



What Works in Education?

Using Evidence to Improve Education

13

January 2019

Education inspection: which models work best?

Álvaro Choi

Most education systems in developed countries have an inspection service and although their models are not identical, they all seek to improve the quality of education. The question is: Do they achieve it? This review approaches education quality through student performance and gathers evidence relating inspection activity with academic results. In the following pages we will answer this question and others, such as: Should inspection activity play an advisory role or should it also entail other types of implications? Is it better to publish evaluation reports or not? What kind of feedback is most effective? And finally, should inspection activity be strengthened?

“For too long, education has been subject to inertia and based on traditions, and educational changes have been grounded in unfounded intuitions and beliefs. The ‘What Works’ movement irrupts into the world of education with a clear objective: to promote evidence based policies and practices. [Ivàlua](#) and the [Jaume Bofill Foundation](#) have come together to push this movement forward in Catalonia.”



What Works in Education?

Using Evidence to Improve Education

Education inspection: which models work best?



Álvaro Choi

Álvaro Choi is a Professor in the University of Barcelona's Department of Economics and a researcher at the Barcelona Institute of Economics (IEB). His main lines of research are the economics of education and the evaluation of public policies.

Motivation

The education systems of most developed countries have an inspection service. This is an instrument to control and promote the quality of education that is external to the schools and aims to improve the level of knowledge of their students, in addition to other objectives.

There are many inspection models with very different characteristics [1]. In recent decades, the various education inspection models have observably diversified. This diversification is mainly due to three lines of tension: between the professionalization and politicization of the body of inspectors, between the administrative role they play and their educational role and, regarding teaching efforts, between supervision and advisory work [2].

Every autonomous community in Spain enjoys freedom in developing its education inspection system. Catalonia has chosen an inspection model that is regulated by Articles 178, 179, 180 and 181 of the Education Law of Catalonia (Llei d'educació de Catalunya), which establishes a professional body of inspectors that oversees, advises and evaluates teaching and administrative efforts, controlling schools and education services. Therefore, returning to the three aforementioned lines of tension, inspection combines both administrative and educational functions, with advisory

work becoming more important in the latter [3] [4]. However, as is the case in most education systems, its effectiveness in influencing academic performance has not yet been evaluated.

Therefore, an exhaustive review of the existing literature will yield information helping us to answer the following question: What elements characterize the most effective education inspection systems? By identifying those elements, we will be able to make recommendations for education policy that lead to better student performance.

What type of intervention are we talking about?

Education inspection is an external system for evaluating and guiding schools' activity through a body of inspectors belonging to a local or national education authority. As such, works that analyze the schools' internal assessments fall outside the scope of this review.

In fact, inspection activity consists of a set of very dissimilar activities. First, it enforces compliance with legislation and minimum quality standards at the schools; as such, it serves as a mechanism of accountability. This is the classic function of inspection activity. Yet it is also a mechanism to promote organizational improvement and innovate in education practices. Finally, inspection may participate in evaluating the education system and the professionals working in it [5]. Modern inspection systems tend to cover these functions to a certain extent.

In fact, inspection activity enforces compliance with legislation and minimum quality standards at the schools yet it is also a mechanism to promote organizational improvement and innovate in education practices.



Apart from the various roles that they may play [6], inspection systems are also distinguished by a set of essential **characteristics** [7]:

- Whether the assessment focuses on results or on educational processes.
- The types of consequences of the inspection (high-stakes versus low-stakes; when the stakes are high, the inspection can have punitive consequences, such as getting fired, whereas low-stakes inspections focus on providing information and making suggestions for improvement in the schools).
- Whether the inspection reports are published.
- The frequency and intensity of the visits.

Visiting the schools is one of the primary mechanisms for inspectors to gather information about the level of compliance with legislation, the quality of the service provided and student performance. Direct observation, documentary analysis and interviews with teachers are some of the most common practices during these visits. Education and school staff receive feedback from the inspectors based on the information they collect during them. This feedback is one of the most important mechanisms that inspections have for influencing student performance [8].

The literature makes a distinction between three types of **effects** that inspection considers priorities, although in some cases the line that separates them is not so clear [9]:

- a. **Conceptual** effect: ability to influence teachers and principals to perceive the need to think about educational processes.
- b. **Instrumental** effect: effective changes implemented in the school based on the information and recommendations provided by the inspectors.
- c. **Symbolic** effect: ability of the inspection activity to modify or confirm the teachers' and principals' previous ideas and convictions.

Moreover, the literature has also called attention to **externalities** (the undesired effects that inspection has on third parties) that may affect student performance. These externalities may be classified as [10]:

- a. **Voluntary strategic behavior:** modification of behavior in order to provide a good image of the school and achieve a positive result in the inspection. Some examples include organizing an event during the days when the visits take place or holding meetings with the teachers to tell them how they should conduct themselves during the visit.
- b. **Involuntary strategic behavior:** distraction from regular school activity resulting from the notification of a visit or resulting from an inspection. Some examples include altering scheduled school activities to prepare for the inspection or enhancing the importance of teaching in the classes assessed especially by the inspectors.
- c. **Emotional impact:** stress and anxiety for the school staff as a result of the inspection. This effect may appear before or after the inspection and is mainly observed if the inspection may lead to punitive measures.

As we can see, an inspection's various concurrent direct and indirect effects may lead to a very uneven impact on students' academic performance. Therefore, this review will refer to the impact of different combinations of these effects on academic performance.

Questions influencing the review

The review examines the literature that analyzes the link between education inspections and students' academic performance. Specifically, it aims to answer the following questions:

1. Does inspection activity influence academic performance?
2. If so, through which channels? What kind of feedback is most effective in improving student outcomes? Does the frequency of inspections matter? Should inspection results be published?

The answer to these questions should allow us to identify the most effective inspection models and make recommendations on education policy applicable to the Catalan education system.

Reviewing the evidence

Reviews and studies considered

The review of the evidence was conducted as follows. First, we explored the meta-analyses in the repositories of prestigious institutions like the Institute of Education Sciences through its What Works Clearinghouse initiative and through the Education Endowment Foundation. Despite the informational wealth held in these repositories, we did not find any content relevant to our subject of study. Second, we searched online for meta-analyses relating inspection activity and academic performance. We only found a few reviews and given the publication date of the latest one (2014), we decided to complement them with an ad hoc review of works published since 2014.

With regard to the ad hoc review, we searched online for empirical works in Spanish or English published between 2000 and 2018 that used an experimental or quasi-experimental methodology to evaluate the effects of inspection on the academic performance of primary and secondary school students. Thus, works that did not provide causal evidence were excluded. It is important to clarify that we conducted a systematic review of the literature, not a meta-analysis in the strictest sense of the word. We found 12 studies in total. Even though it relies on a small sample of studies, this review is actually the most extensive so far.

Even though it relies on a small sample of studies, this review is actually the most extensive so far.



The scarcity of the studies identified is related to the complexity involved in determining the relationship between inspection activity and academic performance. First, both the availability and quality of data related to inspection activity vary widely from country to country. Second, inspection systems tend to be centralized, so education systems under the same educational jurisdiction are subject to similar treatment (meaning that there is hardly any variation in how the schools are “treated”),

which makes it difficult to conduct impact assessments for comparing the effectiveness of different models. Third, the potential effect of inspection activity on student performance is fundamentally indirect (with all the methodological complexity that this entails), since its effectiveness depends on the ability to transform or modify the behavior and practices of other education professionals, among other issues. Thus, in practice there is complex interaction between inspection systems, inspectors and schools, the latter including students, teachers and administrators, all of which have dissimilar characteristics [11]. As such, identifying the impact of inspection activity provides little information about the mechanisms that can explain it [7].

Finally, it is difficult to isolate the effect of inspection activity from the effects of other related policies or practices. Some examples of this would be the school's degree of autonomy, self-assessments performed at the school and the existence of standardized external evaluations, as well as how they are used and the visibility they are given.

Regarding this last issue, although there are similarities between inspection activity and the introduction of other mechanisms of accountability at schools, we considered it appropriate to discard research focused on analyzing the effects of implementing other mechanisms of accountability at schools (mainly based on external assessments of student performance) [12-14] since we could not isolate the effect of the inspection activity from the effect of the other types of external assessments.

a) Review of the existing meta-analyses and reviews

Four relevant reviews were identified (Table 1). Of them, only two [15] [16] focused specifically on analyzing the relationship between inspection and academic performance. In fact, they are the only two reviews we found that systematically examine the literature on this issue. The other two [1] [10] look into issues related to school inspection and analyze some studies linking inspection and academic performance as part of them. These four reviews collected little causal evidence. Thus, the review that provides the most information based on quasi-experimental evidence is the one by Nelson and Ehren (2014) [16] (eight of the 14 quantitative articles included in it). As can be observed in Table 1, most of the studies covered by the reviews refer to only two countries: the United Kingdom and the Netherlands.

Table 1.
Reviews of the effect of inspection activity on student performance (N = 4)

Authors	Number of studies	Place and period	Effect indication and size
Klerks (2013) [15]	3	The United Kingdom and the Netherlands. Since 2000.	Positive, but smaller effect.
Nelson and Ehren (2014) [16]	8	The United Kingdom and the Netherlands. Since 2000.	The six oldest studies tend to show a very small or no effect. Studies after 2010 do identify a positive effect.
OECD (2013) [1]	2	The United Kingdom and the Netherlands. No restrictions.	Positive effect of indeterminate size on performance.
De Wolf and Janssens (2007) [10]	2	The United Kingdom. No restrictions.	Mild positive effect.

Note: The “Number of studies” column refers to the number of quantitative studies included in each review that employed an experimental or quasi-experimental methodology.

Source: Author’s creation

Some of the main conclusions that can be drawn from our examination of the previous reviews include:

- The existing causal evidence for the direct effect of inspections is too small for us to fully confirm the effect that education inspections have on academic performance. This finding, which some authors support [10] [17], will be explored in greater detail in the second part of this review.
- In any case, there seems to be a mild positive (though still unquantified) effect on academic performance that is associated with inspection activity.
- This effect has only been evaluated over the short term (during the following year or, at the most, two years after the inspection). This is a clear limitation, since inspections should have long-term effects.
- Solid evidence about the possible externalities of inspection activity is very limited. There is only some evidence related to involuntary strategic behavior and mounting stress among the administration and teaching staff as a result of the inspection visits.
- Studies that describe the inspection’s effect on issues like improvements in teaching practice, organizing the school and changing teacher behavior are based on analyses that do not allow us to identify causal relationships. These analyses seem to indicate a positive association between inspection activity and school performance [8] [18].

b) Specific assessments of special interest

The low number of studies that we identified made it advisable to widen our review with an additional systematic one. Virtually all the articles in [Table 2](#) refer to developed European countries (the United Kingdom and the Netherlands). There are two reasons for this: the greater availability of data and the differences between

the systems of accountability in European schools and those of the United States. In recent years, the accountability systems have been based on the students' performance in external evaluations, so they fall outside the scope of this review. Five of the 12 studies analyze the situation in the United Kingdom [19-23]. All the articles employ quasi-experimental methodologies (differences in differences or panel models) to identify causal effects of inspection activity on students' academic performance. The estimated effects are usually short-term (three years at the most).

It should be emphasized that the literature focuses primarily on analyzing the impact of the number of inspections on student performance, not on whether there is an inspection or not, which would be difficult to assess in European countries since practically all of them have an inspection system. In reality, the lack of information is one of the factors explaining why there are so few studies that can identify the causal effects of inspections.

Box 1.

The impact that implementation of an inspection system has on performance

Identifying the effect of implementing an inspection system is complex, since they are usually set up across the country at the same time and therefore there is no variation possible to evaluate. Furthermore, most developed countries already have an inspection system. However, one of the articles reviewed (Schueler, 2016 [24]) studies the impact of implementing an inspection system compared to not doing so. It evaluates the effect of replacing an “inspection” system controlled by the Catholic Church with a centralized inspection system in Prussia during the second half of the 19th century, using this historical perspective to determine the impact of the reform on enrollment rates. The results indicate that the introduction of a centralized inspection system contributed to higher rates of enrollment. The study also discovered an important lesson: this rise in enrollment was milder in areas where the new inspection system was not supported by the local community and the schools.

For further information:

Schueler, R. (2016). “Centralized Monitoring, Resistance, and Reform Outcomes: Evidence from School Inspections in Prussia”, *Ifo Working Paper*, no. 223.

Can inspections improve students' educational outcomes?

Our review allows us to state that the general effect that inspections have on performance is positive. Various studies indicate that performance improves in the years immediately after the inspection, both in primary [22] [25] and secondary school [19] [24] [26].

It is remarkable that several studies have found that inspections can have dissimilar effects [19] [21] [22] [27]. More specifically, they indicate the schools and students whose performance improved after inspections were found at opposite ends of the distribution of academic results, meaning that they had both the lowest results and the highest.

What are the most effective inspection models like?

Most of the articles estimate the effects that inspection visits have on the performance of the schools' students by resorting to some standardized test. However, some studies delve into the characteristics of the inspection that may influence its effectiveness. Despite the fact that the external validity of these articles (the ability to transfer the conclusions they reach to other environments) is limited by the fact that each refers to a unique education system, the rigor with which they were created merits a discussion of the results.

Before describing the characteristics of the most effective inspection models, it should be noted that the studies reviewed do not allow us to differentiate whether the inspection activity should focus on analyzing the results of the schools or, on the contrary, it should also evaluate the educational processes, since they focus on education systems where the inspection assesses both the results and the educational processes.

Low-stakes versus high-stakes

The first important issue for implementing an inspection system is deciding on whether it will be a high-stakes or low-stakes system. High-stakes systems are more focused on student outcomes and tend to have similar consequences for teachers or school principals. Low-stakes systems tend to be focused on the educational process, though not necessarily. The advisory function of the inspection predominates in the low-stakes system.

Table 2.
Results of impact studies on principals (N = 12)

Authors	Database	Location	Technique	Variable of analysis	Outcome variable	Effect indication and size
Allen and Burgess (2012) [21]	Panel of students and schools (National Pupil Database) 2002-2011.	England.	Discontinuous regressions.	School inspection and effect on leadership.	Standardized student academic results.	Obvious improvements after two or three years (0.1 SD the year of the inspection and higher from the second year on).
De Hoyos et al. (2017) [27]	Census of primary schools and results in standardized tests from 2006 to 2013.	State of Colima (Mexico).	Differences in differences and discontinuous regressions.	Participation in the Specific Attention Program.	Results in the ENLACE (Academic Achievement Evaluation in School Centers).	0.12 SD several months after the feedback.
Ehren and Shackleton (2016) [28]	Data panels (authors' creation) of primary and secondary schools, 2011-2013.	The Netherlands.	Panel models (random effects).	Two different types of inspection: basic and intensive (in schools with a poor previous evaluation).	Opinion of the teachers and principals on the different effects of the inspection.	No effect on the school's effectiveness.
Gustafsson (2014) [29]	Panel of primary and secondary schools, 2011-2013.	Sweden.	Growth model.	Inspector's visit.	Indications about different aspects of the school's operations.	Positive effect on improving teachers' teaching skills and on incorporating measures that boost teaching effectiveness.
Hussain (2012) [22]	Panel of primary schools and students.	England.	OLS and differences in differences.	Time of the inspection.	Get a failing result in the inspection.	Improvements in standardized tests (0.1 SD).
Kemethofer (2017) [30]	Data panels (author's creation) of primary and secondary schools, 2011-2013.	Sweden and Styria (Austria).	Growth model with latent variable.	Inspector's visit.	Indication of effectiveness and of teaching conditions.	Small but positive effect in both countries.
Luginbuhl et al. (2009) [25]	Primary school, panel of students and schools, 1999-2002.	The Netherlands.	Panel models (fixed effects).	Inspector's visit.	Standardized student academic results.	Improvements during the next two years Visits lasting two or three days improve by an SD of 2-3%.
Pietsch et al. (2014) [26]	Panel of secondary school students, 2010-2011.	Hamburg (Germany).	Difference in differences.	Inspector's visit.	Results in the university entrance exam.	Approximately 0.2 points over 10 in the test.
Rosenthal (2004) [20]	Panel of students and schools, 1994-1998.	England.	Panel models (fixed effects and random effects).	Inspector's visit.	Percentage of 15-year-old students who obtain high scores on the GSCE. (General Secondary Education Certificate).	Negative effect in the year of the inspection; drop of approximately 2%; a posteriori, none.
Schueler (2016) [24]	Census and education inspection data, 1864-1886.	Prussia.	Difference in differences.	Introduction of a centralized inspection system.	Enrollment rate.	Positive effect, but only in areas supported by the local community.
Shaw et al. (2003) [19]	Panel of secondary schools and students, 1992-1997.	England.	Multi-level models (fixed effects).	Inspector's visit.	Percentage of 15-year-old students who obtain high scores on the GSCE. (General Secondary Education Certificate).	Mixed results based on the type of school Slight effects (2%) Positive effect for the best and worst schools.
Sims (2016) [23]	Panel of primary and secondary school teachers, 2010-2013.	England.	Difference in differences.	Negative assessment of the school.	Teacher rotation.	Negative assessments increase the teacher rotation rate by 25%.

Note: SD = standard deviation

Source: Author's creation

First, it is worth noting that both types of inspection can lead to better academic performance. The studies reviewed indicate that for low-stakes systems, changes in behavior occur after the inspection, whereas they come before or after the inspection in high-stakes systems (Sweden, England in the 1990s) [20] [23].

For low-stakes systems, changes in behavior occur after the inspection, whereas they come before or after the inspection in high-stakes systems.



Thus, in high-stakes models, teachers and principals may change their behavior during the year prior to the inspection, which can lead to lower academic performance [20]. High-stakes models tend to generate a higher level of stress among teachers and principals [20]. There is another undesired effect linked to this type of inspection: negative assessments may lead to a higher rate of teacher rotation at the school for fear of possible sanctions [23]. In summary, low-stakes inspections seem to promote the schools' capacity for self-assessment once they have been evaluated, whereas high-stakes systems seem to promote the capacity for improvement before inspections [30].

Box 2.

A study that compares the effects of low-stakes and high-stakes models

The study conducted by Kemethofer et al. (2017) [30] is of great interest, since it compares the effectiveness of two different inspection models in two different places: Austria (Styria) and Sweden. In Austria, the stakes of the inspection were low, as it was purely informative in nature and had no punitive implications. The frequency of the visits was between two and four years and they could last up to three days with one, two or three inspectors, depending on the size of the school. In Sweden, on the other hand, inspections (conducted every five years for one or two days by two inspectors) could lead to sanctions for teachers, so the stakes were high. The results were published (website) and the schools had three months to implement an action plan negotiated with the Swedish education authorities. Using their own database, they get inspections to have a slightly positive effect on teaching quality in both cases. Through auxiliary analysis, they estimate that this effect is primarily due to the fact that the schools accept the feedback from the inspection. Acceptance is greater in Austrian schools. In Austria, inspections improve indicators related to self-evaluation after the inspection activity, whereas in Sweden inspections influence the indicators related to the schools' "capacity for improvement" before they take place. In both cases, inspections translate into improvements in the quality of education, although there is greater activity in schools in Sweden before an inspection (to avoid sanctions) and in Austria after an inspection, which we understand as an opportunity to introduce long-term changes.

For further information:

Kemethofer, D.; Gustafsson, J. E. and Altrichter, H. (2017). "Comparing effects of school inspections in Sweden and Austria". *Educational Assessment, Evaluation and Accountability*, vol. 29, no. 4, pp. 319-337.

Feedback

In low-stakes and high-stakes systems, the **feedback** that the school receives from the inspection is crucial [15]. The capacity for improvement connected to this feedback depends on two fundamental factors:

- The first is the **school's acceptance of this information**. Knowledge of the environment and of the specific characteristics of the school appears to be key [19] [30]. It should be added that there seems to be greater acceptance of feedback in low-stakes systems, though the evidence is inconclusive [30]. In any case, it is good for the school and the inspection to come to an agreement on the recommendations [11]. This is related to the importance of creating a bond of trust between the inspectors, the principals and the teachers so that the latter two are receptive to the findings of the inspection [18]. It is interesting to observe that over the short term, high-performing schools seem to accept the feedback better, although the threat of possible sanctions means that all schools end up accepting the recommendations [28].

It is good for the school and the inspection to come to an agreement on the recommendations. This is related to the importance of creating a bond of trust between the inspectors, the principals and the teachers.



- The second factor on which the transformational ability of inspection feedback for schools depends is the **type of information provided**. Schools usually have abundant information and can find self-assessment useful for identifying problems, but not necessarily for solving them [19]. Therefore, it is important that the inspector's feedback include specific suggestions for action [18] [31]. Naturally, for these recommendations to be accepted, they must be realistic in light of the school's resources and circumstances [32]. This brings us back to the previous point: the importance of the inspector's knowledge of the school and its environment and the creation of a bond of trust.

It is important that the inspector's feedback include specific suggestions for action.



- A related issue hinges on **who should receive the feedback from the inspection within the schools**. An inspection's ability to influence improvements at a school occurs mainly through the feedback it provides to the administration [15]. However, the effect of the feedback provided to teachers following direct observation in the classroom is not very clear. While some authors assert that one way that inspectors have a positive impact on performance is reflected in changes in teaching activity prompted by the feedback provided to the schools based on the results of the evaluations [27], others claim that inspections especially modify the behavior of principals, but not teachers [28].

An inspection's ability to influence improvements at a school occurs mainly through the feedback it provides to the administration.



Box 3.

A study on the importance of the type of information provided in the inspection and academic performance

The study conducted by de Hoyos et al. (2017) [27] did not exactly evaluate inspection activity per se, but the participation of the schools that perform poorly in a program to improve educational results, structured around a standardized diagnostic test (ENLACE). The only consequences of this test are the ability to identify problems in the school. The results are made public. Schools that perform poorly in ENLACE received (and this is the important issue for this review) at least three visits per month by a technical advisor, who was always the same one for each school, helped them to develop a school plan and advised them on teaching practices. Those who benefitted most from this program were the students who had performed better previously, leading the authors to suggest that more comprehensive inspections are probably needed to enhance the performance of students with greater difficulties. While it is true that the results of the study conducted by de Hoyos et al. (2017) combine the effects of the advisory role of inspections and the benchmarking effect of the external evaluation, this experience underscores the importance not only of providing information to the schools, which they probably had before the improvement program, but also of supporting them and helping them to turn this information into improvement. Cooperation, shared responsibility and trust between the advisor-inspector-administration and schools seem key to the program's success. Therefore, a diagnostic test without punitive effects may be enough to activate improvement processes.

For further information:

De Hoyos, R.; Garcia-Moreno, V. A. and Patrinos, H. A. (2017). "The impact of an accountability intervention with diagnostic feedback: Evidence from Mexico". *Economics of Education Review*, vol. 58, issue C, pp. 123-140.

Publication and accountability

Regarding the controversial issue of **publishing inspection reports**, most studies refer to education systems in which the content of these reports is made public. Studies that deal with the effects of publishing school results are not usually linked to inspection of the school, according to the restricted definition employed for this review, but to the study of other external systems for evaluating the education system (basically, standardized tests). An analysis of the situation in England [22] finds that the inspector's assessment provides families with valuable information about school quality. However, a study focused on the Netherlands [28] clarifies that families and school councils usually focus on the results of the inspection during the following year, but then forget about them. In any case, a noteworthy study evaluated the effect of devolving powers of inspection to Wales in 1999 [33]. Unlike England, Wales decided to stop publishing inspection results based on the view that it had a negative impact on the performance of Welsh schools. Note that even though previous studies show that the publication of inspection results has a positive effect on the average academic level, they do not consider the possibility that those effects are different based on the school profile or calibrate the potential effects on other relevant spheres of education.

Previous studies show that the publication of inspection results has a positive effect on the average academic level.



Inspection duration and frequency

Finally, evidence related to the **duration and frequency of the inspections** is also scarce. However, the different experiences evaluated by the studies gathered allow us to take away some related indications. The scanty evidence on this point finds that a higher number of visit days has a positive effect on school performance. One study based on panel data for primary schools in the Netherlands (where visits are unexpected and random) concludes that school inspections are a cost-effective policy, in comparison with other alternatives [25].

The frequency (number and intensity) of the visits has a positive effect on the performance of schools that have poor academic results [22] [34].

In any case, there seems to be a certain consensus about the need to intensify inspection activity due to the growing autonomy of schools, not only to guarantee compliance with the regulations and minimum standards set by the education system, but also to provide better and greater support to schools that are increasingly diverse [1] [35] [36].

There seems to be a certain consensus about the need to intensify inspection activity due to the growing autonomy of schools.



Summary

Most modern education systems have an inspection service in charge of boosting the quality of education, among other goals. Therefore, effective inspection systems should be able to improve students' academic performance. This literature review shows that inspection activity can effectively improve academic performance, though its impact seems to be modest. The evidence also suggests that there is no single effective model.

There are not many studies that establish causal relationships between inspection activity and academic performance. This is because these studies require a large amount of data and it is difficult to identify the impact, since inspections operate through indirect channels (mainly by influencing principals and teachers); moreover, inspection activity is similar in the schools of different countries. The ad hoc review of the studies that employ quasi-experimental techniques allows us to draw the following conclusions, highlighted in a summarized way in [Table 3](#):

- First, the inspection activity may have a **positive effect** on academic performance. However, the studies conducted thus far focus on the short term (three years maximum), whereas some effects of the inspection may show up in the medium or long term. There are different inspection models that may lead to better performance, though some of them involve certain risks, like modification of a school's usual activities before inspections or the added stress for principals and teachers, which can lower that positive effect.
- Second, the positive **effect** of inspections is **not uniform**. The schools that benefit from inspection activity the most are those with either very good or very poor results.
- Third, they may improve academic performance through both **high-stakes** and **low-stakes** models. The choice of the type of model to use is shaped by cultural and political factors. It should be noted, in any case, that for the inspection system to be a success, it is important that education professionals accept the recommendations.
- Fourth, the **intensity and frequency** of the visits of the inspection team seem to be positively related to school performance. Longer and more frequent visits have a positive impact on school performance. Inspection activity is also more necessary the greater the school's level of autonomy.
- Fifth, the **feedback** resulting from the inspection that is provided to the schools is key to the effectiveness of inspection activity. The information provided to the schools should not only identify problems, but also provide specific suggestions for action. In any case, the school's ability to accept the information, mainly by the principals, will depend on the level of trust established between the school and the inspection team. In this sense, the inspection team's level of knowledge about the school and its environment is important.
- Finally, there is evidence seeming to indicate that the **publication of inspection reports** has a positive effect on the school's average academic level. In any case, these results leave out the possibility that there may be dissimilar effects among different profiles of schools, in addition to the possible effects of other different variables on academic performance.

Table 3.
Strengths and weaknesses of the inspection activity

Strengths	Weaknesses
<ul style="list-style-type: none"> • School inspection activity influences academic performance in the short term. Different models can lead to better performance. • The impact of inspection activity is greater in low-performing and high-performing schools. • The intensity and regularity of the inspectors' visits is positively related to academic performance. • Performance may be improved through either high-stakes or low-stakes inspection models. • The reports of the most effective inspection models contain specific suggestions for action. • The trust between the schools and the inspection team boosts the effectiveness of the inspection activity. • Knowledge of the characteristics of the school and the context help to improve the quality of the inspection. • Publishing the inspection reports may have positive effects on academic performance. 	<ul style="list-style-type: none"> • Causal evidence is scarce. The study is based on a small group of articles focused on an even smaller number of countries. As such, the results cannot be directly exportable to other settings. • The existing literature analyzes the short-term effects of the inspection, neglecting the possibility that some reforms made as a consequence of the inspection may take shape over the long term. As occurs in other public programs, it is also possible that the effect of the inspection may fade over time. • There is virtually no causal evidence for the types of activities and channels that make an inspection more effective. • There is no information about what characterizes the most effective inspectors. • There is hardly any evidence for the effect of implementing an inspection system compared to the alternative of not having an inspection system. • There is no evidence for possible dissimilar effects among schools caused by publishing inspection reports.

Implications for practice

The results presented throughout this review allow us to take away some guiding principles for designing effective inspection systems. The fact that inspections have more of a positive effect in schools located at both ends of the spectrum of scores (with the highest and lowest scores) opens the door to an interesting and cost-effective way to reduce inequalities between schools. Thus, intensifying inspection activity in schools with poorer academic performance may not only increase the average performance of the education system, but also improve equity within it. The finding that more intensive and regular inspections have a positive impact on academic results provides clear indications about the type of inspection activity.

One important result is related to the effectiveness of different inspection models. Empirical evidence indicates that different inspection models can lead to similar degrees of improvement in academic performance. The differences between models (high-stakes versus low-stakes models; whether they publish their reports or not; how much importance they give to results in external evaluations) tend to follow the different socio-political, educational and cultural development of the various countries. The fact that the literature points to the importance of schools' acceptance of the information and recommendations resulting from their inspections suggests a certain difficulty (or, at least, the existence of transition periods) in successfully transferring inspection models from one country to another. For example, the adoption of high-stakes models, which are applied and work effectively in countries

such as United Kingdom or Sweden, could have negative effects, at least in the short term, as some of the reviewed works indicate.

The importance of creating bonds of trust between the school and the inspection team may lead to the recommendation to grant a certain degree of stability to inspection teams working the schools. Even though no studies were identified that specifically analyze the features and characteristics of the most effective inspectors, our review of the literature allows us to conclude that knowledge of a given school's characteristics and means, as well as of the socioeconomic environment of the area where it is located, is crucial for them to be able to make the right recommendations. In this sense, the feedback that the inspectors provide to the schools must include specific suggestions for solving specific problems, in addition to adapting to the school in question. Inspection activity seems to prompt self-assessment processes in schools, though these often face limitations when it comes to dealing with the problems and challenges identified. The inspectors' training is relevant in this regard, as they should be aware of the initiatives and activities carried out both in schools similar to those they inspect and in other environments. Tools for synthesizing evidence could be useful in this regard, such as the What Works Clearinghouse initiative or the set of studies that include this review.

With regard to the advisability of publishing the results of the inspections, it should be borne in mind that inspection systems are primarily external systems of evaluation. As such, publishing the results of the inspections is subject to the same potentialities (more information for families or increase in the school's average performance), but also to similar risks (rise in inequalities and suppression of activities not evaluated by the inspection, among others) as occur when publishing the students' results in external evaluation tests. The recent experience with the school evaluation system in Catalonia (known as the *Avaluació anual de centre*, or AVAC), which combines the results into external evaluations, the information of the school's system of indicators (*Sistema d'indicadors de centre*, or SIC) and the analysis of the inspection, is an attempt to promote those opportunities and limit risk.

In any case, we are still far from being able to evaluate the impact of the inspection system on Catalan students' educational performance (in short, to evaluate the evaluators). To do so, we would need to organize the means and resources to evaluate the inspection system in Catalonia. For this purpose, it would be necessary to provide information about the inspectors' activities, profiles and professional careers and link them to the schools with which they have worked. Another promising field for analysis would be the study of differences in inspection practices among regional services, or even within them. Again, this exercise would require an amount and type of information that is very difficult to access at the present time.

Bibliography

- [1] OECD (2013). *Synergies for better learning: an international perspective on evaluation and assessment*. Paris: OECD Publishing.
- [2] Aguerrondo, I. (2013). "El rol de la supervisión educativa en la gestión de las políticas públicas". *Educar*, no. 49, vol. 1, pp. 13-27.
- [3] Generalitat de Catalunya (2011). *Pla director de la Inspecció d'Educació 2011-2015*. Barcelona: Generalitat de Catalunya.
- [4] Generalitat de Catalunya (2017). *Pla director de la Inspecció d'Educació 2017-2020*. Barcelona: Generalitat de Catalunya.
- [5] de Grauwe, A. (2003). "Les réformes des services d'inspection. Modèles et ideologies". *Revue Française de Pédagogie*, no. 145, pp. 5-20.
- [6] Whitby, K. (2010). *School inspections: recent experiences in high performing education systems*. Reading, England: CfBT Education Trust.
- [7] Gustafsson, J. E. et al. (2015). "From inspection to quality: Ways in which school inspection influences change in schools". *Studies in Educational Evaluation*, no. 47, pp. 47-57.
- [8] Dederling, K. and Müller, S. (2011). "School Improvement through Inspections? First empirical insights from Germany". *Journal of Educational Change*, no. 12, vol. 3, pp. 301-320.
- [9] Rossi P. H.; Lipsey M. W. and Freeman H. E. (1999). *Evaluation: A systematic approach*. Thousand Oaks, CA: Sage Publications, Inc.
- [10] de Wolf, I. F. and Janssens, F. J. (2007). "Effects and side effects of inspections and accountability in education: an overview of empirical studies". *Oxford Review of Education*, no. 33, vol. 3, pp. 379-396.
- [11] Ehren, M. and Visscher, A. J. (2008). "The relationship between school inspections, school characteristics and school improvement". *British Journal of Educational Studies*, no. 56, vol. 2, pp. 205-227.
- [12] Hanushek, E. A. and Raymond, M. E. (2005). "Does School Accountability Lead to Improved Student Performance?". *Journal of Policy Analysis and Management*, no. 24, vol. 2, pp. 297-327.
- [13] Figlio, D. N. and Loeb, S. (2011). "School accountability" in Hanushek, E. A.; Machin, S. and Woessmann, L. (eds.), *Handbook of the Economics of Education*. Amsterdam: North Holland.
- [14] Feng, L.; Figlio, D. and Sass, T. (2018). "School accountability and teacher mobility". *Journal of Urban Economics*, no. 103, pp. 1-17.
- [15] Klerks, M. C. J. L. (2013). "The effect of school inspections: a systematic review". *School Improvement*, pp. 2-32.
- [16] Nelson R. and Ehren, M. (2014). *Review and synthesis on the (mechanisms of) impact of school inspections*. Twente: University of Twente.
- [17] Spink, J.; Cassity, E. and Rorris, A. (2017). *What Works Best in Education for Development: A Super Synthesis of the Evidence*. Canberra: Department of Foreign Affairs and Trade.
- [18] Matthews, P. and Sammons, P. (2005). "Survival of the Weakest: The Differential Improvement of Schools Causing Concern in England". *London Review of Education*, no. 3, vol. 2, pp. 159-176.
- [19] Shaw, I.; Newton, D. P.; Aitkin, M. and Darnell, R. (2003). "Do OFSTED inspections of secondary education make a difference to GCSE results?". *British Educational Research Journal*, no. 29, vol. 1, pp. 63-75.
- [20] Rosenthal, L. (2004). "Do School Inspections Improve School Quality? OFSTED Inspections and School Examination Results in the UK". *Economics of Education Review*, no.23, vol. 2, pp. 143-151.
- [21] Allen, R. and Burgess, S. (2012). *How should we treat under-performing schools? A regression discontinuity analysis of school inspections in England*. Bristol, UK: CMPO.
- [22] Hussain, I. (2012). *Subjective performance in the public sector: evidence from school inspections*. London: London School of Economics and Political Science. Centre for Economic Performance.
- [23] Sims, S. (2016). "High-stakes accountability and teacher turnover: how do different school inspection judgements affect teachers' decisions to leave their school?". *UCL-IOE DoQSS Working Papers*, no. 16-14.
- [24] Schueler, R. (2016). "Centralized Monitoring, Resistance, and Reform Outcomes: Evidence from School Inspections in Prussia". *Ifo Working Paper*, no. 223.
- [25] Luginbuhl, R.; Webbink, D. and de Wolf, I. (2009). "Do Inspections Improve Primary School Performance?". *Educational Evaluation and Policy Analysis*, no. 31, vol. 3, pp. 221-237.
- [26] Pietsch, M.; Janke, N. and Mohr, I. (2014). "Führt Schulinspektion zu besseren Schülerleistungen? Difference-in-Differences-Studien zu Effekten der Schulinspektion Hamburg auf Lernzuwächse und Leistungstrends". *Zeitschrift für Pädagogik*, no. 60, vol. 3, pp. 446-470.

-
- [27] de Hoyos, R.; Garcia-Moreno, V. A. and Patrinos, H. A. (2017). "The impact of an accountability intervention with diagnostic feedback: Evidence from Mexico". *Economics of Education Review*, vol. 58, issue C, pp. 123-140.
- [28] Ehren, M. C. and Shackleton, N. (2016). "Risk-based school inspections: impact of targeted inspection approaches on Dutch secondary schools". *Educational Assessment, Evaluation and Accountability*, no. 28, vol. 4, pp. 299-321.
- [29] Gustafsson, J. E. (2014). *Impact of school inspections on teaching and learning in primary and secondary education in Sweden. Technical report ISI-TL*. Gothenburg: University of Gothenburg.
- [30] Kemethofer, D.; Gustafsson, J. E. and Altrichter, H. (2017). "Comparing effects of school inspections in Sweden and Austria". *Educational Assessment, Evaluation and Accountability*, no. 29, vol. 4, pp. 319-337.
- [31] Penninckx, M.; Vanhoof, J.; De Maeyer, S. and Van Petegem, P. (2014). "Exploring and explaining the effects of being inspected". *Educational Studies*, no. 40, vol. 4, pp. 456-472.
- [32] Schildkamp, K. and Visscher, A. (2010). "The use of performance feedback in school improvement in Louisiana". *Teaching and Teacher Education*, no. 26, vol. 7, pp. 1389-1403.
- [33] Bevan, G. and Wilson, D. (2013). "Does 'naming and shaming' work for schools and hospitals? Lessons from natural experiments following devolution in England and Wales". *Public Money & Management*, no. 33, vol. 4, pp. 245-252.
- [34] Atuhurra, J. F. (2014). "An empirical analysis of the impacts of the universal primary education policies on educational performances in East Africa: a comparative study of Uganda's UPE and Kenya's FPE". *Tokyo: GRIPS PhD dissertations*.
- [35] Faubert, V. (2009). "School evaluation: Current practices in OECD countries and a literature review". *OECD Education Working Papers*, no. 42.
- [36] Eyles, A.; Machin, S. and McNally, S. (2017). "Unexpected school reform: Academization of primary schools in England". *Journal of Public Economics*, no. 155, pp. 108-121.

“What works in education?” Collection

1. What works in education?: The question that has to be asked
Miquel Àngel Alegre,
Is the introduction of pay-for-performance salary incentives for teachers linked to students' academic performance advisable?
J. Oriol Escardíbul
2. Are individual tutoring programs effective in addressing diversity?
Miquel Àngel Alegre,
What grouping strategies respond to criteria of efficiency and equality?
Gerard Ferrer-Esteban
3. Are summer programs effective in improving learning and educational outcomes in students?
Miquel Àngel Alegre
4. What impact do after-school activities have on children's and youngsters' learning?
Sheila González Motos
5. Are social and emotional learning programs effective tools to improve students' skills?
Queralt Capsada,
Using self-regulation and metacognition in class: what works and under what conditions?
Gerard Ferrer-Esteban
6. Are scholarships and grants effective when it comes to the continuity and improvement of educational results at primary and secondary school level?
Mauro Mediavilla
7. School Choice and Allocation Policies: What Effects Do They Have on School Segregation?
Miquel Àngel Alegre
8. Does school leadership affect student academic achievement?
Álvaro Choi, María Gil
9. Is pupil assessment a mechanism for improving school performance?
Sheila González Motos
10. Do behavioral programs improve pupils' attitudes and outcomes?
Miquel Àngel Alegre
11. Do programs to encourage parental involvement in education improve school performance?
Jaume Blasco
12. What impact do guidance and counselling programs have on students?
Sandra Escapa, Albert Julià

First edition: January 2019
© Fundació Jaume Bofill, Ivàlua, 2019
fbofill@fbofill.cat, info@ivalua.cat
www.ivalua.cat
www.fbofill.cat

Author: Álvaro Choi
Translator: textosBCN (Dustin Langan)
Edited by: Bonal·letra Alcompas
Publishing Technical Coordinator: Anna Sadurní
International Projects Manager: Miquel Àngel Alegre
Ivàlua analyst: Núria Comas
Design and layout: Enric Jardí
ISBN: 978-84-947887-1-0

This work is subject to the Creative Commons license **Attribution-ShareAlike 4.0 International (CC BY-SA)**. You are free to copy and redistribute the material in any medium or format, and remix, transform, and build upon the material for any purpose, even commercially.

