

School inspections in low and middle income countries; Explaining impact and mechanisms of impact

1. Introduction

Most studies on the impact of school inspections are set in Europe (particularly England and the Netherlands), showing that school inspections can have effects and side effects on teachers' behavioural change, school improvement and student achievement (see recent literature reviews by Klerks, 2013, and Nelson and Ehren, 2014). However, local contexts in low- and middle-income countries are very different from those in developed countries and a lack of financial and material resources have a pronounced impact on the conduct of inspections. De Grauwe (2001), MacPherson (2011), Herselman and Hay (2002), Mazibuko (2007), Uwazi (2009) and Wanzare (2002) explain how Inspectorates of Education in a range of low and middle income countries are characterized by infrequent and limited visits to schools, particularly those in remote areas; lack of stationery and computers also limits the publication and dissemination of inspection reports and limits the collection and analysis of relevant school documents and data in the preparation of upcoming inspection visits. Differences between low and high income countries also include the roles and responsibilities of school inspectors. As De Grauwe (2001, 2007) describes the term 'supervision' is generally used for school inspectors in low and middle income countries and their role is often not only to control and evaluate, but also to advise, assist and support head teachers. Sometimes supervisors even have managerial tasks and are responsible for employment of teachers, or deciding on promotion of teachers and head teachers. The range of additional activities school inspectors in low and middle income countries are expected to undertake, in comparison to their counterparts in developed countries, causes a high work load and limits the time they can actually spend on school inspection visits, according to De Grauwe (2001), Moswela (2010), Mazibuko (2007), and Wanzare (2002).

The differences in the implementation and context of school inspections in high income versus low and middle income countries suggests that findings about the effectiveness of school inspections from Western studies are not easily transferable to the developing world. The complex and varied links amongst governance context, policy, design of inspection systems, mechanisms of impact and school outcomes make translation of research findings across studies challenging. Currently there are no systematic reviews available that attempt to link school inspections in low- and middle-income countries with particular school-level outcomes. This paper aims to address this gap by presenting the results of a systematic review about the conditions under which school inspections lead to improvement in schools and to positive learning outcomes for schoolchildren in low- and middle-income countries. A case study on the impact of school inspections in Ethiopia by the fourth co-author is included to contextualize – and in some respects challenge – the findings from the literature review.

2. Realist synthesis¹

This review follows the approach of realist synthesis. We have turned to realist synthesis (Wong, Greenhalgh, Westhorp, Buckingham & Pawson, 2013; Pawson, 2006; Pawson, Greenhalgh, Harvey & Walshe, 2005) as a means of systematic review because of the complexity and dynamism of conditions that influence the outcome of school inspections in low and middle income countries, the wide variability in available literature, and our aim of explaining how particular organizational outcomes arise, given particular conditions. A realist synthesis allows us to use a wide range of resources, including 'grey literature' (such as conference papers) which present more conceptual work to understand mechanism of change. In a realist framing, the overriding question is, "What works for whom under what circumstances, how and why?" (Wong et al., 2013, p. 2). The goal shifts from pinpointing features of effective interventions to testing a conceptual framework (an 'initial rough theory' in realist terminology) which explains the mechanisms

¹ Section 2 and 3 were adapted from the published review protocol, see <...>

through which school inspection systems, operating under certain conditions, are more or less likely to improve school-level outcomes, answering the following research question:

Under what conditions do school inspections improve school-level outcomes in low- and middle-income countries?

Our review followed six overlapping phases to answer this question:

1. scoping of the literature and theory elaboration
2. search process of databases and hand searching using inclusion criteria as specified in the protocol (published in English, in and after 1990, conducted in low- and lower-middle-income countries according to World Bank classification, target inspection systems in primary and/or secondary education, including all types of study designs, policy and theoretical/conceptual framework documents).
3. screening and selection of relevant papers. A total of 10255 records were screened on title and abstract, 823 full reports were screened and 214 studies were included in the systematic map.
4. characterizing included studies in a systematic map. Studies were characterized on the following criteria: location, publication type, year of publication
5. quality appraisal and data extraction: full reports were coded for rigor of the study and relevance in reporting on school inspections. Only studies coded as highly and somewhat relevant, with a cut-off date of 2001 were included in the analysis and synthesis stage. See appendix 1 for details of the quality appraisal of included studies.
6. analysis and synthesis. 22 sources were analysed. Sources included one conference proceeding, one dissertation, four case studies and 11 journal papers. The majority of the sources provide narrative descriptions of school inspections across a number of countries, particularly in Africa, outlining the problems these inspection systems face in inspecting schools. The empirical papers are from small-scale case studies, often including limited descriptions of underlying methodology and presenting self-reports of small (potentially non-representative) samples of respondents. Only three papers report quantitative results from surveys and secondary data.

The breakdown of sources according to geographical locations shows that the majority of papers are from Sub-Saharan Africa (n = 10), and a smaller number of papers from Latin America and the Caribbean (n = 2), South Asia (n = 2), and East Asia and the Pacific (n = 3). The remaining 5 sources talk about low and middle income countries in general but do not provide specific locations. The breakdown of sources by income level suggests (when information was available) that 5 sources are from low income countries, 7 from lower-middle, while another 7 from upper-middle income countries.

3. Initial rough theory development

The first phase involved the development of an initial rough theory, which was used in phase six of the review to organize our analysis and synthesis. The rough theory explains how school inspections can lead to school-level outcomes. The scoping of literature for use in developing the initial rough theory was described in our review protocol (see authors, p. 27) and involved identifying relevant articles from academic journals, scholarly books and reports from multilateral and regional organisations (e.g., World Bank, IIEP/UNESCO, OECD, Brookings Institute). Two researchers read the full text of 25 articles to develop a generic hypothesis about how systemic accountability, and school inspections specifically, intend to influence service delivery, systemic efficiency, and learning outcomes based on the integrated open systems model of school effectiveness put forward by Scheerens (1992). In this model, which has formed the basis for much work in both low, middle and high income countries, schooling is described of four aspects at the organisational level:

- Inputs of technical, human and social capital
- Processes of the technical and administrative core, with ‘technical’ indicating classroom-level

interactions amongst teacher-students-curriculum and ‘administrative’ the organising processes of the school

- Outputs that relate to student learning
- Outputs that relate to the technical efficiency of the school.

For our initial rough theory we discerned two levels of outcomes – those at the organisational level and those at the level of the educational system. At the organisational level, we considered increased student access to education, reflected in increases in enrolment as well as more regular student attendance; time devoted to teaching in classrooms and greater allocation of education expenditure for teaching and learning as an outcome were also included. Finally, these outcomes could be translated across schools in ways that led towards system outcomes, of technical efficiency as well as societal efficiency (Cheng, 1993) – the contributions of the school and school system to an educated, equitable society.

Within this model, we drew on Ehren’s (2016) literature review of school inspections in high income countries in which she provides evidence of the following five mechanisms to explain how school inspections lead to organisational outcomes. In our realist synthesis we will test these mechanisms for school inspections in low and middle income countries:

- Setting expectations
- Providing feedback/consequences
- Capacity development of educators
- Capacity development of local stakeholders.

The first hypothetical mechanism is about **setting expectations**. Setting of expectations acknowledges the fact that inspection standards have a normative or standardization purpose. Such indicators not only serve a measurement function to undertake inspections, but they also communicate expectations about goals and about what a good school, a good lesson, and good performance constitutes. Schools are expected to use the criteria and descriptors set out in the inspection framework to define their own standards of a ‘good school’ and a ‘good lesson’ and to incorporate these standards in their daily work and teaching. The communication and use of standards in school inspections are expected to motivate schools to reflect on the standards, process them and adapt their goals and their practical ways of working in such a way that they come closer to the normative image of schools communicated by the inspection indicators.

The second hypothetical mechanism is about the **feedback** from inspections that is provided to schools. Such feedback may include an outline of strengths and weaknesses on school quality in inspection reports and during an inspection visit. School inspection systems often set targets for school performance and have consequences (e.g. sanctions and/or rewards) in place for low and high performing schools. Such consequences are expected to motivate schools to attend to the feedback provided. Schools are assumed to use the feedback to improve, and stakeholders are expected to take note of the feedback and hold schools accountable for their use of the feedback for improvement.

Capacity-building of educators is our third mechanism and refers to the school’s capacity to enhance the professional learning of teachers and to transform large-scale reform into accountable student-oriented teaching practices. Improvement capacity is considered to be an important condition for school development in general, as well as in response to external school inspections. School inspections are expected to build a school’s capacity for improvement, primarily through impact on school self-evaluation and the school’s internal quality assurance systems and the impact on professional development, school collaboration, school leadership and external support around (improvement on) inspection indicators. High quality self-evaluation is considered to be a critical element in improvement of schools as schools identify and correct problems in the quality of their school in preparation for, and in response to inspections. Internal quality assurance mechanisms, together with external school inspections, are seen as inseparable and integral parts of an informed and evidence-based improvement cycle that build capacity in schools to improve the teaching and learning and lead to improved student outcomes.

Capacity development of local stakeholders, as a fourth mechanism, is about engaging a ‘third’ party in school inspections, providing them with the information and support to have an active role in school evaluation and improvement. Local stakeholders typically include parents and community members, as well as students and local officials. Examples of capacity development might include the public dissemination of inspection reports, as well as forms of participatory evaluations in which a school’s stakeholders take an active role in the evaluation of schools, such as when stakeholders are involved in the development of inspection standards, school inspectors interview parents or school boards during school inspections, or require the school to actively engage with community members in the process of constructing and analysing school inspection reports.

The inclusion of stakeholders as a ‘third’ party in school inspections is expected to reinforce public recognition of inspection standards and make it more likely that schools react to inspection standards in anticipation of the response of local stakeholders. Stakeholders may, however, become more active and raise their ‘voice’ in order to motivate schools to improve. If schools do not give stakeholders sufficient opportunities for participation (in that they accept some ‘stakeholders’ influence’ or enter into ‘negotiation’ with them), stakeholders may retreat to the option of ‘choice’ or ‘exit’ where parents choose to enter or move their child to a higher performing school. ‘Choice’ and ‘exit’ are expected to exert pressure on schools to conform to inspection standards through the introduction of competition between different providers, while voice alternatives allow parents to express preferences and opinions around education service delivery that would motivate schools to improve.

Our interest in this review was in examining the mechanisms that explain if/how school inspections lead to improved school-level outcomes in low and middle income countries. We also used a case study from the fourth co-author in Ethiopia to contextualize our findings from the review. The case study was implemented in August 2015 with the purpose of analysing and evaluating the operation of the national inspection system. Data collection included meetings with World Bank staff, the State Minister for General Education and the Director and members of the GEID team. Additionally interviews were held with inspectors in the Regional Education Bureaux (REBs) of Oromia, Amhara and Southern Nations, Nationalities and Peoples (SNNP) Regions, and in several Zone Education Offices (ZEOs) and Woreda Education Offices (WEOs) within those regions. The directors of five primary and three secondary schools were also interviewed. Documents and data on the inspection process and its outcomes so far were analysed.

4. Findings

Our initial screening and selection of the literature included 214 sources but after a quality appraisal of the rigor and the relevance of these papers, only 22 studies were included in the review. These studies present detailed findings on school inspections in the countries of interest, but the rigor of the limited number of studies varied: 3 studies were counted as high-rigour (2 of these did not have inspection as the main topic of study), 8 studies were medium-rigour and 11 were low-rigour. Most sources describe an overall lack of impact of school inspections in low and middle income countries, while a small number suggest inspection does have impact. In the following section, we detail these findings. We then use a case study to illustrate the mechanisms of change from inspections in Ethiopia.

Only eleven studies in our review analysed outcomes of school inspections, describing an overall lack of impact of school inspections in Uganda, South Africa, Timor Leste, Peru, Africa and Pakistan (Crouch and Winkler, 2008; Marlien and Driekie, 2002; Mazibuko, 2009; Macpherson, 2011; Uwazi, 2009; Mazibuko, 2009), while two studies in Ghana and South Africa (Opoku-Asare, 2006; Mazibuko, 2009) also point to unintended consequences. A study by Opoku-Assara (2006) for example shows how teachers in Ghana put on an act during inspection classroom observations, and how principals use inspections to threaten their teachers in South Africa (Mazibuko, 2009). According to Opoku-Asare (2006), school inspections are often pre-announced and enabling the teachers concerned to prepare adequately for the observation lessons. This

enables those teachers to arm themselves with all the teaching materials they can possibly lay hands on and sometimes, rehearse the lessons they intend to teach for the exercise.

Two studies suggest that inspection have an impact. Macpherson (2011) describes how school inspections in Timor Leste have the potential to contain the scale of corruption in the misuse of school grants when policing transparency in the collection and disbursement of schools grants, while not engaging in the processes of collection and disbursement themselves. As the study only looked at how schools are investigated and how inspectors investigate allegations of misuse, no claims can be made about school inspections actually leading to a decrease in corruption.

Brock (2009) draws on a number of somewhat rigorous case studies in Gansu when explaining how increased power to school inspectors to report on the quality of schools, to propose changes and support in/of schools, lead to an improvement of school development planning: schools set out specific goals for their development in close cooperation with the local community in which they take into account the needs of poorest children and developed learning materials to address these needs. These school development goals could subsequently be measured by inspectors. Our case study of Ethiopia also offers early evidence of the impact of inspection:

Box 1. Impact of inspections in Ethiopia

The Ethiopia case study shows how a number of re-inspected schools show improved standards in comparison to their previous (first) inspection visit, particularly in the construction of new classrooms (in the vernacular style, using local materials, with community support), and the provision and equipping of laboratories and libraries. These improvements were often made possible by use of the School Grant, paid to all schools as part of the General Education Quality Improvement Programme (GEQIP, a programme managed by the Ministry of Education and supported by international donors) and intended to support the improvement of teaching and learning. In many cases, local communities also supported the improvement of their schools by raising funds and providing help in kind; for example, contributing building materials and the labour to erect new classrooms.

Improvement of the quality of teaching, support for students and school management also received specific attention in some of the schools. A majority of the schools could demonstrate that such actions had helped them to improve outcomes for their students. Three primary schools were able to demonstrate that academic results improved from 2013-4 to 2014-15, at least in part as a result of the actions taken; and one primary school could show a significant reduction in the student drop-out rate.

4.2 Mechanisms of impact

The following section summarizes our literature review findings by the mechanisms we proposed in our initial rough theory.

Setting expectations

Six studies present findings from Gansu, Uganda, Pakistan, South Africa, Mexico and Peru that indicate how the development of standards, guidelines and frameworks to inspect schools can be an important driver for improvement as they inform schools of where to focus on in their improvement plans, support school self-evaluation and as they ensure consistency of inspection assessments and feedback to schools. Currently many low and middle income countries do not have such guidelines to evaluate the quality of schools which leads to inconstancy in the evaluation of schools and also limits schools in their preparation for visits and in the incorporation of inspection standards and criteria in their school development planning. As Brock (2009) explains:

“Frameworks and guidelines are made available to both schools and inspectors. Availability to schools supported school development planning as a vehicle for change in which schools and local communities are brought together to create a unified approach to the school’s development, and

in which the relationship between the county education bureau and the school changes from top down to bottom up development.”

It seems that openness of frameworks and inspection schedules allows schools to prepare for visits and creates buy-in to the inspection process which would promote improvement. As De Grauwe (2001, p.17; 2008, p.14) suggests ‘openness and transparency are increasingly encouraged, implying that schools will be informed beforehand of visits. As a result, where these reforms are actually being implemented, teachers are beginning to consider inspectors as sources of help rather than of criticism, and start applying the same frameworks and norms to the evaluation and improvement of their work throughout the country. Brock (2009) also found that making frameworks and guidelines available to schools and inspectors created a more bottom-up and unified approach to school development as it brought together schools and local communities in setting priorities for improvement. These processes to school development are, according to Churches and McBride (2013) and Moswela (2010), enhanced when stakeholders, such as principals, proprietors, employers, higher education providers and teachers are involved in the design of frameworks and buy-in is created for the evaluation of teachers and schools. Such buy-in is an important condition for impact of inspections as teachers and teacher unions have strong power positions to resist inspections and often do so (De Grauwe, 2008). Openness and transparency of frameworks, buy-in to these frameworks, and bottom-up processes of school development seem to result in standardisation of quality across a country and may have, according to De Grauwe (2008), a desirable impact on improvement of schools, particularly in homogenous countries with little disparities.

Santiago et al (2012) suggest that such tools and guidelines can support schools in engaging in self-evaluations and enhances a common language of quality in a country. Guidelines and frameworks also support school inspectors in their evaluation of schools and enable them to have a professional dialogue with school staff about potential improvements. Such a dialogue is considered to support the school’s acceptance and use of inspection feedback. An important condition is however, according to Santiago et al (2012) to prevent schools from being overloaded with different types of guidelines and materials as this will confuse them when deciding on which approach to focus on in improvement the quality of their school. Wanzare (2002, p.21) also discusses how schools can use inspection handbooks in their evaluations if these handbooks are not too detailed, bureaucratic or rigid. Churches and McBride (2013) suggest that buy-in and use of these handbooks and standards by stakeholders (such as schools) is enhanced when they are involved in the design of these handbooks and standards.

Box 2. Setting expectations in Ethiopia

The inspection framework was developed for Ethiopia as a whole, and its details reflect the national context as well as international practice. A national framework was considered important because of the need to be able to compare the quality of schools in different regions, and to establish a secure baseline against which future improvement could be measured. The inspection and self-assessment frameworks had been extremely well received in the three regions visited. Both inspectors (at regional, zone and woreda levels) and school staff said that they liked the detail and precision of the standards and indicators, and found them very useful as tools for identifying schools’ strengths and areas for improvement. During the first two years of inspections however a range of revisions were suggested to fine-tune the framework for private schools. Another sentiment expressed fairly commonly was that some of the standards relating to buildings, facilities and the qualifications of staff, cannot be achieved by some schools. Inspectors’ explanations that schools should not be blamed for failing to meet these standards (but responsible authorities) ensured that schools generally accepted the rationale for these standards. Inspectors and school directors commented favourably on the transparency of the inspection framework as a basis for evaluating schools. The training provided for both groups on the standards appeared to have been important in establishing a common understanding of the nature and purpose of the framework, and its function as a tool to drive school improvement.

Providing feedback information

In our initial rough theory we suggested that inspections could have an impact on school-level outcomes through the feedback from school inspectors on strengths and weaknesses in school quality. Feedback refers to which priorities for improvement are set and communicated to schools (adapted to local context), and targeting weak schools for visits and further feedback. Fourteen papers reflect on inspection feedback and communication in relation to school inspections. Of these 14 papers, only eight papers (one high rigour, five medium rigour, two low rigour) present actual study findings while six papers only make claims about the functioning of inspection feedback.

All of the papers describe a lack of impact from inspection feedback which is, in the author's views, caused by specific attributes of the feedback, and caused by a number of conditions. These attributes and conditions fail to 'fire' any improvement mechanism from inspection feedback, such as when school staff accept inspection feedback and use it to improve the school's weaknesses to enhance student outcomes. Each of these conditions will be described below and how they have failed to lead to improved outcomes.

Content of the feedback

Several authors discuss the importance of feedback and communication of inspection findings in school improvement and claim that the content of the feedback is an important cause of limited improvement from inspection feedback in low and middle income countries. Three papers in Indonesia and Ghana present findings from primary research which indicate that inspections particularly focus on bureaucratic and administrative issues, checking figures and compliance to regulations which are not considered to be relevant for school improvement and are often outside of the school's span of control (Chen, 2011; Darvas and Balwanz, 2014; Opoku-Asare, 2006). Inspection recommendations do not focus on vital problems in schools and are often repeated in a routine manner, year after year. As a result, schools fail to accept and/or implement inspection recommendations while Uwazi (2009) presents study findings which suggest that such a focus on administrative and bureaucratic issues takes time away from schools to focus on actual improvement of student outcomes. Similar claims have been made by Santiago et al (2012) in a country review of Mexico.

As Uwazi (2009) and Opoku-Asare (2006) explain, the inspection recommendations are often generic, unrealistic and often require additional resources that the school administration is not able to acquire. These papers highlighted how the Tanzania and Ghana Inspectorate of Education often provided advice to schools which should be aimed at the Ministry of Education, such as hiring more teachers, acquiring more textbooks, or construct/renovate school buildings. Uwazi (2009) suggests that inspections can only be effective if they address issues of poor performing students, how to address dropout rates, and how to improve learning and instruction and/or training gaps in schools. Similar issues are discussed in papers from De Grauwe (2007), Jaffer (2010), Santiago et al (2012), Wanzare (2002) referring to Africa, Pakistan, Mexico and Kenya.

A number of authors explain why inspectors in low and middle income countries fail to give relevant feedback to schools (Churches and McBride, 2013; Darvas and Balwanz, 2014; De Grauwe, 2001, 2007, 2008; Harber, 2006; Jaffer, 2010; Wanzare, 2002). They point to the work overload of inspectors (both in numbers of schools to inspect, as well as in number of indicators to inspect) which lead them to focus on a simple checking and control of administrative protocols. Performance management systems hold inspectors to account for the number of schools visited (instead of impact and quality of feedback) which would lead them to focus on checking of facts and figures. Other conditions of ineffective feedback are a lack of professionalism of school inspectors and lack of training in evaluation of school quality. School inspections are often also prioritized for schools that are suspected of irregular use of resources and misconduct of teachers, while inspectors also seem to feel that control of administrative procedures gives them power over schools, and authority in the evaluation of schools as it would signal a clear mandate from central government. Such status, credibility and authority is often lacking (see section below).

A number of authors also suggest that inspection feedback and standards need to fit within the local context in order to motivate school improvement and address local priorities and issues. None of these papers have however actually investigated a relation between adapting (inspection) standards to local context and improvement of schools, but they suggest that inspection purposes and priorities need to be adapted to the history and culture that underpin the local context of a country, or of different regions within a country to advance school improvement, particularly in heterogeneous countries (De Grauwe, 2008). Inspectors need to have an open mind in order to recognize excellence and understand the existing restraints on pedagogy that exist in a specific context (such as class size and resourcing).

Communication and tone of voice

Studies in Africa also suggest that the hostile and intimidating tone of voice of school inspectors is another explanation for the lack of impact of inspection feedback (De Grauwe, 2007, 2008; Wanzare 2002). Wanzare (2002) and Moswela (2010) for example explain that:

“Inspection of schools in Kenya has at times been marked by impromptu, irregular visits by some inspectors with the object of “catching” the teachers doing the wrong. Some school inspectors have been criticized for being harsh to teachers and for harassing teachers even in front of their pupils.” (Wanzare, 2002, p.10)

“The environment in which instructional supervision takes place in schools (in Botswana) is rather hostile and intimidating to teachers to make any meaningful impression on the improvement of teaching standards” (Moswela, 2010).

As findings from Moswela’s study (2010) suggest, a condescending tone in communicating with schools and presenting feedback leads to a poor relationship between teachers and inspectors and makes no impression on teachers, resulting in a lack of impact on improvement of schools.

Lack of credibility of inspectors and inspection feedback

The lack of perceived expertise, status and credibility of school inspectors by school staff is also expected to limit the implementation of inspection feedback. Eight papers in Nigeria, Botswana, Namibia, Tanzania, Zimbabwe, Eastern and Southern Africa, Pakistan, Timor, Mexico and Kenya present findings from primary research to explain that school inspectors are often not trained in the evaluation of schools and have limited expertise in doing such evaluations which would limit their credibility and the credibility of inspection findings in the eyes of school staff (De Grauwe, 2001; Harber, 2006; Jaffer, 2010; Macpherson, 2011; Moswela, 2010; Santiago et al, 2012; Uwazi, 2009; Wanzare, 2002). They have no expertise in how to objectively evaluate schools/teachers, how to provide accurate and consistent feedback on strengths and weaknesses and how to engage schools in a professional dialogue about school improvement.

“However, it was reported that in general there is much variation in the quality of advice and support supervisors may be able to offer schools. The capacity of supervisors in general to engage in school evaluations in ways which may promote school improvement as well as resulting in accurate evaluation of the quality of a school’s work is limited under present conditions” (OECD review in Mexico, Santiago et al, 2012).

These papers also highlight how the overall lack of systems and structures around human resource management and development to support the hiring and training of a high quality inspection core hampers the credibility of school inspectors. Studies in Botswana, Namibia, Tanzania, Zimbabwe (De Grauwe, 2007; Moswela, 2010), Pakistan (Jaffer, 2010), and Mexico (Santiago et al, 2012). For example, they explain that inspectors are often recruited on an ad hoc basis from a pool of teachers and principals with long service where personal connections and political affiliations are used to transfer people into inspection posts.

Studies in Pakistan (Jaffer, 2010), Botswana and Tanzania (De Grauwe, 2001) also suggest that the lack of credibility is caused by recruitment issues (favouring individuals with high political influence), and the pay scale of school inspectors which is on a lower grade as the head teachers they are inspecting, causing head teachers to believe that inspectors are not of a high status and that their feedback can be disregarded.

“The inspector’s position was equivalent to the teaching grade of a high school teacher, so these teachers and the inspectors were at the same grade and salary scale. Hence, individuals could not

be held accountable for sub-standard performance or rewarded for good performance. No pre-service training was provided to prepare the inspectors for the specific roles and responsibilities of the post. These lacunae further weakened the inspectors' position and authority, and also impacted adversely on the efficiency of the inspection system.” (Jaffer, 2010, p.378)

There also seems to be little incentive in place to improve the overall quality of inspections as the number of visits to schools is the main performance indicator used in the evaluation and monitoring of school inspectors, and not the quality of their work. A study Tanzania by the national audit office (Uwazi, 2009) also highlights that there is no monitoring of inspection systems to learn about what works and doesn't work and to improve the functioning and impact of inspections. According to the Tanzania national audit office (Uwazi, 2009), such monitoring needs to analyse the extent to which schools have implemented recommendations, stakeholders' perceptions of the recommendations and the impact of implemented recommendations.

“It shows that the school inspectorate programme is not functioning properly and therefore fails to safeguard quality of instruction and its improvement by: failing to prioritise the issues of poor performance of students in the inspection cycle, not effectively communicating and following up on implementation of recommendations, failing to monitor the effectiveness of school inspections” (Uwazi, 2009).

Consequences from inspection feedback

Finally, six studies in Indonesia, Uganda, Namibia, Tanzania, Ghana and Kenya find that inspection feedback has no impact because of the lack of consequences for failing schools and the lack of follow-up on inspection visits. Chen (2011) and Crouch and Winkler (2008) also explain that, even if schools can be sanctioned by law, there are no means or mechanisms in place to actually implement such consequences. Too close relationships between schools and inspectors and negative inspection reports not being published may add to such lack of implementation as Opoku-Asare (2006) found in Ghana.

Jaffer (2010) also found that the lack of credibility of inspectors (due to a lack of training, political appointments into post, and pay scale similar to teachers) renders them powerless in holding schools accountable for low performance. Similar findings on the lack of consequences from inspections have been discussed in study by Churches and McBride (2013) in conference proceedings for Nigeria.

Where inspectors do have the power to report, propose changes and support schools, they enhance school development planning, particularly when they measure school goals, according to Brock (2009).

Box 3. Inspection feedback in Ethiopia

Our case study shows how inspectors, on their arrival at the school, inspectors took time to explain the nature and benefits of the process and created a positive atmosphere. The director of one primary school said that this had influenced the attitudes of the teachers in a positive way. On several occasions it was stated that schools had initially been suspicious of inspection, expecting it to be an exercise in fault finding and the allocation of blame for failings, but had become convinced of the benefits when inspectors stressed that the main purpose was to promote and support improvement. Judgements that schools did not meet some of the standards relating to inputs such as the buildings and their facilities sometimes met resistance, but inspectors reported that schools were generally reconciled when they explained that these were not matters for which they would be blamed, but rather issues that needed to be brought to the attention of the responsible authorities.

Where feedback was given to individual teachers following observations of lessons, this was highly valued; although, it was not always given. In the schools visited, the inspection reports and their recommendations were considered credible, accurate and useful, although there were indications that their quality varied. Among the factors that appeared likely to be contributing to the variation of quality were the varying practices in the appointment and training of inspectors, and the high turnover of personnel. The qualifications required of inspectors are defined in the inspection guidelines but in at least one of the

woredas visited these had not been followed in their entirety. Although its content appeared to have been fairly consistent, the number of days training provided had varied from region to region. The high turnover seemed to be related to the relatively low remuneration of inspectors, who are paid less than supervisors and school principals.

Capacity-building of educators

Capacity-building of educators particularly refers to the ability of schools to evaluate and improve their own performance, and the capacity to implement improvements. Studies discuss how linking external inspections and internal school self-evaluation may motivate self-evaluation of schools and suggest how school self-evaluation may lead to more sustainable improvement. None of the authors have however studied the relationship between strengthening internal evaluations in inspection systems and actual school improvement.

De Grauwe (2001, 2007, 2008) and Herselman and Hay (2002) expect that an increase in school internal evaluation will strengthen participation and commitment of teachers in school change and creates a culture of quality in which teachers reflect on their own practices which is expected to lead to more sustainable improvement. External support for internal evaluations and resulting improvement, such as from resource centres, may strengthen linkages between schools and prevent isolation of schools. It is also expected to strengthen school management and culture and the school's capacity to improve.

According to De Grauwe (2008, p.15), internal evaluations of schools in response to external inspections can however only be effective if school inspectors take these evaluations seriously when they inspect the schools, if there is overlap in internal and external frameworks and criteria for making a judgement about school practices, and if the agenda for these self-evaluations fit the improvement priorities of schools and countries (instead of being driven by donor organizations). Schools also need support in the implementation of rigorous self-evaluations. Wanzare (2002, p.21) discusses how schools can use inspection handbooks in their evaluations if these handbooks are not too detailed, bureaucratic or rigid, while Moswela (2010) suggests that teachers who have an active part in inspections throughout the year are better able to improve their teaching. Capacity for improvement however also refers to the resources and knowledge in schools to address and implement improvements and to effectively engage in whole school evaluation and school inspections as is also evidenced in the case study from Ethiopia:

Box 4. Capacity-building of educators in Ethiopia

In Ethiopia, schools are required to use the inspection framework to complete an annual self-assessment. In years 2012-13, 2013-14 and 2014-15, a great majority of the schools in the three regions had completed such a self-assessment. Some schools were reported to have also completed a separate developmental self-assessment, using a framework produced in connection with the School Improvement Programme. This involved a duplication of effort, which it is hoped to avoid in the future by unifying the two programmes. All schools, apart from one had completed a self-assessment using the inspection framework. Several of the principals stated that they found this assessment very helpful although challenging. Inspectors reported that the self-assessments were not always accurate, often tending to be over-sanguine, particularly in the case of private schools, but that their rigour had improved after the schools had been inspected. Consequently, when schools were re-inspected, their self-assessments conformed more closely with the inspectors' judgements.

Evidence of the use of inspection results to build the capacity of teachers and school directors was limited although there were instances when schools had organised teacher training in response to inspectors' recommendations, and one school graded 'achieving the standards' was encouraged by its supervisor to support the improvement of other schools in the cluster.

Capacity Development of Stakeholders

A final condition for effective inspections is the alignment of (actions of) stakeholders in the education system and their capacity to implement and support school improvement. These conditions refer to both the building of capacity of stakeholders in our initial rough theory, as well as the setting of expectations.

De Grauwe (2001), Mazibuko (2009) and Opoku-Asare (2006) describe how limited coordination between the Inspectorate of Education and other national stakeholders, such as teacher training or resource centres in the dissemination and use of inspection findings, potentially limits the impact of school inspections. Limited coordination between the Inspectorate and other stakeholders in the education system particularly leads to a lack of follow-up on school inspection visits and findings, and limited support to schools on the implementation of inspection feedback.

Six papers discuss the relation between alignment and follow-up on inspection assessments and school improvement, but only one study (Jaffer, 2010 in Pakistan) actually reports findings that support such a relationship:

“The problem, as the respondents indicated, was that others rarely followed up on the supervisor’s recommendations. As one respondent put it, ‘the higher authorities just write ‘seen’ on the supervisor’s recommendations without taking any action. There is no decision on the actions that we have suggested for school improvement. They ignore our note. And so we know that nothing will come out of these reports and efforts’.” (Jaffer, 2010, p.386)

Other authors discuss similar issues, such as De Grauwe (2001, p.143), who explains:

“Co-ordination is difficult, especially between the supervision service and other services which work towards pedagogical improvement, such as teacher training, teacher resource centres, curriculum development and examinations. The follow-up to school visits suffers from this lack of coordination. Recommendations made in inspection reports and addressed to the administrative and/or pedagogical authorities, remain words in the wind, which frustrates the school staff as well as the supervisors.”

Coordination and alignment is needed to disseminate knowledge from inspection visits and make sure that relevant actors follow-up on inspection recommendations. Alignment is also strongly related to buy-in to inspection standards and recommendations from relevant actors (such as teacher unions) who will then support and pressure school staff to act on inspection findings. Follow-up on inspection recommendations through better alignment of actions of stakeholders in the education system is needed to prevent an overall sense of inertia and demotivation

Examples of how such alignment can be improved are given in study by De Grauwe (2001) who describes:

“Relationships between supervision and the other pedagogical services are close and institutionalized in Botswana, supervisors being members of committees and panels in charge of curriculum development, teacher training and examinations. In Zimbabwe, their involvement is less institutionalized but still quite intense: supervisors serve as resource persons in training and participate in writing test items, marking examinations and preparing evaluation reports. In Tanzania, supervisors sit on curriculum panels and help with examinations, but in practice their involvement in pedagogical improvement is limited because of the practical constraints on their work.”

Alignment of, and coordination between different agencies and offices is however difficult, according to De Grauwe (2008) as it goes against the sense of independence many of these agencies have, and their differences in opinion about for example adequate teaching methods. Mazibuko (2009) also found that in South Africa (Kwa-Zulu Natal) there is a lack of clarity in roles of support offices, districts and inspectors with hampers the support of schools in using inspection recommendations to improve.

Box 5. Capacity development of local stakeholders

Our case study in Ethiopia shows how the actions taken by individual schools in response to the recommendations made by inspectors often involved the engagement of members of the local community, who in many cases made material contributions, in cash or in kind. It appears that the results of inspections were generally shared with parents, through the representation of the parent-teacher association at the oral feedback meeting at the end of the inspection or in meetings at the start of the academic year. When sharing the inspection report, parents' contribution to the school was discussed and they were encouraged to help improve their children's behaviour. Inspectors explained how inspections had galvanised parent-teacher associations; while some communities had raised funds for their schools and had put pressure on their woreda authorities to lobby the regional education bureau for more resources. Inspection seems to have raised both the awareness and the expectations of communities of what their schools should be like.

Evidence of the engagement of other service providers was less plentiful, but a model was provided by South Gonder Zone, whose organising inspector had sent reports to the woreda education offices identifying areas in which, on the evidence from inspections, teachers and school directors required more training.

6. Conclusion and discussion

We used realist synthesis to understand the connections between school inspections and school outcomes in low- and middle-income countries, and the mechanisms of change that motivate school improvement from school inspections. Additionally, a previously unpublished case-study from Ethiopia was drawn upon.

The review for example indicated that development of standards, guidelines and frameworks can be an important driver for school-level improvement as openness of frameworks and inspection schedules allows schools to prepare for visits and creates buy-in to the inspection process (*setting expectations*). Openness of frameworks and inspection schedules allows schools to prepare for visits and creates buy-in to the inspection process which can promote improvement, while tools and guidelines can support schools in engaging in self-evaluations. Studies also show how a lack of high quality *feedback and a lack of consequences* may have caused an overall lack of impact of school inspections in developing countries; available studies particularly point to specific attributes of inspection feedback, such as lack of credibility of inspectors (e.g. due to low pay scale), disrespectful tone of voice, and recommendations on administrative procedures and conditions out of the school's control. The lack of impact is exacerbated by the absence of inspection sanctions for failing schools. One study in particular (Brock, 2009) suggests that inspectors' power to report, propose changes and support may enhance school development planning, particularly when inspectors measure school goals.

Capacity development of educators was also suggested an important mechanism of change in a range of studies, describing how an increase in school internal evaluation can (when used in school inspections) strengthen participation and commitment of teachers in school change and sustainable improvement, particularly when schools are supported in the development of their internal evaluations and have access to guidelines and handbooks that would support their evaluation. None of the authors have however studied the relationship between strengthening internal evaluations in inspection systems and actual school improvement.

Capacity development of stakeholders particularly referred to coordination between the Inspectorate of Education and other national stakeholders, such as teacher training or resource centres in the dissemination and use of inspection findings. Strong alignment between these key stakeholders was found to enhance the dissemination and use of inspection findings and the follow-up on school inspection visits and findings, as well as ensuring that school improvement efforts across the system focus on the same standards (preventing confusion of schools). Alignment of, and coordination between different agencies and offices is however difficult, according to De Grauwe (2008) as it goes against the sense of independence many of these agencies have, and their differences in opinion about for example adequate teaching methods.

Our findings also indicate that these four mechanisms are inter-related and cannot be separated when explaining how inspection leads to improvement. Schools' acceptance and use of inspection feedback is for example indicated as an important mechanism of change across the papers, but many indicate the high level of support schools need to effectively use feedback and implement improvements, as well as adequate communication and distribution of feedback (e.g. the tone of voice of school inspectors, and proper distribution of inspection reports). Schools (and national policymakers) need support in the improvement of identified weaknesses which often requires system-wide improvement of available resources. Such support equally sets expectations in schools around standards of good education and institutionalises external inspection standards. Support for the use of feedback therefore strongly links to our description of capacity-building of educators and stakeholders and setting of expectations. Similarly, developing inspection systems with schools and local stakeholders (e.g. developing inspection standards and data collection methods which include schools' self-evaluations) enhances the capacity of schools and stakeholders, but also sets expectations around evaluation and improvement and institutionalizes external inspection norms.

'Capacity development' of both educators and stakeholders therefore shifts position in our initial rough theory. We initially hypothesized that capacity development was triggered by the provision of feedback/consequences from inspection, but our review suggests that capacity development serves as a precondition for triggering the setting of expectations or intended responses from the provision of feedback/consequences.

Research has however yet to trace clear connections between change in processes at the school-level that occur as a result of school inspections and changes in student learning outcomes, particularly in low and middle income countries. Our elaboration of the interdependent mechanisms of impact from school inspections however provides insight into the most salient relationships to explore in further research and suggests that future studies should specifically look at interlocking mechanisms and conditions of change. The most salient implication for policy-makers and donors is to build capacity of educators and stakeholders first before implementing a school inspection system, or as part of the implementation phase. Capacity-building should be focused on ensuring strong alignment and coordination between Inspectorates of Education and other education service providers or stakeholders in the education system, and should ensure the capacity of school inspectors to provide high quality feedback which yields desirable school, system, and student outcomes.

Appendix 1. Reviewers' judgements about rigour and relevance of each study

Studies	Rigour			Relevance		
	High	Medium	Low	High	Medium	Low
Alcazar (2006)	✓				✓	
Barrett (2011)			✓		✓	
Brock (2009)		✓			✓	
Chen (2011)		✓			✓	
Churches (2013)			✓	✓		
Crouch (2008)			✓	✓		
Darvas (2014)	✓				✓	
De Grauwe (2007)			✓	✓		
De Grauwe (2001)	✓			✓		
De Grauwe (2008)			✓	✓		
Harber (2006)		✓			✓	
Herselman (2002)			✓		✓	
Jaffer (2010)			✓	✓		
Kingdon and Muzammil (2012)		✓		✓		
Macpherson (2011)			✓		✓	
Mazibuko (2007)			✓	✓		
Moswela (2010)		✓		✓		
Opoku-Asare (2006)		✓		✓		
Santiago (2012)		✓			✓	
Uwazi (2009)		✓		✓		
Wanzare (2002)			✓	✓		
World Bank (2010)			✓		✓	

References

- ADEA. (2001). *What Works and What's New in Education: Africa Speaks!* (No. ISBN: 92-9178-026-X). Paris France: ADEA.
- Alcazar, L., Rogers, F. H., Chaudhury, N., Hammer, J., Kremer, M., & Muralidharan, K. (2006). Why Are Teachers Absent? Probing Service Delivery in Peruvian Primary Schools. *International Journal of Educational Research*, 45, 117-136.
- Barrett, A. M. (2011). A Millennium Learning Goal for education post-2015: a question of outcomes or processes. *47(1)*, 119-133.
- Brock, A. (2009). Moving mountains stone by stone: reforming rural education in China. *International Journal of Educational Development*, 29 (5), 454-462.
- Chen, D. (2011). *School-Based Management, School Decision-Making and Education Outcomes in Indonesian Primary Schools*: World Bank.
- Churches, R., & McBride, C. (2013). *Making external school review effective: CFBT*.
- Crouch, L., & Winkler, D. R. (2008). *Governance, management and financing of Education for All: basic frameworks and case studies*: UNESCO.
- Cueto, S. (2005, 2005/10//). *Empirical Information and the Development of Educational Policies in Latin America*. Paper presented at the Regional Dialogue on Education of the Inter-American Development Bank.
- Darvas, P., & Balwanz, D. (2014). *Basic Education beyond the Millennium Development Goals in Ghana : How Equity in Service Delivery Affects Educational and Learning Outcomes* (No.: 978-1-4648-0098-6). Washington US: World Bank Studies.
- De Grauwe, A. (2001). *School supervision in four African countries: Vol. I: Challenges and reforms*. Paris France: International Institute for Educational Planning/UNESCO.
- De Grauwe, A. (2007). Transforming School Supervision into a Tool for Quality Improvement. *International Review of Education*, 53(5/6), 709-714.
- De Grauwe, A. (2008). *School monitoring systems and their impact on disparities*: UNESCO.
- Eddy Spicer D, Ehren M, Bangpan M, Khatwa M (2014) *Under what conditions do inspection, monitoring and assessment improve system efficiency, service delivery and learning outcomes for the poorest and most marginalised? A realist synthesis of school accountability in low- and middle-income countries. Protocol*. London: EPPI-Centre, Social Science Research Centre, Institute of Education, University of London.
<http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=x-3hzguBXhY%3D&tabid=3174>
- Gershberg, A. & Gonzalez, P. (2012) Understanding and Improving Accountability in Education: A Conceptual Framework and Guideposts from Three Decentralization Reform Experiences in Latin America. *World Development*, 40, 1024–1041.
- Gvirtz, S. (2002). Curricular Reforms in Latin America with Special Emphasis on the Argentine Case. *Comparative Education*, 38 (4), 453-469.
- Harber, C. (2006). Democracy, development and education: Working with the Gambian inspectorate. *International Journal of Educational Development*, 26 (6), 618-630.
- Higgins, L., & Rwanyange, R. (2005). Ownership in the Education Reform Process in Uganda. *Compare A Journal of Comparative Education*, 35 (1), 7-26.
- Jaffer, K. (2010). School inspection and supervision in Pakistan: approaches and issues. *Prospects*, 40 (3), 375-392.
- Kingdon, G., & Muzammil, M. (2012). The School Governance Environment in Uttar Pradesh, India: Implications for Teacher Accountability and Effort. *Journal of Development Studies*, 49 (2), 251-269.
- Klerks M (2013) The effect of school inspections: a systematic review. Paper presented at: ORD, Wageningen, The Netherlands, 20-22 June 2012. <http://schoolinspections.eu/wp-content/uploads/downloads/2013/12/ORD-paper-2012-Review-Effect-School-Inspections-MKLERKS.pdf> (accessed 7 June 2014).

- Macpherson, R. (2011). Educational Administration in Timor Leste: Language Policy and Capacity Building Challenges in a Post-Conflict Context. *International Journal of Educational Management*, 25 (2), 186-203.
- Marlien, H., & Driekie, H. (2002). Quality assurance in the Foundation Phase in the Eastern Cape province: a case study. *South African Journal of Education*, 22 (3), 239-245.
- Mazibuko, S. P. (2009). *The managerial role of the principal in whole-school evaluation in the context of disadvantaged schools in Kwazulu-Natal*.
- Moswela, B. (2010). Instructional supervision in Botswana secondary schools: an investigation. *Educational Management Administration & Leadership*, 38 (1), 71-87.
- Nelson R, Ehren M (2014) Review and synthesis of evidence on the (mechanisms of) impact of school inspections. <http://schoolinspections.eu/wp-content/uploads/downloads/2014/02/Review-and-synthesis-of-evidence-on-the-mechanisms-of-impact-of-school-inspections.pdf> (accessed 1 April 2014).
- Opoku-Asare, N. (2006). Quality Control for Effective Basic Education in Ghana. *Journal of Science and Technology*, 26 (3).
- Pawson R (2006) Evidence-based policy: a realist perspective. London: Sage.
- Pawson R, Greenhalgh T, Harvey G, Walshe K (2005) Realist review: a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research and Policy*, 10: 21-34.
- Santiago, P., McGregor, I., Nusche, D., Ravela, P., & Toledo, D. (2012). *OECD Reviews of Evaluation and Assessment in Education: Mexico 2012*: Paris and Washington, D.C.: Organisation for Economic Co-operation and Development.
- Scheerens J (2000) Improving school effectiveness: fundamentals of educational planning. Paris: Unesco, International Institute for Educational Planning.
- Uwazi. (2009). *When School Inspection Doesn't Deliver: Highlights from the CAG audit of the Secondary Schools Inspection Programme in Tanzania*. Tanzania Africa: Uwazi InfoShop at Twaweza.
- Wanzare, Z. (2002). Rethinking School Inspection in the Third World: The case of Kenya.
- World Bank. (2010). *Improving Education Management in African Countries*: World Bank.
- World Bank (2011) Learning for all: investing in people's knowledge and skills to promote development: World Bank education strategy 2020. <http://sdindicators.org/wp-content/uploads/2013/07/World-Bank-Education-Strategy.pdf> (accessed 7 June 2014).
- Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R (2013) RAMESES publication standards: realist syntheses. *BMC Medicine*, 11: 21. <http://www.biomedcentral.com/content/pdf/1741-7015-11-21.pdf> (accessed 7 June 2014).