in Mexico (Bando 2010; Gertler, Patrinos, and Rubio-Codina 2012; Murnane, Willett, and Cardenas 2006; Rodriguez, Sanchez, and Armenta 2010; Skoufias and Shapiro 2006). However, in none of these studies is failure precisely defined, in terms of which subjects are included in the assessment of a student's failure at the end of a year.

Gertler, Patrinos, and Rubio-Codina (2012) estimate a significant reduction in grade failure for AGE, a finding which is robust to checks on pre-intervention trends between treatment and comparison schools. Rodriguez, Sanchez, and Armenta (2010) also identify a reduction of 1.4 percentage points in PER schools over control schools. Three studies examine failure rates for the programme succeeding AGE, PEC (Murnane, Willett, and Cardenas 2006; Skoufias and Shapiro 2006; Bando 2010), finding mixed results. Skoufias and Shapiro (2006) estimated participation in PEC to reduce failure rates by 0.24 percentage points, while Murnane, Willett, and Cardenas (2006)

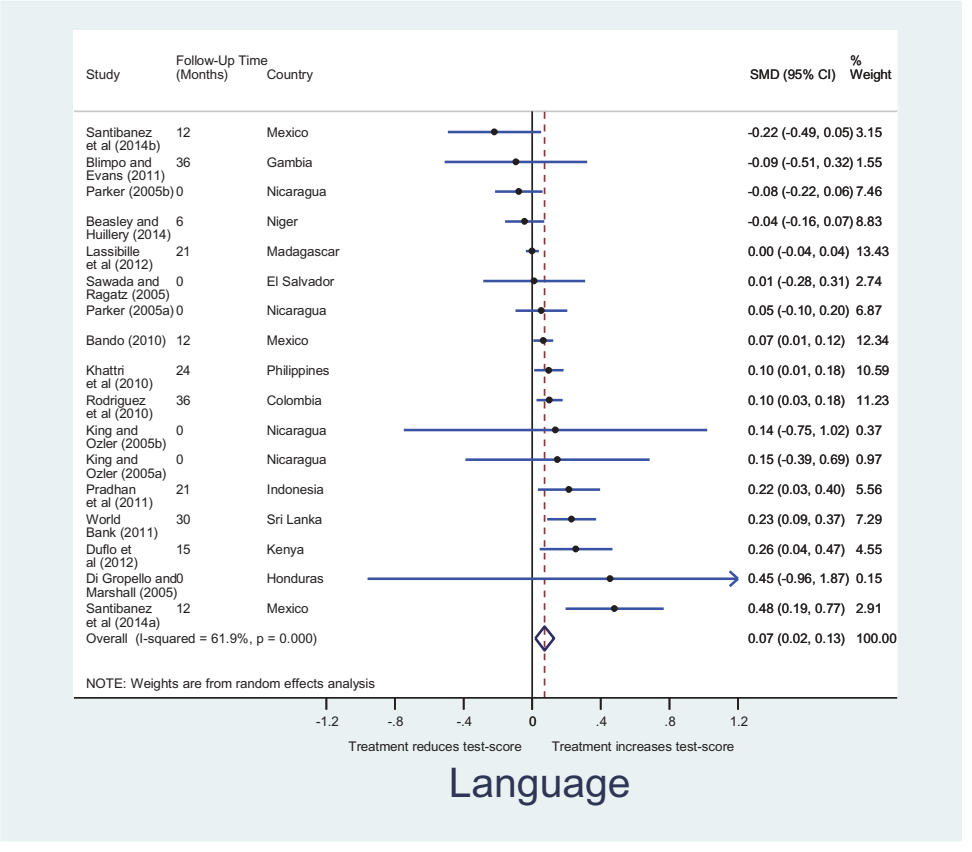

Figure 8. Effects on language test score.

argue PEC schools were more successful than control schools in retaining many students. On the other hand, Bando (2010), using census data in her analysis, suggests a positive association with failure rates that strengthens over time.

Two studies also investigated impacts on student progression and/or continuation (Barr et al. 2012; Jimenez and Sawada 2003), offering discrepant findings. Barr et al. (2012) found no impact on the probability of continued enrolment, as a result of the participatory scorecard intervention. However, Jimenez and Sawada (2003) identify an increase in continuation in EDUCO schools than others.

Teacher attendance

Figure 5 reports results from seven studies that measure the impact of a school-based decision-making intervention on teacher attendance. Five estimates are from Africa and the other two from Latin America and Asia. Evidence suggests effects on teacher attendance are positive overall, at 0.10 SD, but not statistically significant (95% CI = −0.05, 0.26). Analysis suggests there is significant heterogeneity in the estimates (I-squared = 72%). Indeed, two studies in Kenya and Uganda found significantly positive effects on teacher attendance.

Table 2. Analysis of effects by level of decentralisation and income context.

	Moderator variable	Pooled SMD	95% CI		I-squared (%)	Num obs.
Student drop outs						
	Low decentralisation	−0.23	−0.27	−0.19	n/a	1
	Medium decentralisation	−0.04	−0.07	−0.00	27.4	9
	High decentralisation	–	–	–	–	0
	Low income	−0.05	−0.13	0.04	0.0	2
	Lower middle-income	0.21	−1.48	1.87	42.0	2
	Upper middle-income	−0.04	−0.07	0.00	46.9	6
Teacher attendance						
	Low decentralisation	–	–	–	–	0
	Medium decentralisation	0.03	−0.13	0.20	65.8	5
	High decentralisation	0.28	0.10	0.48	7.8	2
	Low income	0.10	−0.04	0.25	74.6	5
	Lower middle-income	0.05	−1.05	1.15	81.8	2
	Upper middle-income	–	–	–	–	0
Math test scores						
	Low decentralisation	–	–	–	–	0
	Medium decentralisation	0.10	0.03	0.17	74.9	11
	High decentralisation	0.05	−0.11	0.22	59.3	7
	Low income	0.01	−0.09	0.11	55.1	6
	Lower middle-income	0.11	0.02	0.20	75.0	10
	Upper middle-income	0.09	0.03	0.14	0.5	3
Language test scores						
	Low decentralisation	0.10	0.03	0.18	n/a	1
	Medium decentralisation	0.08	0.00	0.15	75.1	9
	High decentralisation	0.05	−0.06	0.16	18.4	7
	Low income	0.02	−0.06	0.09	25.6	6
	Lower middle-income	0.09	0.03	0.16	41.5	8
	Upper middle-income	0.10	−0.19	0.40	83.7	3

Test scores

Effects on test scores are larger and more robust. We find a positive and significant improvement of 0.21 SDs in aggregate test scores[11] on average (95% CI = 0.09, 0.32) (Figure 6). The five estimates of aggregate test scores come from two countries (one from Kenya and four from the Philippines, all of which use the same test data). Two are positive and significant (both in the Philippines) with SMD around 0.3, and none is negative and significant.

We also find positive and significant average improvements of around 0.08 SD in scores on mathematics (95% CI = 0.02, 0.13) (Figure 7) and language (SMD = 0.07, 95% CI = 0.02, 0.13) (Figure 8). The 19 estimates for mathematics tests come from a range of contexts (Africa, Asia and Latin America). Only one estimate is negative and significant, while five, from a variety of contexts, are positive and significant – SMD exceeds 0.2 in Sri Lanka, Kenya and the Philippines. There is significant heterogeneity in effects (I-squared = 69%) suggesting further moderator analysis is needed to explain differences between studies. The 17 estimates for language tests[12] are from Asia, Africa and Latin America. Six of the 17 estimates are positive and significant, with SMD exceeding 0.2 in Indonesia, Kenya, Sri Lanka and in one Mexico study, while none is negative and significant. The analysis suggests significant residual heterogeneity (I-squared = 62%), which is explored further in moderator analysis.

Adverse outcomes

Devolving decisions to the level of the school can have negative consequences, such as elite capture and disharmony between ethnic groups. Two impact evaluations reported unintended consequences of school-based decision-making reforms. Murnane, Willett, and Cardenas (2006) identified a significant increase in the administrative burden on schools as a result of the PEC programme in Mexico. Duflo, Dupas, and Kremer (2012) note that school management committees in Kenya seemed to be more likely to hire male teachers than females.

Results of moderator and sub-group analyses

We conducted moderator analysis for variables which we believed are likely to affect the impact of school-based decision-making reforms: the level of decentralisation (high, medium or low) and the country income level (Table 2).[13] The results do not suggest consistent differences across outcomes for decentralisation categories, possibly due to the small numbers of observations by moderator variable groups.

The cross-country study by Hanushek, Link, and Woessmann (2011) found that the impact of school autonomy depends on the level of development of the country implementing the reform. Our own moderator analysis does not suggest effects differ by income group for dropouts and teacher attendance, despite evidence for the latter being dominated by studies from low-income countries where issues relating to teacher attendance may be particularly acute. However, analysis of test scores does suggest impacts pertaining to middle-income countries for mathematics and language, where the overall positive pooled effect is driven by the results for lower-middle income countries (0.09 SMD; 95% CI = 0.03, 0.16). By contrast, and with the exception of one study in Kenya (Duflo, Dupas, and Kremer (2012), now a middle-income country, there is no significant improvement in student learning in low-income country settings (SMD = 0.01 SMD; 95% CI = −0.09, 0.11).

We examined the possibility impact heterogeneity depending on the length of exposure to the reforms under investigation. Evidence from the US suggests that there can be a time lag of up to eight years between the implementation of a school-based management model and any observable impact on student test scores, although intermediate effects may be more rapidly identifiable (World Bank 2007, p. 13). This could be because schools initially see a decline in performance as school personnel adapt to the new structures, or because school-based management reforms are likely to have a more

Table 3. Meta-regression analysis of test score outcomes.

	Coefficient	Standard error	*p*-Value	95% CI	
Follow-up period in months	0.0145	0.0060	0.021	0.0024	0.0267
Follow-up in months (squared)	−0.0003	0.0002	0.049	−0.0007	−0.0000
=1 if low-income country	−0.0937	0.0485	0.062	−0.1925	0.0051
Constant	−0.0031	0.0477	0.949	−0.1003	0.0941

Observations = 36; I-squared = 52.7 per cent; adjusted R^2 = 0.378; Model F(3,32) = 3.18 (p = 0.037).

immediate impact on proximal outcomes (for example, teacher attendance), which then have a more gradual impact on student learning over time. Seven of the studies do explicitly include time-lag in heterogeneity analysis, and present inconsistent evidence. Some studies (for example, Duflo, Dupas, and Kremer 2012; Gertler, Patrinos, and Rubio-Codina 2012; Jimenez and Sawada 1999; and Santibañez, Abreu-Lastra, and O'Donoghue 2014) identify a possible 'Hawthorne effect', whereby schools show positive results in the first year (possibly due to the energy and momentum created by the new reform), which do not continue to increase with prolonged exposure. A similar effect is identified by Khattri, Ling., and Jha (2010) and Yamauchi (2014), although neither study explicitly presents data on this point. However, other studies (for example, Bando 2010; King and Ozler 2005; Murnane, Willett, and Cardenas 2006) identify stronger results in communities with longer exposure to the intervention. As studies in both groups examine similar outcomes, it is difficult to draw any conclusions around the differential impact of length of exposure.

We also explored the correlation between the time-lag between the start of the intervention and the impacts observed for test scores, where there were sufficient observations to examine variation by follow-up time, using meta-regression analysis (Table 3). The meta-regression also conditions on country income status and suggests that it may take on average 22 months for reforms to have a maximum effect of 0.16 SD in middle-income country contexts. This suggests results from shorter term follow-ups may be biased downward, if we are to believe indirect treatment comparisons across contexts.

Almost half of the studies, covering 14 interventions, examined impact heterogeneity according to the following factors:

- *student-level factors* including baseline ability (Pradhan et al. 2011), sex (Pradhan et al. 2011), socio-economic background (Rodriguez, Sanchez, and Armenta 2010) and grade level (Beasley and Huillery 2014; Gertler, Patrinos, and Rubio-Codina 2012; King and Ozler 2005; Parker 2005; Rodriguez, Sanchez, and Armenta 2010; Santibañez, Abreu-Lastra, and O'Donoghue 2014);
- *school-level factors* including school size (Beasley and Huillery 2014; King and Ozler 2005) and characteristics of teachers (Glewwe and Maïga 2011; Barr et al. 2012; Jimenez and Sawada 2003; Duflo, Dupas, and Kremer 2012) and head teachers (Rodriguez, Sanchez, and Armenta 2010);
- *community-level factors* including the level of community disadvantage (Gertler, Patrinos, and Rubio-Codina 2012; Murnane, Willett, and Cardenas 2006; Rodriguez, Sanchez, and Armenta 2010; Skoufias and Shapiro 2006), education levels of parents and school management committee members (Beasley and Huillery 2014; Blimpo and Evans 2011) and the level of community participation (Jimenez and Sawada 1999; King and Ozler 2005);
- *national-level factors* including one study examining a sub-group of teachers under a centralised pay-for-performance scheme that rewarded teachers for strong results on student assessments (Gertler, Patrinos, and Rubio-Codina 2012) and
- *co-interventions* such as training (Blimpo and Evans 2011; Bold et al. 2013; Duflo, Dupas, and Kremer 2012; Pradhan et al. 2011), accountability mechanisms like report cards (World Bank 2011; Barr et al. 2012), the election of school management committees (Pradhan et al. 2011) and whether the implementing body is government or NGO (Bold et al. 2013; Glewwe and Maïga 2011; Lassibille et al. 2010).[14]

Implementation fidelity was also discussed to a very limited extent (Pradhan et al. 2011; Blimpo and Evans 2011; Bold et al. 2013; Yamauchi 2014) but not in formal sub-group analysis.

The findings from these analyses varied across contexts, and are reported in full elsewhere (Carr-Hill, Rolleston, and Schendel 2016). Briefly, there is some evidence to suggest that school-based decision-making reforms have a stronger impact on wealthier students with more educated parents. It also appears that reforms may be particularly effective for lower grade levels. By contrast, reforms appear to be less effective in disadvantaged communities, particularly if parents

and community members have low levels of education and low status relative to school personnel. This is a particularly important result, given that some studies showing positive impacts explicitly acknowledged having avoided including more remote areas in their analysis (for example, Glewwe and Maïga 2011; and Lassibille et al. 2010). Devolution also appears to be ineffective when communities are not able to participate actively in decision-making processes. Small schools, however, may find school-based decision-making to be effective, particularly if community members establish a collaborative, rather than an adversarial, relationship with teachers. Two studies in particular (Jimenez and Sawada 1999; King and Ozler 2005), both investigating programmes in Latin America, conclude that community participation levels are critical.

Barriers to and enablers of effective school-based decision making

This section draws on findings from impact evaluations and the linked studies, which were critically appraised prior to synthesis (Table A1). The finding that devolving decisions to the school level does not have a positive effect on the poorest, most disadvantaged communities, is supported by qualitative evidence from Nicaragua. Fuller and Rivarola (1998) found that schools in severely impoverished areas were unlikely to raise additional revenue from the surrounding communities. Gershberg and Meade (2005) found parental contributions to be a significant component of autonomous school budgets, suggesting that disadvantaged communities would be unable to raise sufficient resources under the autonomous schools model.

Low levels of capacity within communities are also barriers. Communities with high levels of illiteracy and/or with few educated parents do not seem to benefit from devolution of decisions to the community level. In their study of Whole School Development programme in the Gambia, Blimpo and Evans (2011) argue that devolution may even be detrimental in such contexts 'because the communities are not well equipped to act on [such information]' (p. 29).

The cross-country study by Hanushek, Link, and Woessmann (2011) finds autonomy reforms improve student achievement in more developed countries but may undermine it in less developed ones. Reimers and Cardenas (2007), in analysis of Mexico's PEC programme, find that lack of leadership or 'coherence of vision among school staff' are significant barriers (p. 38). Teachers in Indonesia felt they did not have the capacity to implement the curricular component of that country's school-based management reform points, nor did they feel adequately supported to use the autonomy given to them (Bjork 2003).

There are a variety of reasons why the capacity of institutions and communities can act as a barrier to effective school-based decision-making reforms. First, for such reforms to be effective, school personnel and community members must understand the nature of the reform and crucially must also be able to propose changes that are likely to affect student learning within the school. Santibañez, Abreu-Lastra, and O'Donoghue (2014) and Parker (2005) note that communities in Mexico and Nicaragua did not always fully grasp the nature and the objective of school-based decision-making reforms. Bandur (2008) raises similar concerns in analysis of the national school-based management reform in Indonesia. In the Nicaraguan context, the lack of ownership led to active resistance in some communities (Fuller and Rivarola 1998).

Beasley and Huillery (2014) found that school management committees in rural communities in Niger frequently opted to spend their grants on agricultural projects, instead of school materials, teacher incentives or other initiatives likely to affect educational outcomes. Secondly, community members – particularly parents – must have a certain amount of status to play an active role on school management committees. This does not tend to be the situation in rural, poor communities, where school personnel are often perceived as authority figures due to their relatively high levels of education (Beasley and Huillery 2014; Gertler, Patrinos, and Rubio-Codina 2012). All of these reasons may explain why early interventions devolving decisions to the school level, such as EDUCO in El Salvador, restricted participation in school management decisions to literate members

of the community, a requirement which does not appear to feature in similar models of school-based management implemented more recently in other low-income contexts.

Under-resourced governments may simply be unable to implement and monitor complex decentralisation reforms. Bold et al. (2013) finds that a contract teacher programme, which was effective under NGO implementation, had no effect when scaled up by the government at the national level. Lassibille et al. (2010) and Glewwe and Maïga (2011) find impacts among schools benefiting from direct training by NGO representatives in Madagascar, but not among schools that had been trained by district or sub-district employees (who had themselves been trained by the NGO). National-level programmes unlikely to be effective without sufficient monitoring capacity and accountability mechanisms, both of which are often limited in low-income contexts. Indeed, there may be reason to suspect that government officials may actively hinder the effectiveness of school-based management reforms, as was identified by both Bandur (2008) and Vernez, Karam, and Marshall (2012) in Indonesia, where provincial and district officials were found to actively interfere in school decision-making processes.

Finally, the studies highlight the fact that school-based decision-making reforms can only affect the immediate circumstances of a given school or community. Even if a reform is effective within a community, school-based management reforms cannot address many external factors that can act as significant barriers to impact. Although there are myriad external factors affecting educational outcomes, the included studies reference five that appear most relevant:

- The strength of the national teacher's union (Bold et al. 2013);
- The strength of the teacher job market (Barr et al. 2012; Parker 2005);
- Teacher ability/quality (Lassibille et al. 2010; Blimpo and Evans 2011);
- Constraints imposed by the central system, for example, inefficient mechanisms for distributing salaries in rural areas (Blimpo and Evans 2011; Lassibille et al. 2010) and
- Security (Beasley and Huillery 2014).

Studies also point to several enablers of effective school-based decision-making reforms. First, it appears that smaller schools (for example, one-teacher schools as in Beasley and Huillery 2014) are more likely to benefit from local decision-making authority, because it is easier for school management committees to monitor teachers and stay informed about conditions at the school. Second, devolving personnel decisions and financial and other management decisions enables school-based decision-makers to affect teacher behaviour, including attendance (Sawada and Ragatz 2005; King and Ozler 2005). Finally, Duflo, Dupas, and Kremer (2012) suggest that giving parents the majority voting power on school management committees in Kenya was one of the reasons why local hiring addressed issues of elite capture.

Discussion

We report the first mixed-methods systematic review of school-based management. The included studies were systematically critical appraised and met the criteria for relatively low risk of bias.

Our findings are broadly similar to other comprehensive reviews of evidence (for example, Santibanez 2007; World Bank 2007). However, we offer a body of evidence substantially bigger in size and geographic breadth than these reviews, hence adding to the generalisability of the evidence. Our review includes 26 impact studies and 9 qualitative studies, representing 17 distinct interventions in 13 countries across Latin America (5 countries), sub-Saharan Africa (5 countries) and South/Southeast Asia (3 countries).

Overall, we find that devolving decision-making to the level of the school appears to have a positive effect on repetition and beneficial effects on reducing dropouts and improving teacher attendance in certain contexts. Effects on test scores are more robust, and range between 0.10 and 0.20 SD. In comparative terms, these effects may be considered sizeable when compared to the

balance of results for educational interventions, not least because effect sizes in the field of education tend to be relatively small and effect sizes approaching 0.2 SMD are comparatively large (Snilstveit et al., 2016). In broader terms, reported effects on learning outcomes vary widely but are often small and/or statistically non-significant. Kremer, Brannen, and Glennerster (2013) review a number of RCTs which employ test scores as outcomes and find that, very exceptionally, effect sizes can be as high as 0.6 SD (providing village schools in Afghanistan), while more generally an effect 0.2 SD can be considered large.

There is also much heterogeneity in effects. Evidence suggests that school-based decision-making reforms appear to be less effective in disadvantaged communities, particularly if parents and community members have low levels of education and low status relative to school personnel.

In models of school-based decision-making classified as 'high' decentralisation, schools and communities have decision-making authority over nearly all aspects of school management. Most importantly, the school (or, typically, the school management committee) has authority over both financial and personnel decisions, including the authority to hire/fire teachers and to pay salaries. As is evident from the studies examining the impact of differential levels of participation on outcomes, devolving decision-making to school level does not always result in increased stakeholder participation in school activities. However, when participation does increase – and when school management committees have the authority to hire and fire teachers – the evidence suggests that teacher attendance improves (Figure A1). We know less about how this may translate into student learning. In fact, improved teacher attendance does not appear to result in increased teacher effort or improved quality of teaching in many contexts. The link between teacher attendance and student learning is likely to depend on several other external factors, including teacher ability, community characteristics and the specific design of the school-based decision-making reform.

In 'medium' decentralisation models, schools have authority over non-personnel financial decisions. This authority usually comprises oversight of grants related to School Improvement Plans and/or the school budget, as well as legal authority to raise independent monies on behalf of the school. There is evidence to suggest that devolving financial decisions to the school level often results in an increased amount of money available to the school, either due to the receipt of a grant or to the fundraising activities of school management committees. However, increased funding does not appear to translate into educational outcomes, particularly in poorer communities (Figure A2).

Implications for policy, practice and research

Our findings carry several implications for policy and practice. First, it appears that school-based decision-making reforms in highly disadvantaged communities are less likely to be successful. Parental participation seems the key to the success of such reforms and this is linked to the real authority or status and cultural capital of community members. One benchmark, proposed by Blimpo and Evans (2011), is that communities need a minimum of 45 per cent overall literacy in order to benefit from school-based management. This suggests that policymakers are likely to see greater impact of school-management reforms in more advantaged areas, although this raises obvious equity concerns.

Second, the involvement of school management committees in personnel decisions appears to play a role in improving proximal outcomes, such as teacher attendance. However, the impact of devolving personnel decisions is also likely to be linked to the overall teacher job market and the possibility of long-term employment. Policy proposals should therefore consider the current and prospective job market conditions for teachers when anticipating the potential impact of school-based decision-making reforms.

Third, the specifics of programme design appear to be crucial. It appears that the details of such supplementary elements (for example, restrictions on the use of grants; the implementing body responsible for training; and so forth) may play an important enabling role. The evidence also suggests that, at least in some contexts, impact on student learning may take longer than is often

allowed within evaluation timelines. Where donors are involved, this also means that decentralisation reforms may require sustained donor commitment over the longer term (minimum 2 years).

Finally, we suggest policymakers should proceed with caution when using the results from small-scale pilot programmes to inform national programming, although further research is needed on this point.

The review also suggests a number of fruitful directions for future research. There needs to be further robust analysis of the impacts of large-scale school-based decision-making reforms that have recently been implemented. More studies are also needed that analyse the relative impacts of different kinds of school-based decision-making interventions (that is, implementation RCTs with active controls). The few studies with active controls (for example, Pradhan et al. 2011) offer important insights into the specific effects of different models, which should be replicated elsewhere.

We were unable to locate many studies investigating possible negative or unintended consequences of school-based decision-making reforms, given that such outcomes do not feature explicitly in any of the included impact studies. There is therefore a clear need to examine negative effects, given widespread adoption of such policies, in impact evaluations. But it would also be possible to incorporate adverse effects drawing on non-experimental studies. A future review of school-based decision-making could expand the inclusion criteria to examine adverse effects by incorporating the full range of non-experimental and qualitative evidence. More generally, we have identified a large amount of qualitative evidence which could also be used to synthesise a broader range of barriers and enablers of implementation to complement the findings of this study.

This review excluded reforms which evaluated interventions designed to improve the functioning of existing school-based decision-making mechanisms, and studies of interventions designed by agencies external to the school (for example, donor agencies, NGOs). A future review could include such studies. Finally, a review of evidence on cost-effectiveness would also be warranted although this would presumably need to incorporate relevant programme documentation to identify unit costs, since the studies we located did not provide such information.

Notes

1. A recent paper by Evans and Popova (2015) argues that divergent conclusions from systematic reviews tend to be driven by a reliance on different samples of research studies, which, in turn, are driven by differing criteria for inclusion. However, the sample of studies included in that review of reviews largely draws on studies which do not use systematic methods of search, appraisal or synthesis.
2. Income classifications reflect the World Bank's income classification system. Classifications were linked to the start date of the intervention under investigation, rather than the current classification.
3. Studies written in other languages were excluded, unless English translations were available, as we did not have any further linguistic ability represented within the review team.
4. We developed a risk of bias assessment tool based on 'Suggested risk of bias criteria for EPOC reviews' (Cochrane EPOC, 2014), with additional questions suggested by Hombrados and Waddington (2012) .
5. As existing systematic reviews (for example, Petrosino et al, 2012) have indicated a lack of relevant studies on education decentralisation in developing countries published prior to 2000, we limited our electronic searches to studies published in or after 2000. We did set any such data boundary for our other search methods (for example, review of reviews).
6. We were unable to complete forward citation chasing of included studies.
7. An additional four studies were identified through reference searching and expert checking.
8. In two of the three studies (Paes de Barros & Mendonca, 1998; De Umanzor et al. 1997), we identified a substantial risk of confounding factors influencing the impact estimates, while there was a high risk of bias due to attrition in the final study (Cueto et al. 2008). Other risks were also identified, including risk of motivation bias and clustering, in one of the three studies (De Umanzor et al. 1997).
9. Carnoy et al. (2008) was excluded from meta-analysis due to missing data.
10. Comparisons of effect sizes measured in standard deviations are comparisons of relative measures, requiring, for example, assumptions concerning the distribution and measurement of a phenomenon or trait (for example, educational performance as measured by a test) in the samples to be compared. It was not possible in every case to calculate SMD, particularly for studies which did not report standard deviations of the outcome variable and/or the number of observations in the study or the statistics required to compute or estimate the standard deviation or other required statistic (for example, *t, z* or *F* statistics, *p*-values and

standard errors). However, we employed appropriate methods to generate comparable effect-sizes wherever possible, including using the Campbell Collaboration online effect size calculator http://www.campbellcollaboration.org/resources/effect_size_input.php.

11. Aggregated tests are multi-subject tests. The National Achievement Test in the Philippines comprises Math, English, Filipino, Science and Social Science. The test used by Bold et al. (2013) covers only Math and English.
12. Of the 14 studies that measured the impact of a school-based decision-making intervention on student language test scores, some reported test data for more than one language. The languages tested are usually the language of instruction in school, where available.
13. Results of moderator analyses by type of evaluation method used (with or without randomised assignment) and risk of bias assessment is available in the technical report (Carr-Hill, Rolleston, and Schendel 2016). The results for RCTs and quasi-experimental studies are similar overall nor could we identify any significant differences in the effects indicated by low and medium risk of bias studies.
14. In some instances, schools were given grants for explicit purposes, for example, the hiring of contract teachers (Blimpo and Evans 2011; Bold et al. 2013; Duflo, Dupas, and Kremer 2012). However, no study in the sample was able to estimate the marginal impact of allocating grants, because all studies included a grant component in treatment and control arms.

Notes on contributors

Professor Roy Carr-Hill has taught in Sussex University (1971-74), Universidade Eduardo Mondlane, Mozambique (1978-81) and the UCL Institute of Education (1993-ongoing); worked on social indicators at the OECD Paris (1974-77); and since 1981 has been a researcher at the MRC Medical Sociology Unit Aberdeen (1981-84); Centre for Health Economics (1984-2011), School of Political and Social Sciences, Hull (1990-93). He has worked in 35 developing countries as a consultant for several of the bilateral and multilateral agencies on a wide variety of evaluation and review consultancies, spanning, education, health and justice. He is currently based at UCL-IoE and researches on accountability, adult literacy, rationality of drop-out, exploitation of women teachers, interplay between education and extremism; and is currently writing book on understanding social data.

Dr. Caine Rolleston is Senior Lecturer in Education and International Development at UCL Institute of Education and Lead Education Researcher at the Young Lives Project, University of Oxford. His research focuses on analysis of education systems in low and middle income countries, including school quality and effectiveness and educational access and equity.

Dr. Rebecca Schendel is Lecturer in Education and International Development at the UCL Institute of Education. Her research focuses on the relationship between higher education and human development in sub-Saharan Africa, with particular emphasis on questions of pedagogy, student learning and processes of institutional change.

Hugh Waddington is Senior Evaluation Specialist and currently Acting Head of 3ie's Synthesis and Reviews Office. He has a background in research and policy, having worked previously in the Government of Rwanda, the UK National Audit Office and the World Bank, and before that with Save the Children UK and the Department for International Development. He is managing editor of the *Journal of Development Effectiveness* and co-chair of the International Development Coordinating Group (IDCG) of the Campbell Collaboration.

Acknowledgements

This review was supported by the Department for International Development's (DFID's) Systematic Reviews Programme. We acknowledge contributions to the review process by Tejendra Pherali, Edwina Peart and Emma Jones at UCL Institute of Education and Claire Stansfield and Carol Vigurs at the Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre).

Declaration of interest

None of the team members have any financial interests in the review, nor have any team members been involved in any other systematic review focused on this topic or in the development of any of the interventions investigated.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Caine Rolleston http://orcid.org/0000-0002-4564-4331

References

Included studies

Bando, R. (2010). The Effect of School Based Management on Parent Behavior and the Quality of Education in Mexico. Unpublished PhD thesis.

Bandur, A. (2008). A study of the implementation of school-based management in Flores primary schools in Indonesia. Unpublished PhD thesis.

Barr, A., L. Bategeka, M. Guloba, I. Kasirye, F. Mugisha, P. Serneels, and A. Zeitlin (2012). Management and Motivation in Ugandan Primary Schools: An Impact Evaluation Report. PEP Working Paper. Nairobi: Partnership for Economic Policy.

Beasley, E., and E. Huillery (2014). Willing but Unable: Short-Term Experimental Evidence on Parent Empowerment and School Quality. Unpublished manuscript.

Bjork, C. (2003). 'Local Responses to Decentralization Policy in Indonesia.' *Comparative Education Review*, 47 (2): 184–216. doi:10.1086/376540.

Blimpo, M., and D. K. Evans (2011). School-Based Management and Educational Outcomes: Lessons from a Randomized Field Experiment. Unpublished manuscript. http://siteresources.worldbank.org/EDUCATION/Resources/Blimpo-Evans_WSD-2012-01-12.pdf.

Bold, T., M. Kimenyi, G. Mwabu, A. Ng'ang'a, and J. Sandefur (2013). Scaling-Up What Works: Experimental Evidence on External Validity in Kenyan Education. CSAE Working Paper WPS/2013-04. Oxford: Centre for the Study of African Economies.

Carnoy, M., A. K. Gove, S. Loeb, J. H. Marshall, and M. Socias (2008) 'How Schools and Students Respond to School Improvement Programs: The Case of Brazil's PDE'. *Economics of Education Review* 27(1): 22–38 doi:10.1016/j.econedurev.2006.04.005.

Cueto, S., M. Torero, J. León and J. Deustua (2008). Asistencia docente y rendimiento escolar: el caso del programa META. Documento de Trabajo 53. Lima: GRADE.

De Umanzor, S., I. Soriano, M. R. Vega, E. Jimenez, L. Rawlings, and D. Steele (1997). El Salvador's EDUCO Program: A First Report on Parents' Participation in School- Based Management. Working Paper Series on Impact of Education Reforms, Paper No. 4. Washington, DC: World Bank.

Di Gropello, E., and J. H. Marshall (2005). 'Teacher Effort and Schooling Outcomes in Rural Honduras.' In E. Vegas (edited by), *Incentives to Improve Teaching*. Washington DC: World Bank, 307–358.

Duflo, E., P. Dupas, and M. Kremer (2012). School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools. NBER Working Paper No. 17939. Cambridge, MA: National Bureau of Economic Research.

Evans, D. K., and A. Popova (2015). What Really Works to Improve Learning in Developing Countries? *An Analysis of Divergent Findings in Systematic Reviews*. Policy Research Working Paper 7203. Washington, DC: World Bank.

Fuller, B., and M. Rivarola (1998). Nicaragua's Experiment to Decentralize Schools: Views of Parents, Teachers and Directors. Working Paper Series on Impact of Education Reforms, Paper No. 5. Washington, DC: World Bank.

Gershberg, A. I., and B. Meade (2005). 'Parental Contributions, School-Level Finances and Decentralization: An Analysis of Nicaraguan Autonomous School Budgets.' *Comparative Education*, 41 (3): 291–308. doi:10.1080/03050060500211658.

Gertler, P., H. A. Patrinos, and M. Rubio-Codina (2012). 'Empowering Parents to Improve Education: Evidence from Rural Mexico.' *Journal of Development Economics*, 99(1): 68–79. doi:10.1016/j.jdeveco.2011.09.004.

Glewwe, P., and E. Maïga (2011). The Impacts of School Management Reforms in Madagascar: Do the Impacts Vary by Teacher Type? Unpublished manuscript.

Gunnarsson, V., P. F. Orazem, M. A. Sanchez, and A. Verdisco (2008). Does Local School Control Raise Student Outcomes?: Theory and Evidence on the Roles of School Autonomy and Community Participation. Working Paper No. 09012. Ames, IA: Iowa State University.

Hanushek, E. A., S. Link, and L. Woessmann (2011). Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA. NBER Working Paper No. 17591. Washington, DC: National Bureau of Economic Research.

Jimenez, E., and Y. Sawada (1999). 'Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program.' *The World Bank Economic Review*, 13(3): 415–441. doi:10.1093/wber/13.3.415.

Jimenez, E., and Y. Sawada (2003). Does Community Management Help Keep Kids in Schools? Evidence Using Panel Data from El Salvador's EDUCO Program. CIRJE Discussion Paper F-236.

Khattri, N., C. Ling., and S. Jha (2010). The Effects of School-Based Management in the Philippines: An Initial Assessment Using Administrative Data. World Bank Policy Research Working Paper 5248. Washington, DC: World Bank.
King, E. M., and B. Ozler (2005). What's Decentralization Got to Do with Learning? 21COE Discussion Paper No. 54. Kyoto: Kyoto University.
Lassibille, G., J. P. Tan, C. Jesse, and T. Van Nguyen (2010). 'Managing for Results in Primary Education in Madagascar: Evaluating the Impact of Selected Workflow Interventions.' *World Bank Economic Review*, 24(2): 303–329. doi:10.1093/wber/lhq009.
Murnane, R. J., J. B. Willett, and S. Cardenas (2006). Did the Participation of Schools in Programa Escuelas de Calidad (PEC) Influence Student Outcomes? Unpublished manuscript.
Paes de Barros, E. and R. Mendonca (1998). The impact of three institutional innovations in Brazilian Education. In Organization matters: agency problems in health and education in Latin America, edited by W. D. Savedof. Washington, DC: Inter-American Development Bank.
Parker, C. E. (2005). 'Teacher Incentives and Student Achievement in Nicaraguan Autonomous Schools.' In E. Vegas (edited by), *Incentives to Improve Teaching*. Washington DC: World Bank, 359–388.
Pradhan, M., D. Suryadarma, A. Beatty, M. Wong, A. Alishjabana, A. Gaduh, and R. Prama Artha (2011). Improving Educational Quality through Enhancing Community Participation: Results from a Randomized Field Experiment in Indonesia. World Bank Policy Research Working Paper 5795. Washington, DC: World Bank.
Reimers, F., and S. Cardenas (2007). 'Who Benefits from School-Based Management in Mexico?' *Prospects: Quarterly Review of Comparative Education*, 37 (1): 37–56. doi:10.1007/s11125-007-9015-0.
Rodriguez, C., F. Sanchez, and A. Armenta (2010). 'Do Interventions at School Level Improve Educational Outcomes? Evidence from a Rural Program in Colombia.' *World Development*, 38(3): 415–428. doi:10.1016/j.worlddev.2009.10.002.
San Antonio, D. M. (2008). 'Creating Better Schools through Democratic School Leadership.' *International Journal of Leadership in Education*, 11(1): 43–62. doi:10.1080/13603120601174311.
Santibañez, L., R. Abreu-Lastra, and J. O'Donoghue (2014). 'School Based Management Effects: Resources or Governance Change? Evidence from Mexico.' *Economics of Education Review*, 39: 97–109. doi:10.1016/j.econedurev.2013.11.008.
Sawada, Y., and A. B. Ragatz (2005). 'Decentralization of Education, Teacher Effort, and Educational Outcomes.' In E. Vegas (edited by), *Incentives to Improve Teaching*. Washington DC: World Bank, pages 255–306.
Skoufias, E., and J. Shapiro (2006). Evaluating the Impact of Mexico's Quality Schools Program: The Pitfalls of Using Non-Experimental Data. Impact Evaluation Series No. 8. Washington, DC: World Bank.
Vernez, G., R. Karam, and J. H. Marshall (2012). Implementation of School-Based Management in Indonesia. *Monograph*. Santa Monica, CA: RAND Corporation. http://www.rand.org/pubs/monographs/MG1229.html
World Bank. (2011). An Impact Evaluation of Sri Lanka's Policies to Improve the Performance of Schools and Primary School Students through Its School Improvement and School Report Card Programs. South Asia Human Development Unit Report No. 35. Washington, DC: World Bank.
World Bank. (2013). Republic of the Philippines Basic Education Public Expenditure Review Phase 2 School Based Management in the Philippines: An Empirical Investigation. East Asia and Pacific Unit Report No: ACS2212.
Yamauchi, F. (2014). An Alternative Estimate of School-Based Management Impacts on Students' Achievements: Evidence from the Philippines. World Bank Policy Research Working Paper 6747. Washington, DC: World Bank.
Yamauchi, F., and Y. Liu (2012). Impacts of an Early Stage Education Intervention on Students' Learning Achievement: Evidence from the Philippines. World Bank Policy Research Working Paper 6246; Impact Evaluation Series No. 72. Washington, DC: World Bank.

Supporting Literature

Anderson, L. M., M. Petticrew, E. Rehfuess, R. Armstrong, E. Ueffing, P. Baker, D. Francis, and P. Tugwell (2011). Using Logic Models to Capture Complexity in Systematic Reviews. *Research Synthesis Methods* 2 (1), 33–42. DOI: 10.1002/jrsm.32.
Banerjee, A. V., R. Banerji, E. Duflo, R. Glennerster, and S. Khemani (2008). *Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India. Policy Research Working Paper 4584*. Washington, DC: World Bank. DOI: 10.1596/1813-9450-4584.
Bardhan, P., and D. Mookherjee (2000). Capture and Governance at Local and National Levels. *American Economic Review* 90 (2), 135–139. DOI: 10.1257/aer.90.2.135.
Bardhan, P., and D. Mookherjee (2005). Decentralizing Antipoverty Program Delivery in Developing Countries. *Journal of Public Economics* 89 (4), 675–704. DOI: 10.1016/j.jpubeco.2003.01.001.
Barrera-Osorio, F., and L. Linden (2009). The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia. Policy ResearchWorking Paper (Impact Evaluation Series) 4836 (29). Washington, DC: World Bank.
Barrera-Osorio, F., T. Fasih, H. A. Patrinos, and L. Santibanez (2009). *Decentralized Decision-Making in Schools: The Theory and Evidence on School-Based Management*. Washington, DC: World Bank.
Bruns, B., D. Filmer, and H. A. Patrinos (2012). *Making Schools Work: New Evidence on Accountability Reforms*. Washington, DC: World Bank. DOI: 10.1596/978-0-8213-8679-8.

Carr-Hill, R., C. Rolleston, and R. Schendel (2016). The Effects of School-Based Decision Making on Educational Outcomes in Low and Middle Income Contexts: A Systematic Review. *Campbell Systematic Reviews*, 2016:9. 2016.9. Oslo: The Campbell Collaboration. DOI: 10.4073/csr.

Carr-Hill, R., C. Rolleston, T. Pherali, and R. Schendel (2014). *The Effects of School-Based Decision Making on Educational Outcomes in Low and Middle Income Contexts: A Systematic Review (Protocol)*.

Carr-Hill, R., M. Hopkins, J. Lintott, and A. Riddell (1999). Monitoring the Performance of Educational Programmes in Developing Countries. DFID Education Research Paper No. 37. London: DFID.

Cochrane Effective Practice and Organisation of Care Group. (2014). *Suggested Risk of Bias Criteria for EPOC Reviews*. Ottawa: EPOC.

Condy, A. (1998). *Improving the Quality of Teaching and Learning through Community Participation: Achievements, Limitations and Risks*. SDD/DFID Working Paper. London: DFID.

De Grauwe, A., C. Lugaz, D. Balde, C. Diakhate, D. Dougnon, M. Moustapha, and D. Odushina (2005). *Does Decentralization Lead to School Improvement? Findings and Lessons from Research in West Africa*. Paris: IIEP.

DFID. (2014). *How to Note: Assessing the Strength of Evidence*. London: DFID.

Gertler, P., H. A. Patrinos, and M. Rubio-Codina (2008). *Impact Evaluation for School-based Management Reform. Policy Research Working Paper. Doing Impact Evaluation Series 10*. Washington, DC: World Bank.

Glassman, D., J. Naidoo, and F. Wood (2007). *Community Schools in Africa: Reaching the Unreached*. New York, NY: Springer.

Glewwe, P., N. Ilias, and M. Kremer (2003). *Teacher Incentives. NBER Working Paper Series (Working Paper 9671)*. Cambridge, MA: National Bureau of Economic Research.

Guerrero, G., J. Leon, M. Zapata, C. Sugimaru, and S. Cueto (2012). *What Works to Improve Teacher Attendance in Developing Countries? A Systematic Review*. London: EPPICentre, Social Science Research Unit, Institute of Education, University of London.

Hombrados, J., and H. Waddington (2012). *Internal Validity in Social Experiments and Quasi-Experiments: An Assessment Tool for Reviewers. Mimeo*. London: International Initiative for Impact Evaluation (3ie).

Kremer, M., C. Brannen, and R. Glennerster (2013). The Challenge of Education and Learning in the Developing World. *Science* 340: 297–300. DOI: 10.1126/science.1235350

Krishnaratne, S., H. White, and E. Carpenter (2013). *Quality Education for all Children? What Works in Education in Developing Countries. Working Paper 20*. New Delhi: International Initiative for Impact Evaluation (3ie).

Lugaz, C., A. De Grauwe, D. Balde, C. Diakhate, D. Dougnon, M. Moustapha, and D. Odushina (2010). *Schooling and Decentralization: Patterns and Policy Implications in Francophone West Africa*. Paris: IIEP.

Oliver, S., K. Dickson, and M. Newman (2012). Getting Started with a Review. In D. Gough, S. Oliver, and J. Thomas (edited by), *An Introduction to Systematic Reviews* (p. 66–82). London: Sage Publications.

Petrosino, A., C. Morgan, T. A. Fronius, E. E. Tanner-Smith, and R. F. Boruch (2012). Interventions in Developing Nations for Improving Primary and Secondary School Enrollment of Children: A Systematic Review. *Campbell Systematic Reviews* 2012: 19.

Pherali, T., A. Smith, and T. Vaux (2011). *A Political Economy Analysis of Education in Nepal*. Kathmandu: EU.

Rocha Menocal, A., and B. Sharma (2008). *Joint Evaluation of Citizens' Voice and Accountability: Synthesis Report*. London: DFID.

Rose, P. (2003). Community Participation in School Policy and Practice in Malawi: Balancing Local Knowledge, National Policies and International Agency Priorities. *Compare: A journal of comparative and international education* 33 (1), 47–64. DOI: 10.1080/03057920302597.

Santibanez, L. (2007). *School-Based Management Effects on Educational Outcomes: A Literature Review and Assessment of the Evidence Base*. Toluca: Centro de Investigación y Docencia Económicas.

Snilstveit, B., J. Stevenson, D. Phillips, M. Vojtkova, E. Gallagher, T. Schmidt, H. Jobse, M. Geelen, M. Pastorello, and J. Eyers (2015). *Interventions for Improving Learning Outcomes and Access to Education in Low- and Middle-Income Countries: A Systematic Review, 3ie Systematic Review 24*. London: International Initiative for Impact Evaluation (3ie).

Snilstveit, B., J. Stevenson, R. Menon, D. Phillips, E. Gallagher, M. Geleen, H. Jobse, T. Schmidt, and E. Jimenez. (2016). *The Impact of Education Programmes on Learning and School Participation in Low- nd Middle-income Countries: A Systematic Review Summary Report, 3ie Systematic Review Summary 7*. London: International Initiative for Impact Evaluation (3ie).

Thomas, J., A. Harden, and M. Newman (2012). Synthesis: Combining Results Systematically and Appropriately. In D. Gough, S. Oliver, and J. Thomas (edited by), *An Introduction to Systematic Reviews* (p. 179–226). London: Sage Publications.

UNESCO. (2014). *Teaching and Learning: Achieving Quality for All. EFA Global Monitoring Report 2013/4*. Paris: UNES.

Unterhalter, E. (2012). Silences, Stereotypes and Local Selection: Negotiating Policy and Practice to Implement the MDGs and EFA. In A. Verger, H. K. Altinyelken, and M. Novelli (edited by), *Global Education Policy and International Development: New Agendas, Issues and Policies* (p. 79–100). London: Contiuum.

Westhorp, G., B. Walker, and P. Rogers (2014). *Under What Circumstances Does Enhancing Community Accountability and Empowerment Improve Education Outcomes, Particularly for the Poor? A Realist Synthesis*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

World Bank. (2004). *World Development Report: Making Services Work for Poor People*. Washington, DC: World Bank.

World Bank. (2007). *What Do We Know About School-Based Management?* Washington, DC: World Bank.

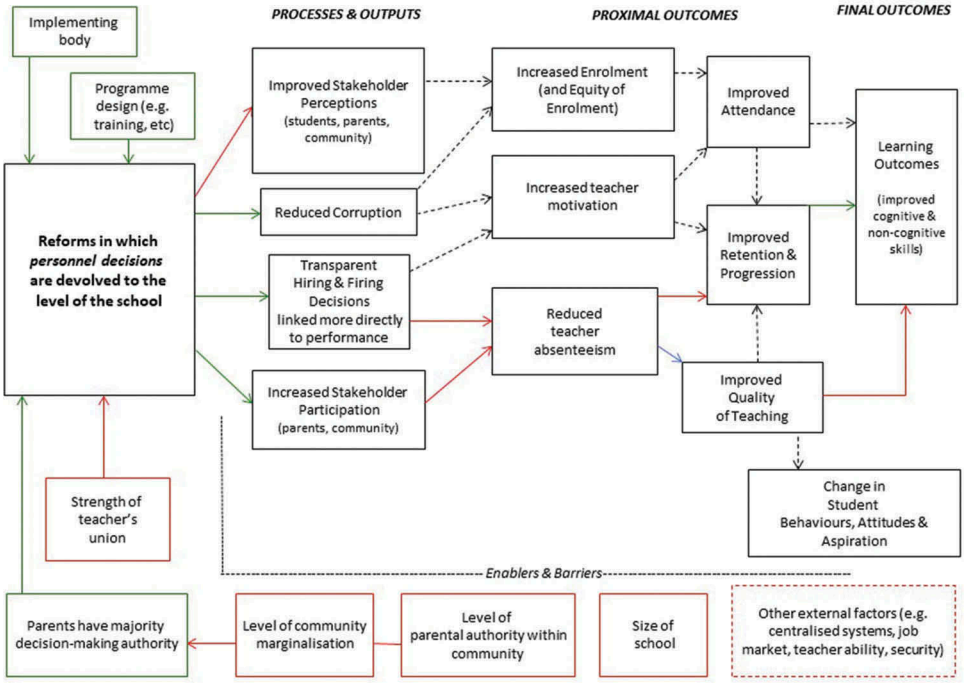

Figure A1. Revised framework for personnel decisions ('high decentralisation').

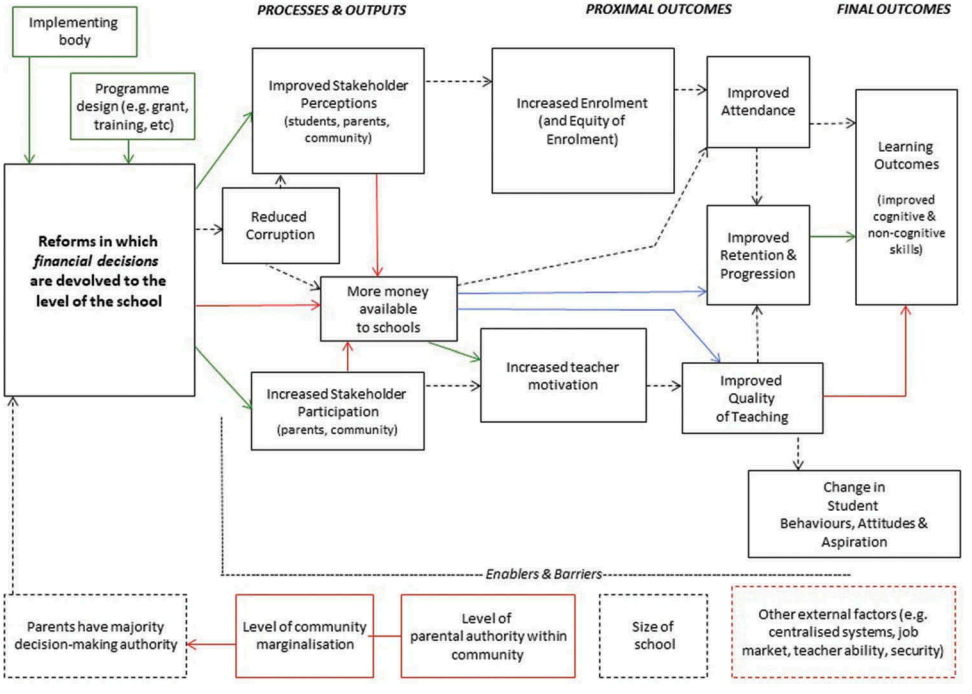

Figure A2. Revised framework for financial decisions ('medium decentralisation').