School Inspection, Evaluation and Support Policies:

A Guide for Policy Dialogue

Agustina S. Paglayan[1]

This version: December 2010

Draft – Do not circulate without the author's permission

I.   Introduction

Recent developments in education systems around the world have included an increased interest in strengthening the evaluation of educational quality at the school level, and providing school-level support for quality improvements. Indeed, several developing countries have recently approached the World Bank for advice on how to improve their school inspection and evaluation systems, so that they are more conducive and relevant to school improvement efforts – including countries of different regions and levels of development such as Chile, Bulgaria, Egypt, Ethiopia and Poland.

As developing countries seek to create or strengthen their school inspection systems, they can take stock of the lessons learned by other –often more developed– countries that have already traveled a part of this journey. Since the 1990s, major developments in school inspection and evaluation have been observed in several high-performing education systems, including –but not limited to– the cases of the Netherlands, New Zealand and Scotland and Ireland (World Bank 2010a, 2010b, 2009). The roads they've traveled so far can provide important insight into the types of questions that need to be addressed, and the key issues and challenges that can be anticipated, as countries embark on their own journey toward a school inspection and evaluation system that is more conducive to improvements in educational quality.

In this context, this note makes four contributions. First, in Section II it presents a **conceptual framework** that identifies and discusses the key questions, issues and trade-offs that are likely to be encountered when seeking to create or strengthen a country's school inspection, evaluation and support system. This conceptual discussion can serve as a guide for policy dialogue between developing countries and World Bank staff. Second, in Section III it outlines a **diagnostic framework** that can be used to collect comparable data about the school-level inspection, evaluation and support policies that exist in different

---

[1] Agustina S. Paglayan is a doctoral student at the Stanford University. She began to write this note while working at the World Bank's Human Development Network, and finished it at Stanford. Comments are welcome and should be sent to paglayan@stanford.edu.

countries. Third, in Section IV it applies the proposed diagnostic framework to document and compare the school inspection, evaluation and support **policies of eighteen different education systems**, all of them in Europe but with different levels of performance in PISA. The main purpose of this section is to provide examples for the conceptual issues and trade-offs that are discussed in Section II and to illustrate the application of the diagnostic framework presented in Section III; a secondary purpose is to identify, where possible, any preliminary patterns about the policies that might distinguish high-performing systems from other systems. Fourth, Section V assembles a list of potentially **useful tools and materials** developed by different education systems, such as examples of school evaluation frameworks, inspection-related instructions for schools, individual school quality reports, and the Code of Ethics that applies to inspectors.

The focus of the note is on the *inspection*, *evaluation* and *support* policies that apply to *schools* –not students, teachers, school heads or school boards, but schools as entities in and of themselves. Moreover, the focus is on the policies that apply to *public* schools at the *primary and secondary education* levels, concentrating solely on *mainstream* education (as opposed to, for example, special or vocational education). That said, it is easy to expand the diagnostic framework proposed in Section III to document, at a minimum, whether there are differences in the policies that apply to public vis-à-vis private schools; primary and secondary vis-à-vis pre-primary and tertiary education institutions; and mainstream vis-à-vis vocational and special.

Before moving on to the next section, Table 1 specifies the definitions I use when referring to certain educational terms such as inspection, evaluation and support.

Table 1. Relevant Definitions

Term	Definition
Inspection	The verification of schools' compliance with educational laws and regulations.
Evaluation	The assessment of the quality of schools' educational inputs, processes and/or outcomes.
Internal evaluation	Self-evaluation or evaluation that is conducted by a school authority (e.g. the school head, teachers, members of the school community, the school board).
External evaluation	Evaluation that is conducted by an authority that works outside the school (e.g. a local, sub-national or national educational authority).
Support	The provision of resources (time, advice, expertise, financial and material resources, educational opportunities, etc.) to schools and/or their staff.
Primary education	The level of education that usually begins at ages 5-7 and lasts for four to seven years of schooling. In some countries it may be called elementary education. It typically marks the beginning of *systematic* studies on reading, writing and mathematics, although children may begin learning basic literacy and numeracy skills at the pre-primary level.
Secondary education	The level of education that usually begins at ages 11-14 and lasts for three to six years of schooling. Secondary education may include a lower secondary period followed by an upper secondary period. In some countries, these are called middle school and high school, respectively. Although it generally continues the basic programs of study of the primary education level, teaching typically becomes more subject-focused, often employing more specialized teachers who conduct classes in their field(s) of specialization.
Public school	This refers to schools that are managed by a public authority.
Private schools	This refers to schools that are managed by a non-public authority (e.g., an individual owner, a corporation, a foundation, a religious organization, etc.).

Sources: Author, World Bank (2010c), *Teacher Policies Around the World*, and Eurydice (2004).

II.   Conceptual Framework – Key Questions, Issues and Trade-Offs

The goal of this section is to provide a guide for policy dialogue between developing countries and World Bank staff, as they work together to strengthen school inspection, evaluation and support systems. When a country embarks on a journey to improve its school inspection and evaluation system so that this system is more conducive to improvements in the quality of schools, a number of questions are likely to be encountered along that journey – and decisions on key issues will need to be made. While each country will likely face questions and issues which are context-specific, it is useful to anticipate some key questions and issues which are likely to become relevant in almost all journeys. This section seeks to identify these key questions and issues. It also discusses possible answers to these questions, and the trade-offs involved in different answers.

The conceptual discussion relies primarily on (i) theoretical arguments, most of them coming from a philosophical, political or technical perspective; (ii) the reflections and thoughts –including debates and agreements– of European inspectors about the role, challenges and future of school inspection and evaluation, as imprinted in the speech transcripts and discussion papers published by SICI, the organization of European education inspectorates which has been active for the past 15 years regarding; (iii) three comparative reports about school inspection, evaluation and support in Europe (Standaert 2000; Eurydice 2004; Van Bruggen 2010); and (iv) the analytical framework of a recent OECD initiative, *Review of Evaluation and Assessment Frameworks for Improving School Outcomes*. It is worth noting that while the literature reviewed is sufficient to point out several challenges that may arise from school inspection and evaluation policies and procedures, no empirical studies were found that look at the effect of a specific policy on school quality. Similarly, the discussion about the eighteen selected systems' policies should not be understood as an analysis of what policies lead to improvements in education quality. Instead, the main purpose of looking at these systems is to illustrate the conceptual issues and trade-offs with concrete examples; and a secondary purpose is to identify hypotheses about policies that might distinguish high-performing systems from other systems.
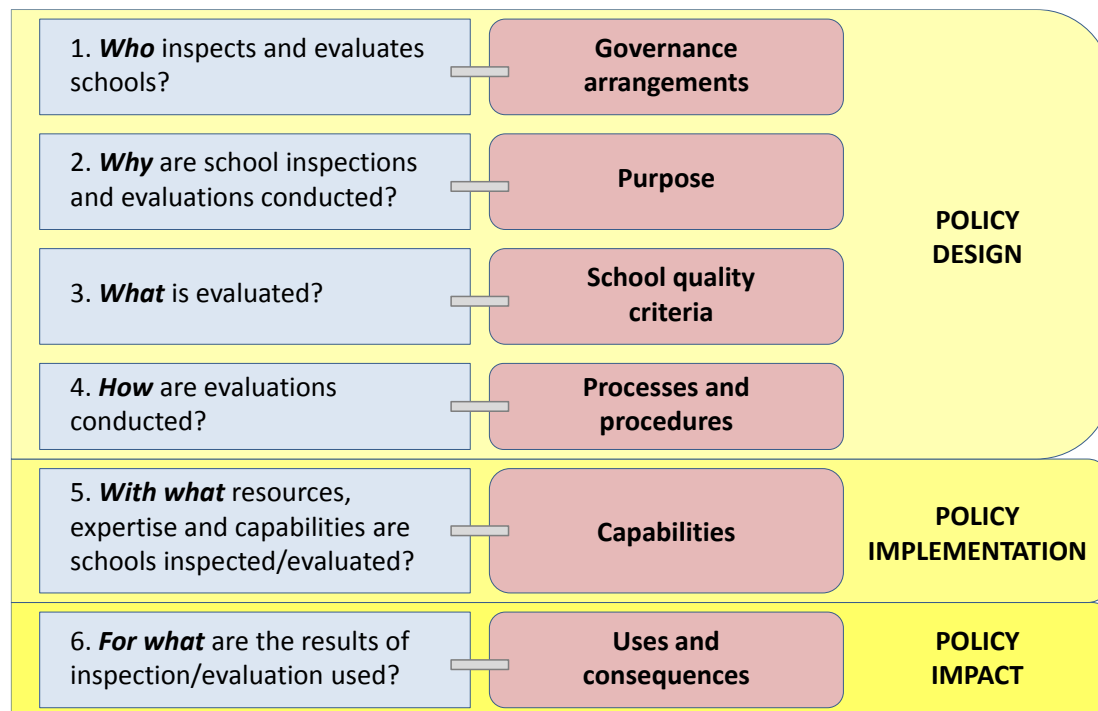
Five key questions will need to be answered in order to understand a country's current school inspection and evaluation system, detect areas for improvement, and draw a roadmap toward a stronger system:

1. *Who* inspects and evaluates schools?
2. *Why* are school inspections and evaluations conducted?
3. *What* is evaluated?
4. *How* are evaluations conducted?
5. *With what* resources, expertise and capabilities are schools inspected and evaluated?
6. *For what* are the results of inspection and evaluation used?

Each of these questions can be associated with a concept. The "who" question refers to the governance arrangements for inspection and evaluation; the "why" question, to their purpose; the "what", to the school quality aspects that are assessed; the "how", to the processes and procedures in place for evaluating schools; the "with what" question, to the capabilities to implement inspection, evaluation and support policies; and the "for what", to the use and consequences of inspection and evaluation results.

Moreover, each of these questions can be related to different stages of the policy process –policy design, implementation and impact. The "who", "why", "what" and "how" questions refer mostly to policy design issues – when seeking to create or strengthen school inspection, evaluation and support policies, it is important to assess the coherence of these four elements, and improve that coherence as needed. The "with what" question refers mostly to policy implementation – a very coherent design is necessary but not sufficient for inspection and evaluation policies to be conducive to school quality improvement; the right resources, expertise and capabilities among evaluators and those being evaluated are essential so that inspection and evaluation can add value to schools. The "for what" question refers mostly to the extent to which inspection and evaluation results are used in ways that promote or support school improvement efforts – this can be thought of as the policy impact, which is tightly linked to the evaluation culture within the education system. Figure 1 summarizes the key questions and their related concepts, as well as the stages of the school inspection, evaluation and support policy process.

Figure 1. Key Questions, Concepts and Stages of the School Inspection, Evaluation and Support Policy Process

Question	Concept	Stage
1. *Who* inspects and evaluates schools?	Governance arrangements	POLICY DESIGN
2. *Why* are school inspections and evaluations conducted?	Purpose	
3. *What* is evaluated?	School quality criteria	
4. *How* are evaluations conducted?	Processes and procedures	
5. *With what* resources, expertise and capabilities are schools inspected/evaluated?	Capabilities	POLICY IMPLEMENTATION
6. *For what* are the results of inspection/evaluation used?	Uses and consequences	POLICY IMPACT

Source: Author.

The following paragraphs discuss, for each key question, the main issues that need to be considered in order to address that question.

1.   *Who* inspects and evaluates schools? Governance arrangements.

Different governance arrangements may be in place to inspect and evaluate schools. Understanding these arrangements requires looking into several issues including who is in charge of inspection and evaluation functions; whether these functions are required or recommended; what is the relationship

and balance between the different types of inspection and evaluation that exist; to what extent evaluation and judgments about the quality of education are independent from political interference; and what mechanisms are in place to monitor the quality of inspecting and evaluating agencies and hold them accountable for their performance.

- **Who is in charge of inspection and evaluation functions?**

This question concerns the distribution of school inspection and evaluation functions across national, sub-national, local and school authorities. In this paper, inspections/evaluations carried out by a non-school authority (e.g., a national, sub-national or local government authority) are referred to as **external inspections/evaluations**. Evaluations carried out by a school authority (e.g., a school board or the school head) are referred to as **internal evaluations[2]**. Understanding who is in charge of school inspection and evaluation should be the starting point in any policy dialogue in this area. The answers to all the other questions will depend on the answer to this first one.

In most countries, an external agency conducts some form of school inspection. In addition, in some countries an external agency evaluates the quality of individual schools and the education system as a whole. This evaluating agency may or may not be the same one that inspects schools. External inspectors and evaluators may report to different levels, including the national, sub-national or local levels. In some countries, there is more than one external inspectorate and more than one evaluating agency –for example, two separate agencies for inspecting school buildings and schools' compliance with educational laws and regulations[3]; or different agencies for evaluating different aspects of educational quality; or inspecting/evaluating agencies with the same focus but reporting to different levels of government.

In addition, over the past decades several countries have introduced regulations that require or recommend schools to conduct their own internal evaluations.

In any case, policymakers face **two important trade-offs** when deciding on the distribution of school inspection and evaluation functions across national, sub-national, local and school authorities: **one between objectivity and relevance; and one between expertise and logistics** for school inspection/evaluation. External inspection and evaluation provide a potentially objective and independent mechanism to monitor compliance with laws and regulations and assess the quality of the school system. However, schools may perceive external inspection and evaluation as invasive, irrelevant, corrupt or useless, making it less likely for inspection and evaluation to make a positive contribution to school improvement. Conversely, internal evaluations may potentially provide information that is more relevant to an individual school's particular context and challenges. However, schools may not have the capacity to conduct reliable, evidence-based internal evaluations. From a resource perspective, there is also a trade-off. On one hand, it is easier to have one external inspector/evaluator assigned to several

---

[2] It would be extremely rare for schools to be in charge of inspection functions, which entails a monitoring of schools' compliance with laws and regulations.

[3] External evaluating agencies may include an educational authority, but also a non-educational one, such as a Ministry of Planning and Infrastructure, which can monitor the condition of school buildings, or a Ministry of Health, which can monitor and evaluate schools' hygiene and sanitation conditions.

schools, than to build every school's in-house expertise and capacity to conduct its own evaluations. On the other hand, schools can be reached more easily and frequently through internal rather than external inspection/evaluation.

A key challenge for external school inspection/evaluation is to be relevant to a school's particular context, and to reach schools with enough frequency so as to inform school development; while a key challenge for internal evaluation is to build the right capacities in order to be systematic and evidence-based. The existence of trade-offs implies that there is no obvious way to decide who should inspect and evaluate schools –this decision should be made after weighing the trade-offs in a particular country, and taking into account the country's overall governance arrangements for the evaluation of students, teachers, school principals and the overall quality of the education system.

- **Are inspection and evaluation required or recommended?**

There are multiple possible combinations of required or recommended inspection and evaluation. In most countries, some form of school inspection is usually required –of different aspects and with different frequencies, but some minimum form of inspection is in most cases required. In the case of school quality evaluations, when they do exist, the situations are more diverse: in some countries, both external and internal evaluations exist, but the former is required while the latter is only recommended; in other countries, both external and internal evaluations are required; in yet other countries, internal evaluation is required but external evaluation is conducted on an ad hoc basis, without any requirement for all schools to be evaluated; and some countries require participation in external evaluations but do not require or recommend internal ones. In addition, when more than one type of external evaluation exists, one form of evaluation may be required (e.g., that focusing on school inputs) while another form may be recommended (e.g., that focusing on educational processes and outcomes). Moreover, in a given country, what is required and what is recommended can change over time. It is important to understand whether inspections and evaluations are required or recommended, as this is likely to affect schools' autonomy, incentives and accountability as well as the resources and capabilities required to implement inspection and evaluation policies.

Policymakers can face several trade-offs when deciding whether to require or recommend a certain inspection or evaluation. On one hand, inspections/evaluations that are required may be more useful to keep track of individual schools' performance, hold schools accountable, provide them with incentives to continue improving the services they provide, and ensure that schools are exposed to exchanges and debates on how to bring about quality improvements. On the other hand, when inspection and evaluation are required, the sense of autonomy within schools may be somewhat diminished, schools may shift their focus so as to perform well in inspections and evaluations at the expense of other important school activities, and the need for resources within the education system is likely to be greater.

This implies that key challenges for a country that requires a certain inspection/evaluation include to minimize schools' perception that this requirement constitutes an invasion of their autonomy; reduce schools' incentives to devote excessive time to evaluation activities; curtail schools' ability to "game" the

inspection/evaluation; and ensure that adequate financial, organizational and human resources are in place to conduct the required inspection/evaluation. On the other hand, key challenges for a country that decides that a certain school inspection/evaluation will be recommended but not required include to find alternative mechanisms for monitoring the quality of the education system regularly, holding schools accountable, and ensuring access to support services that help schools identify and address their main weaknesses.

- **What is the relationship and balance between different types of inspection and evaluation?**

Different types of inspection and evaluation (e.g., external and internal evaluation) need not be substitutes of one another; they can co-exist in at least two different ways. They may be: (i) **independent**, focusing on the same aspects but not taking the results of one another into account, or focusing on different aspects altogether; or (ii) **interdependent**, focusing on overlapping aspects and where at least one form of inspection/evaluation takes into account the results of the other form (e.g., external evaluation takes into account the results of internal evaluation, and/or vice-versa).

Ideally, if both external and internal evaluations meet minimum standards of reliability, different types of evaluations should not be independent from one another, especially when their focus is on similar aspects of school quality – for external evaluators, the results of internal evaluation provide information that should not be disregarded, and vice-versa. However, in reality, different types of evaluations are not always interdependent, probably reflecting how difficult it is to make them so. A key issue in many countries, especially in lower-performing ones, is that internal evaluations are of poor quality and therefore their results are not a reliable source for informing external evaluators' judgments. Moreover, State-led evaluations may be embedded in a paternalistic culture where the inherent belief is that the State knows better than schools, leading external evaluators to disregard the information provided by schools' own evaluations. In addition, schools may not perceive the judgments made by external evaluators as legitimate, and may question the criteria and procedures used for external evaluation, or altogether the State's role in evaluating the quality of individual schools. All of these issues can be summarized by saying that, for external and internal evaluations to be interdependent, **they both need to be perceived as reliable and legitimate assessments**, as well as take place in an evaluation environment where **external evaluators and school staff show mutual respect**.

- **To what extent are external evaluations and judgments independent from political interference?**

School-level evaluations can be useful for different purposes, including to hold schools accountable, identify areas where they need more support, provide information to parents about individual schools, as well as to oversee and report on the overall quality of education provided in a country. However, in order to provide an accurate picture of the overall quality of education services that informs the general public as well as policymakers, external evaluating agencies in charge of assessing schools must have insulation form any political actors whose interests could be threatened by an honest analysis and the dissemination of the conclusions made by evaluators. Independence from political interference is a

necessary condition for external evaluation judgments to be objective and reliable, which in turn is a necessary condition for schools to trust the process and outcomes of external evaluation.

A particularly important aspect of independence from political interference is the independence of the inspecting or evaluating agency from the Ministry of Education. This independence is important because, among other reasons, the results of external evaluation can suggest that certain education policies have a weak design or implementation. Independence is needed for the inspectorate or evaluating agency to be able to disseminate this finding, as it typically implies a criticism to the work of the Ministry of Education. Where independence from the Ministry exists, an inspecting/evaluating agency will often publish Annual Reports that provide an overall picture of the main weaknesses and strengths of the education services that are available; provide solicited or unsolicited advice to Minister on certain topics and policy issues that the agency has identified as important; and report to different institutions (e.g., Parliament, the Cabinet, the Head of Government) on the state of the education system and priority areas. Such actions are indeed a way of holding the Ministry of Education accountable for its actions, decisions and performance.

Independence form the Ministry of Education can take many different forms which can be placed on a continuum. On one end, a few countries have established a national agency in charge of evaluating schools which is fully independent from the Ministry of Education. For example, in New Zealand, the Education Review Office (ERO) has its own Minister, which is distinct from and has the same hierarchy as the Minister of Education. On the other end, in many countries evaluation is carried out by a department within the Ministry of Education that has no independence whatsoever to determine the criteria and procedures for evaluating schools, the uses and consequences of school-level inspection and evaluation, or the strategies to disseminate the results of inspection and evaluation. Between these two ends, different arrangement can be observed. In some countries, schools are evaluated by an agency that has operational independence, and that reports to the Minister of Education as well as the Head of Government (e.g., the Prime Minister, President) or Parliament. This is the case of the Netherlands and Scotland, among others. In other countries, school evaluations are conducted by a department within the Ministry of Education that reports to the Minister, although it has independence for some of its operations. This is the case of, for example, Ireland.

Although from a pure design perspective there is a strong case to be made in favor of independent evaluating authorities, when implementation challenges are taken into account this advice becomes less clear-cut. Indeed, having a department within the Ministry of Education that conducts evaluations can help exploit economies of scale, ensure coordination with the Ministry, facilitate the management of educational data, and simplify (minimize) the channels used to communicate with schools. In other words, in addition to the political dynamics which affect the decision on whether to have an independent evaluating agency or not, policymakers face a **trade-off between design and implementation** considerations when making this decision. The trade-off is most likely to be resolved in favor of independence when the additional implementation (and resource) burdens are manageable.

- **Do (what) mechanisms exist to assure the quality of inspection and evaluation, and hold inspecting and evaluating authorities accountable?**

For inspection and evaluation to be relevant and make a positive contribution to school quality improvement, the quality of inspection/evaluation must be monitored and inspecting/evaluating authorities held accountable for their performance. This is particularly true for external evaluation – at the same time that the external evaluators will "demand" schools to be up to standard, so will schools demand these evaluators' work to be of high quality and relevant to improve their daily functioning.

Many mechanisms can be put in place to provide incentives for external authorities to carry out quality inspections and evaluations. Examples of quality assurance and accountability mechanisms include: (i) **disseminating and regularly revising the methodological frameworks** used to carry out an inspection/evaluation; (ii) developing and enforcing a **Code of Ethics** that outlines what schools should expect from inspectors/evaluators and the values that should guide inspection/evaluation work (e.g., integrity, impartiality, transparency, openness, respect); (iii) conducting **ad-hoc parallel inspections** to assess the consistency of inspectors/evaluators' judgment and their use of inspection frameworks; (iv) the administration of **post-inspection questionnaires** among school heads, teachers, parents and/or students, inquiring about their satisfaction with the inspection/evaluation process; (v) **complaints mechanisms** available to schools; (vi) well-established **opportunities for schools to contest** the results of inspection/evaluation; (vii) **regular evaluation of the performance of individual units and individual employees** within the inspecting/evaluating agency; (viii) requiring inspecting/evaluating agencies to publish an **Annual Report** that includes data about the activities carried out, the allocation of the budget, the consequences of inspection/evaluation, and any relevant changes within the agency (e.g., number of schools inspected, reports published, schools judged to be non-compliant with laws or under-performing in terms of educational quality, number of schools intervened, changes in the criteria, procedures or uses of inspection/evaluation, changes in leadership or organizational structure); and (ix) **impact evaluations** that assess the extent to which external evaluation leads to the desired outcomes, including the development and implementation of school improvement plans, improvements in teaching, and better educational outcomes (student achievement, progression in school, tertiary education and labor market opportunities, etc.). Measuring the extent to which external evaluation contributes to improve the quality of education will be crucial for the evaluating entity's legitimacy to evaluate schools' quality.

Quality assurance mechanisms can serve **an additional purpose: to identify unintended consequences of inspection and evaluation**. Negative side-effects typically include "putting on an act" on the day of the inspection, excluding low-performing students on the day of the assessment, increased stress among teachers, too much time dedicated to administrative tasks, and fear of introducing educational innovations. The feedback provided by schools, teachers, parents and students through questionnaires and complaints, as well we the evidence gathered through random, unannounced inspections/evaluations conducted to double-check the objectivity and reliability of previous inspection/evaluation results, can help to shed light on the extent to which inspection or evaluation leads to unintended consequences.

Mechanisms to ensure that internal evaluations are reliable and up to standard are also necessary. A natural way of doing this is through external evaluation (i.e., looking at the quality of internal evaluation processes and results as one of the elements of the evaluation of a school's overall quality). Other ways can include developing frameworks/guidelines at the national level that schools are required or

recommended to apply when conducting their own evaluations; or having complaints procedures available for teachers, parents and students to report cases of inadequate internal evaluation.
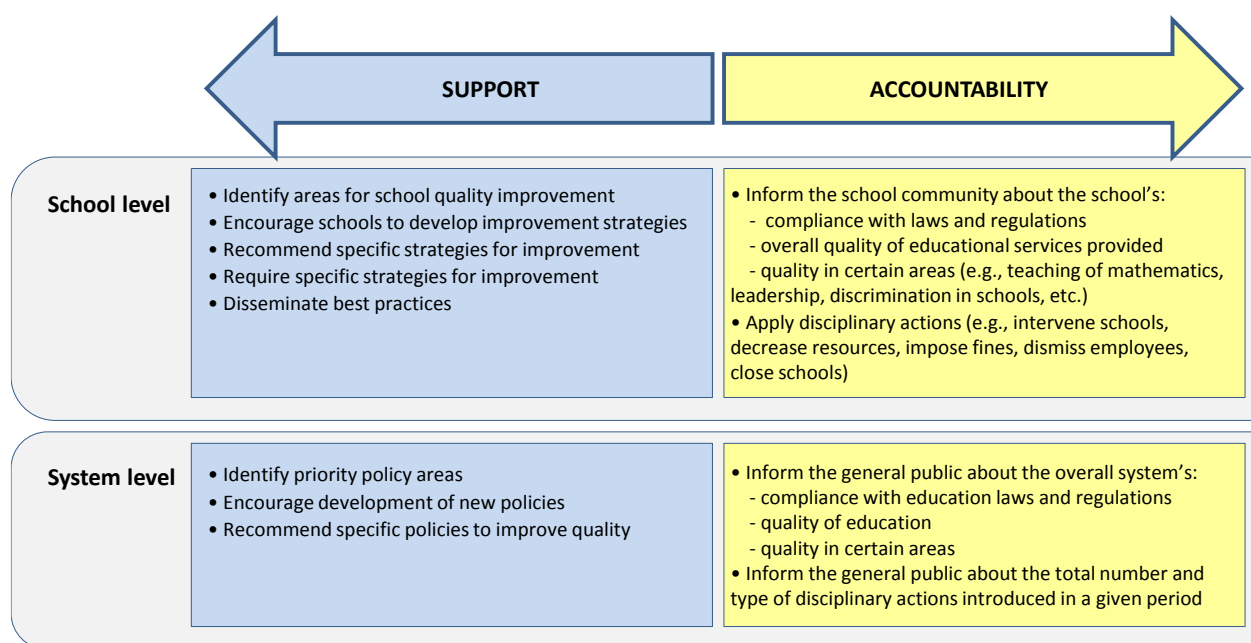
2. *Why* are school inspections and evaluations conducted? **Purpose.**

The purpose of school inspections and evaluations can vary considerably across countries. In some countries, the purpose of inspection/evaluation may not be well-defined in laws or regulations. In addition, even if it is well-defined on paper, in practice it may differ from how inspection/evaluation is actually used. Understanding the purpose of inspection/evaluation policies entails understanding what these policies were designed for; why they were introduced. Later on, under question 6, I will discuss the importance of also documenting what inspection and evaluation are actually used for. Documenting both the written purpose and actual use given to school inspection, evaluation and support policies are critical to identify any gaps between the design and implementation of these policies. A look at the written purpose alone is also useful as it can contribute to identify some weaknesses in the way these policies were designed.

By definition, inspection entails a purpose to monitor schools' compliance with educational laws and regulations, and evaluation (external or internal) entails a purpose to assess the quality of education provided by schools. But inspection and evaluation can have additional goals: to encourage schools to embark on school improvement efforts based on the results of inspection/evaluation, or furthermore, to support schools as they think about what changes to introduce; to hold schools accountable; to assess the overall quality of the education system and identify key areas where improvement is needed; and to hold education policymakers accountable for the quality of the education system as a whole. The different goals that inspection and evaluation can have fall into two main categories, support and accountability, which can be observed at two levels, individual schools and the education system as a whole. Figure 2 summarizes the possible goals or purposes of inspection and evaluation.

There is an **inherent tension between the support and accountability purposes of inspection and evaluation**. On one hand, in evaluation systems that focus on supporting schools to improve the quality of education there is an implicit presumption that schools need help in order to make improvements; that information and data are not enough for schools to address their weaknesses, and that a more active advisory role is necessary. On the other hand, in evaluation systems that focus on holding schools accountable, there is an implicit presumption that reporting to relevant stakeholders about the quality of schools will per se lead schools to introduce changes that are conducive to quality improvements. This ultimately means that in systems where the main purpose of evaluation is on support, there is relatively little confidence in the capacity of schools to be autonomous; while in systems where the main focus in on accountability, there is a strong confidence in that capacity to be autonomous. A corollary of this is that in evaluation systems that mainly focus on supporting schools, the level of State intromission in the day-to-day operation, management and decisions of schools is greater than in systems whose main focus is on accountability. There is growing consensus among educational experts that the right incentives are necessary but cannot by themselves lead schools to make improvements in quality; while supporting schools without giving them the right incentives to improve is also likely to be a futile investment.

Figure 2. Purposes of inspection and evaluation: Support and Accountability

	SUPPORT	ACCOUNTABILITY

School level	• Identify areas for school quality improvement
	• Encourage schools to develop improvement strategies
	• Recommend specific strategies for improvement
	• Require specific strategies for improvement
	• Disseminate best practices
		• Inform the school community about the school's:
		  - compliance with laws and regulations
		  - overall quality of educational services provided
		  - quality in certain areas (e.g., teaching of mathematics, leadership, discrimination in schools, etc.)
		• Apply disciplinary actions (e.g., intervene schools, decrease resources, impose fines, dismiss employees, close schools)

System level	• Identify priority policy areas
	• Encourage development of new policies
	• Recommend specific policies to improve quality
		• Inform the general public about the overall system's:
		  - compliance with education laws and regulations
		  - quality of education
		  - quality in certain areas
		• Inform the general public about the total number and type of disciplinary actions introduced in a given period

Source: Author.

Documenting the purpose of internal evaluation is also important. In some countries, internal evaluation may be conducted without a clearly-defined purpose. When the purpose is defined, it may vary considerably from serving as a tool for schools themselves to identify their weaknesses and embark on school improvement strategies, to informing local- or national-level decisions (e.g., resource allocation, inspection and external evaluation plans and procedures, investment in infrastructure) or providing an overall picture of the state of education in the country.

The **balance between support and accountability purposes is likely a dynamic one**; it can be expected to change over time. Factors that usually affect a country's decision on whether to use inspection/evaluation mostly for support, or mostly for accountability, include: (i) the overall level of regulation in the education system; (ii) the degree of school autonomy; (iii) the historical role of the State in society in general and in the education system in particular; and (iv) prevailing ideologies about the role that the State ought to have in education, including national ideologies but also international trends. In general, systems are more likely to prioritize the role of inspection/evaluation as a means of holding schools accountable in systems that are fairly deregulated, where schools have a lot of autonomy, and where the State has historically delegated many aspects of the day-to-day management of schools. In contrast, systems are more likely to prioritize inspection/evaluation as a means to identify and provide support to schools in systems where there is a heavy amount of educational regulation, where schools have little autonomy, and where historically the State has had a role in determining details of the day-to-day management of schools. International trends, such as the current move toward systems of inspection/evaluation that give similar weight to the purposes of support and accountability, can contribute to change the purpose assigned to inspection/evaluation in a country, although the influence of international trends will vary across countries. Regardless of the influence of international trends, the

purpose of inspection/evaluation policies should be consistent with the country-specific context (e.g., level of regulation, degree of school autonomy).

3. *What* **is evaluated? School quality criteria.**

A key aspect of school-level evaluations is the definition of what is evaluated. To understand what is evaluated in a country, three main issues need to be documented: (i) whether common criteria are used to evaluate schools, and when they are, (ii) what those school quality criteria include and (iii) how they were determined.

- **Do common criteria for evaluating the quality of schools exist?**

In many countries, although external and/or internal evaluations are conducted at the school level, the criteria for such evaluations are not clearly defined. For example, external evaluators may assess schools in a very subjective manner, following no common framework or methodology, or applying very broad guidelines of what things should be taken into account. Similarly, there may be no required or recommended frameworks for schools to conduct internal evaluation, with the definition of what elements to look at and how to assess each particular element left to be decided by each school. The absence of a common framework that defines what constitutes quality education and describes what a school should look like to be able to conclude that it provides up-to-standard services, has several limitations: the results of evaluation will be very subjective, dependent on the evaluator's analytical and technical capacity, difficult to compare across schools, and will reflect a narrow view (that of the evaluator) of what constitutes quality education, which may not reflect how parents, civil society, teachers, school heads or educational experts understand "quality".

When **common frameworks developed at the national level** do exist, it is important to document whether internal/external evaluators are **required or recommended** to apply them when assessing the quality of a school. It is important to note that common frameworks may also be developed at lower levels of government (sub-national or local). While it is easy to document the existence of those frameworks, documenting their contents and how they were developed is time-consuming.

From a theoretical and purely technical perspective, without taking into account any of the implementation challenges, it could be argued that it is desirable for countries to have a common framework for evaluating schools, developed at national or sub-national levels, and for this framework to include indicators of outcomes, processes and inputs/context, and be as comprehensive (include as many relevant indicators) as possible. As we've already noted, in the absence of a common framework that defines what constitutes quality education and describes what a school that provides up-to-standard services should look like, the results of evaluation will be very subjective, dependent on the evaluator's analytical and technical capacity, difficult to compare across schools, and will reflect a narrow view (that of the evaluator) of what constitutes quality education, which may not reflect how parents, civil society, teachers, school heads or educational experts understand "quality". Common frameworks may be defined at national, sub-national and local levels; national and sub-national ones facilitate the translation of school-level evaluation results into analyses of the system-level quality of education.
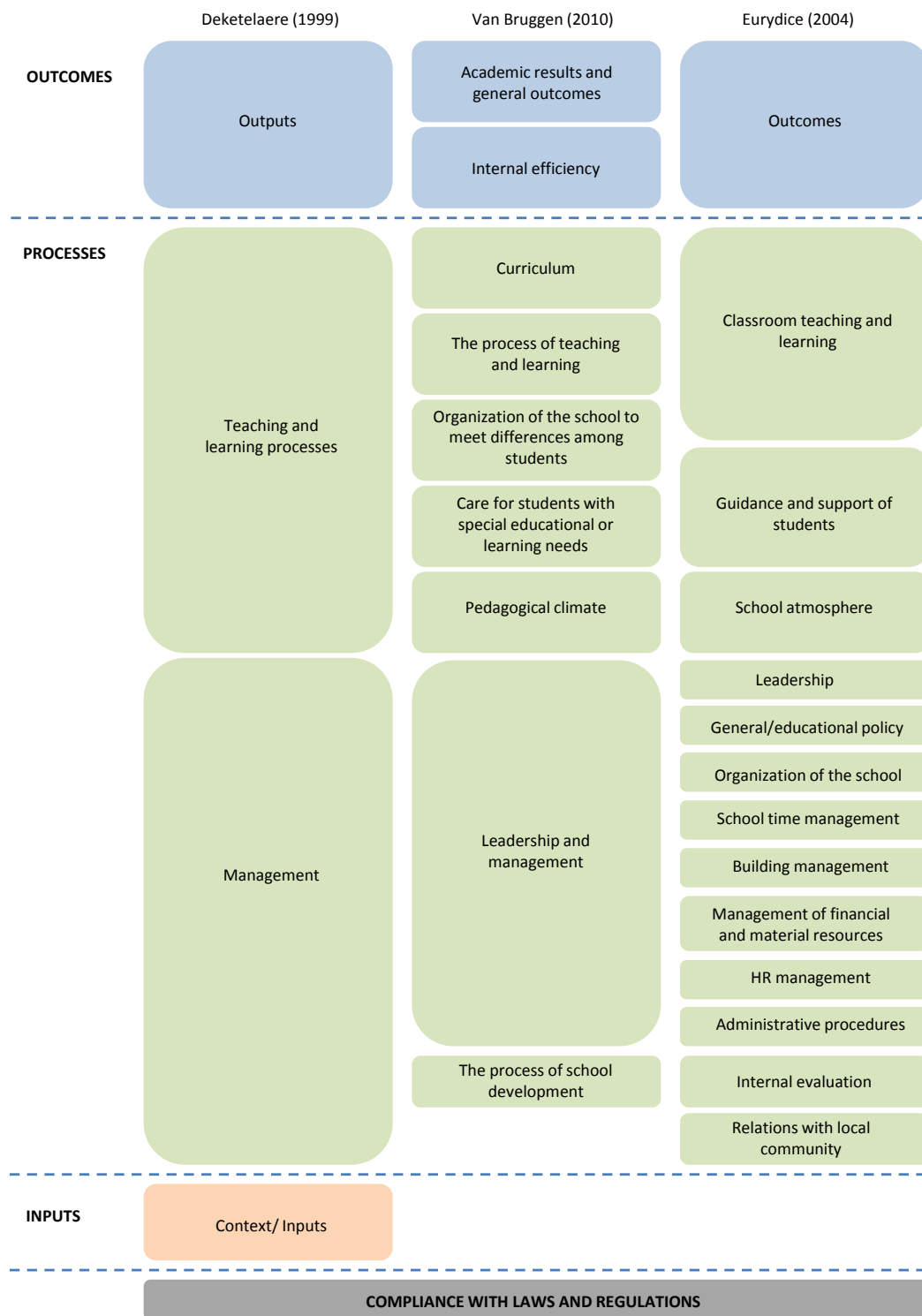
A common framework for evaluating schools can include one or more of the following items: (i) **quality areas or dimensions** of school quality, which are the broad aspects or elements of the quality of a school; (ii) **quality indicators**, which are the specific issues included within each of these dimensions; (iii) **quality variables**, which are the specific ways in which each indicator can be observed or measured; (iv) **possible values** that each variable can take; (v) **standards, norms or descriptors** of the circumstances under which a variable will take each of the different possible values; and (vi) **methods for aggregating the values** of all the different variables, so as to arrive at an overall value for each indicator and dimension. The criteria may vary for primary and secondary education, or for different types of schools (e.g., multi-grade schools, rural schools, religious schools, etc.).

Quality areas/dimension and indicators

According to Van Bruggen (2000), a good set of indicators is one that (i) recognizes the complex relationships between inputs, outputs and processes that affect the quality of a school; (ii) enables inspectorates to make a judgment not only about the quality of an individual school but also about the quality of the education system as a whole (i.e., most indicators fit both levels); (iii) is fairly consistent over time, so as to be able to analyze if there has been improvement or not, for both schools and the whole system; and (iv) resonates with the school's own thought process toward development and improvement (common language, aspects, etc.). The latter is critical for schools to *identify strengths and weaknesses over which they believe they have direct control* (e.g., clarity in teachers' explanations, or engagement of students in the classroom, as opposed to student achievement).

About a decade ago, Anne Deketelaere conducted a comparative study of school quality criteria used in external evaluations throughout Europe. The study, *Indicators for Good Schools: Analysis and a Proposal*, had been commissioned by SICI (the organization that brings together education inspectorates from European countries) to identify common criteria for school quality evaluations. Based on the review of existing frameworks, Deketelaere (1999) proposed four quality areas/dimensions that "any school quality evaluation should take into account": outputs, teaching and learning processes, management processes, and context/inputs (Van Bruggen 2000). More recently, Johann Van Bruggen (2010), in a study also conducted for SICI in which he reviews the criteria used by educational inspectorate in Europe to evaluate schools, has identified a different set of quality areas/dimensions: academic results and general outcomes; internal efficiency; curriculum; the process of teaching and learning; organization of the school to meet differences among students; care for students with special educational or learning needs; pedagogical climate; leadership and management; and the process of school development. Van Bruggen's areas/dimensions overlap with Deketelaere's outputs, teaching and learning processes, and management processes, but do not include any indicators for school context and inputs. Finally, a comparative study conducted by Eurydice (2004) also identifies an exhaustive set of areas/dimensions which are used across European countries for evaluating schools. Some of these criteria are common to all countries, while others are used in just a few of them.

Figure 3. Quality areas/dimensions by Deketelaere (1999), Van Bruggen (2010) and Eurydice (2004)

	Deketelaere (1999)	Van Bruggen (2010)	Eurydice (2004)
OUTCOMES	Outputs	Academic results and general outcomes	Outcomes
		Internal efficiency	
PROCESSES	Teaching and learning processes	Curriculum	Classroom teaching and learning
		The process of teaching and learning	
		Organization of the school to meet differences among students	
		Care for students with special educational or learning needs	Guidance and support of students
		Pedagogical climate	School atmosphere
	Management	Leadership and management	Leadership
			General/educational policy
			Organization of the school
			School time management
			Building management
			Management of financial and material resources
			HR management
			Administrative procedures
		The process of school development	Internal evaluation
			Relations with local community
INPUTS	Context/ Inputs		
	COMPLIANCE WITH LAWS AND REGULATIONS		

Sources: Author based on Van Bruggen (2000 and 2010) and Eurydice (2004).

Figure 3 summarizes the school quality areas/dimensions proposed by Deketelaere (1999), Van Bruggen (2010) and Eurydice (2004). A more detailed comparison, which includes not only the quality

areas/dimensions but also the indicators that fall under each, is presented in Appendix 1 at the end of the document. Several observations emerge from comparing the areas/dimensions and indicators they propose. First, all definitions of school quality include aspects related to educational outcomes as well as aspects related to educational processes. Second, included among the outcomes are not only academic attainment and cognitive skills but also socio-emotional skills; the destination of students after school (e.g., further education, entry to the labor market); and measures of internal efficiency such as repetition rates, drop-out rates and student attendance. Third, among the processes, these can be divided into two main categories: teaching and learning processes, and management processes. The former include the curriculum offered by the school; the processes of teaching and learning in the classroom; the school's organization and ability to address students' individual and differential needs; and the overall school atmosphere and relations between teachers and students. The latter include a consideration of the school leadership, vision and educational policies, and internal communication; the school's organization and administrative procedures; the management of time, human, building, financial and material resources; communication with the local community; and the processes for internal evaluation and quality assurance. Fourth, the school quality criteria proposed by Anne Deketelaere include, in addition to educational outcomes and processes, the educational inputs and the context in which the school operates – for example, the socio-economic background of students, the socio-economic characteristics and qualifications of teachers, the total funding of the school, and any aspects that affect the educational processes and outcomes but over which the school has relatively little or no control. Finally, all frameworks implicitly or explicitly include the monitoring of schools' compliance with the relevant laws and regulations.

The rest of the note refers to the school quality areas/dimensions and indicators listed in Table 2.

Table 2. Proposed school quality areas/dimensions and indicators

Areas/Dimensions	Indicators
STUDENT OUTCOMES	Academic attainment
	Cognitive learning skills
	Social and emotional skills
	Repetition and dropout rates
	Destination of school leavers
	Attendance
TEACHING AND LEARNING PROCESSES	Curriculum/ Subjects offered by the school
	Teaching processes
	Learning processes
	Addressing of students' individual needs
MANAGEMENT PROCESSES	School climate
	Leadership and vision
	Organization
	Administrative procedures
	Communication in the school
	Time management
	Financial management
	Building management
	HR management
	Internal evaluation and school development
	Relationship with wider community
CONTEXT/INPUTS	Local socio-economic, educational and geographic characteristics
	Students' socioeconomic, health, linguistic/cultural background
	Teachers' socioeconomic, educational, linguistic/cultural background
	Infrastructure (buildings, funding, science labs, ICT, etc.)
	Balance of intake of students (by ethnicity, gender, language skills, special needs)

Evaluations that look into educational outcomes, processes and inputs/context can provide a more complete picture of what is going on at the school than evaluations that look at only some of these elements. If outcomes are not taken into account, it is difficult for evaluators to judge objectively the merit of the teaching and management processes observed in the school. If processes are not taken into account, it is difficult for evaluators to identify what are best practices that lead to desirable outcomes, and what processes should be improved for a school to achieve the goals of education. If inputs/context are not taken into account, a school's own merits may be under or overestimated, by confounding what the school does with what it "inherits" or "receives".

A comprehensive set of indicators is desirable to minimize schools' ability and incentives to "game the inspection". When fewer indicators are considered, schools can focus on them just to ace the evaluation, neglecting other important factors that form part of a school's quality. For example, if students' achievement in standardized tests is the only indicator of educational outcomes that is considered, schools may have an incentive to teach to the test, and may disregard students' overall cognitive skills (e.g., the capacity to make connections across disciplines, apply subject knowledge to real life situations or become independent learners) as well as their social and emotional development and well-being.

It is worth noting that the discussion so far has referred to the frameworks that are in place for **whole school evaluations**, that is, evaluations that seek to assess the *overall quality* of an *individual school*. However, school-level evaluations may be conducted with the aim of assessing the quality of: (i) a particular issue (e.g., how schools deal with discrimination or bullying against students, how they teach foreign languages); and/or (ii) the overall school system.

Evaluations conducted to assess the quality of a particular issue, also known as **topic evaluations**, may co-exist with those that seek to assess the overall quality of a school or system. When they do, a topic module can be included as an addendum to the general framework used for whole school evaluations, and applied at each school together with it overall evaluation. The same principles that apply to whole school evaluation frameworks also apply to topic evaluation frameworks. A common framework or set of criteria is desirable so as to increase the objectivity, comparability and transparency of the evaluation. The criteria used to conduct topic evaluations may include indicators, variables, possible values for each variable, standards/norms/descriptors for each possible value, and methods for aggregating the values of all the different variables. A good topic evaluation framework is one that (i) recognizes the complex relationships between inputs, outputs and processes; (ii) enables evaluators to make a judgment not only about the quality of an individual school but also about the quality of the education system as a whole; (iii) is fairly consistent over time; and (iv) resonates with the school's own thought process toward development and improvement (common language, etc.). The

Assessments of the quality of the school system as a whole may be based on the aggregation and analysis of the whole school evaluations conducted in individual schools. Some countries may not disclose individual school results at all, not even to the school staff or community, using these results only to inform policymakers and/or the general public about the overall quality of education provided in a country. System-level assessments can be based on a representative sample of schools; they do not require an assessment of all schools in the country to provide an accurate picture.

Quality variables

Two systems can have national frameworks for evaluating schools that include the same areas/dimensions and indicators of school quality, and yet are significantly different from one another. While the areas/dimensions and indicators considered provide an overview of what is expected from schools, it is the *variables* used in the assessment that provide a more accurate flavor of what actions are valued and promoted within a particular education system, and how "quality" is defined.

The definition of what variables are used to evaluate schools is important because it gives a flavor of what is considered as "quality" education, and contributes to provide a clear definition of what is expected of schools. But defining the variables is not an easy task. By definition, each variable needs to be measureable, and so the variables that can be included will depend not only on what is considered to be indicative of quality but also on what sources of evidence are available and can be used to measure a particular variable in a particular system. In this regard, less-developed countries may need to invest in strengthening their data systems and technical capacity to measure different aspects related to the quality of education –a precondition for designing and implementing a school evaluation framework that outlines not only dimensions and indicators, but also specific variables.

Frameworks cannot be easily transported from one country to another. The specific variables in a framework reflect a system's particular view of what education is all about, and what schools' contribution to children's lives should be. In some countries, there may be a general agreement that schools are there to impart basic literacy and numeracy skills. In other countries, there may be a tacit understanding that the purpose of schools is to make sure that individuals are properly prepared to face the challenges of higher education and the labor market. In yet other countries, the development of cognitive skills may have just as much importance as the development of democratic values, or religious values, or other values that are central to a specific society. The starting point for any common framework should be the definition of the purposes of schooling in a particular society.

Possible values the variables can take

In addition to defining the dimensions, indicators and variables to be used to evaluate schools, national frameworks (or sub-national or local ones) can also specify the values that each variable can take. These values can then be used to identify the main areas in which the school exhibits weaknesses, as well as its main strengths, and to classify a school depending on how many weaknesses and strengths it has.

There is agreement among some school evaluation practitioners that a scale with an even number of possible values is desirable, as it forces evaluators to make a judgment about the direction in which the balance between strengths and weaknesses is inclined (Van Bruggen 2010). In most countries, when a scale exists, the value assigned to each variable is not reported, and neither are the very detailed rules that evaluators use to assign those values. There is a trade-off in the decision on whether or not to report the possible values of each variable and what they signify: transparency may contribute to increase the legitimacy of the evaluation process and results, but it can also lead to detailed discussions about the specific values assigned to each variable, diverting evaluators and school staff from the more important discussion about the overall quality of the different dimensions and indicators of quality. Indeed, what *is* reported more frequently in countries that have a national evaluation framework is the overall assessment made of the quality dimensions and the indicators within them; a judgment which takes several variables into consideration.
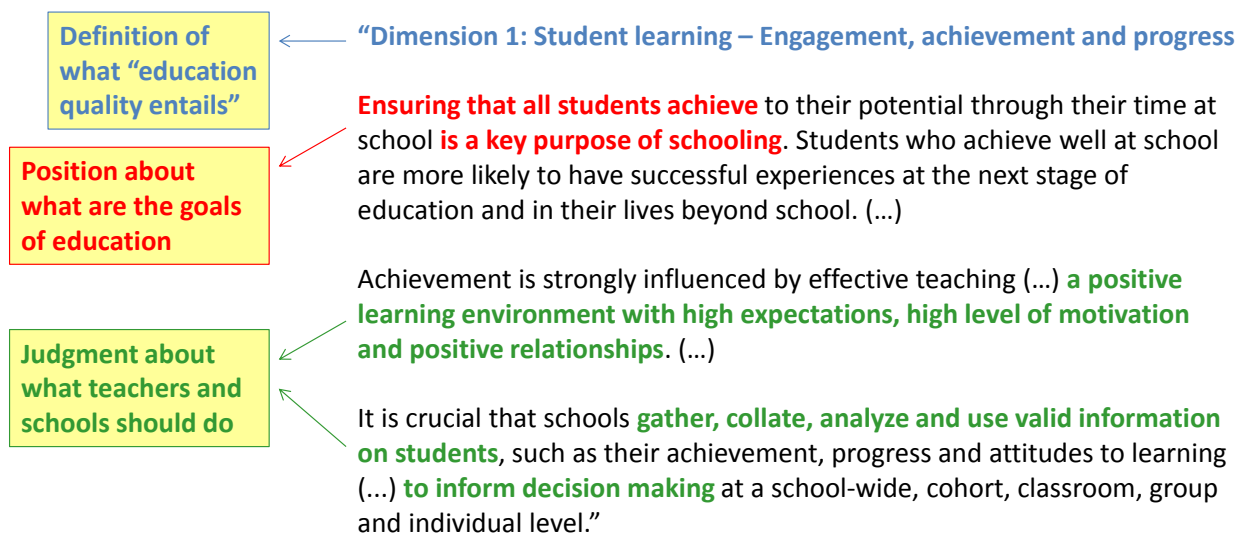
Dimensions and indicators are also often classified using an even number of categories, such as the following four categories: "mostly weaknesses", "more weaknesses than strengths", "more strengths than weaknesses" and "mostly strengths". Regardless of whether or how a school's quality is classified,

evaluations can also help to point out what are a school's main strengths and what are the areas in which improvements are needed.

Standards, norms or descriptors

The quality dimensions and indicators, as well as the variables, entail a description of the aspects of quality that can be included in the evaluation of a school. The determination of what dimensions and indicators to consider in a school evaluation implies a certain definition of what "quality education" is, a position about what the goals of education are, and judgment of what aspects a school should concentrate on to guarantee a minimum level of quality for all its students.[4] An excerpt from the New Zealand Education Review Office's school evaluation framework illustrates this:

Figure 4. Definitions, positions and judgments about education quality made in New Zealand's national framework for school evaluations

> **"Dimension 1: Student learning – Engagement, achievement and progress**
>
> **Ensuring that all students achieve** to their potential through their time at school **is a key purpose of schooling**. Students who achieve well at school are more likely to have successful experiences at the next stage of education and in their lives beyond school. (…)
>
> Achievement is strongly influenced by effective teaching (…) **a positive learning environment with high expectations, high level of motivation and positive relationships**. (…)
>
> It is crucial that schools **gather, collate, analyze and use valid information on students**, such as their achievement, progress and attitudes to learning (...) **to inform decision making** at a school-wide, cohort, classroom, group and individual level."

Source: Author based on Education Review Office (2010), *Evaluation Indicators for School Quality Reviews.*

However, the definition of **standards** involves a further judgment of the form "X (a certain dimension or indicator) has to have a specific shape and/or certain features to be considered 'good enough'". The standard is what provides the norm or rule for how to grade an indicator (positively or negatively). In the example above, the standards would include descriptions of what is considered *good enough* when it comes to setting expectations for students; promoting positive relationships between teachers and students; and collecting, analyzing and using student achievement data. National frameworks used to evaluate schools may or may not include descriptions of what a dimension or indicator should look like to be considered up to standard.

Descriptions of what constitutes the standard for a certain dimension/indicator are tightly linked to the existence of categories for classifying schools on that dimension/indicator. While this may seem obvious,

---

[4] As discussed previously, the definition of the variables is also affected by the existing data and sources of evidence in a country, or the possibility of developing new data and information sources.

it is important to ensure that these descriptions be provided to schools whenever evaluation leads to some sort of classification. The descriptions are not only important for the transparency of the evaluation process and the reliability of its results, but also because they provide a guide to schools on what aspects they should address whenever they are not performing up to standard on a particular dimension or indicator. In that sense, standards are an important mechanism for school-level evaluation to contribute to educational improvement.

Over the past decade, several efforts have been made –and continue to be made- by European inspectorates of education to establish a common set of standards that can be applied to schools in all member countries of SICI. This would enable countries to benchmark their schools against others in the continent, have a better sense of their relative position in terms of the overall quality of education services available at the system-level, and identify best practices from which all schools and systems can learn. However, while agreeing on common school standards may be attractive and theoretically possible, in practice reaching this agreement has proved to be a very difficult task which requires extensive debate, money, time, expertise and testing –all of which are difficult to gather.

- **How were common school quality criteria determined?**

While having a common framework has clear advantages, the application of such a framework may be unwelcomed by some schools, especially those that value their autonomy and feel that the framework is not aligned with the school's own goals and pedagogical orientation. To minimize confrontation and increase the perception that the evaluation of schools using common criteria is legitimate, it is crucial for such criteria to reflect a **consensus of what is generally understood as "quality education"** in that particular society.

Reaching consensus is not an easy task; it requires intensive consultation with an extensive group of education-sector stakeholders. Among the stakeholders to be consulted, examples could include: educational experts and scholars; representatives of schools or school associations; representatives of teachers, non-teaching staff and school principals; school inspectors; parent associations; student associations; local, sub-national and national educational authorities; NGOs that represent the interests of children, and business associations.

In some countries, such as the Netherlands, the national evaluating authority (i.e., the Dutch Inspectorate of Education) is *required by law* to develop the school quality criteria in consultation with teachers, schools and other relevant stakeholders. While consultations are likely to involve a tense negotiation, they can also be constructive. In the Netherlands, the inclusion of context variables in the evaluation of schools was mostly the result of teachers' concern about being judged unfairly if students' background and the conditions of their work were not taken into account. Similarly, the content of school quality reports has been shaped by awareness that what parents look at when selecting a school for their children is less related to academic achievement and more related to the values, ethos and general support and caring of students provided by the school.

A well-respected framework for evaluating schools is one that "resonates with the school's own thought process toward development and improvement" (Van Bruggen 2000); has a track record of being useful

for schools to identify their strengths and weaknesses; can be implemented smoothly, and has minimal undesired side-effects. To achieve this, consultation is necessary but not enough. Another important aspect to build a framework's legitimacy and robustness is the process of **testing and revision of the framework and criteria**.

In many of the high-performing countries that have introduced whole school evaluations in the past fifteen years (e.g., the Netherlands, Ireland, Scotland), this introduction has been very gradual. Typically, whole school evaluations begin to be implemented in a representative sample of schools. The pilot experience provides valuable information on whether the desired outcomes are observed, and whether any undesired side-effects also arise from the implementation of the common frameworks. These lessons are used to refine the frameworks. The development, piloting and refining of the frameworks before they are applied at a massive scale can take several years. For instance, in Ireland it took two years (1998-1999) to develop the Whole School Evaluation framework and apply it in a small sample of 35 pilot schools; and an additional four years before the pilot experience became a general program applied at the national level. Similarly, in the Netherlands, it took four years (1991-1994) for a sample of 180 primary schools to be evaluated individually for the first time; and an additional five years for the evaluation framework to be refined and applied to 13 percent of all primary schools (World Bank 2010a). Usually there are two separate testing and revision processes (and frameworks) for primary and secondary schools.

Once a framework begins to be implemented at a massive scale, there is still a need to revise it periodically, taking into account the positive and negative effects that arise during its application. This issue was already discussed in the section dedicated to the importance of having mechanisms to assure the quality of school-level evaluations.

4. *How* are evaluations conducted? Processes and procedures.

The processes and procedures used to evaluate schools are important to document. The frequency and duration of each evaluation; what sources of information are used to assess how schools are doing; how much administrative work is required from schools and the extent to which evaluation interrupts schools' normal operation; the extent to which teachers, school principals, parents and student actors are involved in the evaluation process; and the structuring of the evaluation process (e.g., pre-evaluation gathering of data, evaluation and follow-up) may all contribute to explain the extent to which evaluation (internal or external) is conducive to improvements in the quality of education. Very little is known about how each of these issues can contribute to make school-level evaluations more effective. Documenting these issues across as many countries as possible is the first step to understand the extent to which each of these issues matters.

- **How often is each school evaluated? How long does each evaluation last?**

The frequency of evaluation is likely to affect the extent to which evaluation is conducive to school quality improvements. Evaluation can provide schools with an overall picture of their main strengths and weaknesses, which can be used to inform a school's improvement strategies. But evaluation provides a snapshot, that is, an overall picture about the quality of a school *at a given point in time*. A school's

challenges, opportunities, strengths and weaknesses are likely to evolve over time. Thus, for evaluation to be relevant to schools, it needs to be up to date and in line with the school's reality. It is unlikely that evaluation will contribute to school quality improvements if it is done too infrequently. On the other hand, if schools design and implement strategies to address the weaknesses identified through evaluation, it might take considerable time before improvements materialize and become observable. Thus, if evaluation occurs too frequently, improvements may not be observed –and schools may become demoralized and unmotivated, and question the legitimacy of evaluations. Finding the right balance is therefore an important task and challenge, and may take several years of trial and error.

As for the duration of evaluation, this may vary from a few hours to a few days. Again in this regard, there is a trade-off in the decision of how long the evaluation process should last. On one hand, evaluations that are relatively short in duration are desirable in that they minimize the extent to which the normal operation of schools is disrupted. On the other hand, a longer process may be needed for evaluation to provide useful information about a school's strengths and weaknesses, for schools to build their own capacity to conduct evaluations, and for external evaluators and school staff to strengthen their relationship.

Finally, it is worth noting that the frequency and duration of evaluation need not be the same for all schools. In some countries, mostly because of limited resources, there has been a recent shift toward **risk-based inspection/evaluation**. What this means is that a preliminary assessment which takes into account a few key indicators of school quality is made of all schools, and then, if a school is suspected to provide below-standard educational services, it is evaluated more comprehensively (i.e., the duration of evaluation is longer). If the results of the comprehensive evaluation show that in fact the school has serious deficiencies, then it is also evaluated more frequently, to monitor its progress more closely.

This principle of evaluating the weakest schools more frequently may translate to the system-level too, so that, on average, schools in high-performing systems are evaluated less frequently than those in lower-performing systems.

- **What sources of information are used to inform judgments about a school's quality?**

As with the frequency of evaluations, when it comes to choosing which sources of information to use to assess a school's overall quality, evaluators face a trade-off. On one hand, the more varied the sources of information, the better the evaluator's ability to get a clear picture of the school's quality, and the lower the incentives and ability of schools to "game the inspection/evaluation". On the other hand, the greater the number of sources of information, the greater the resources and capacity needed to conduct the evaluation. As in other cases already discussed, the trade-off here is between design and implementation concerns. Because a varied set of sources of information is "technically" desirable form a design perspective, the extent to which this is possible will depend on the resources available for conducting evaluations.

Among the sources of information that can be used, the most usual types of sources include: **classroom observations**; **standardized test scores**; **samples of students' work**; **student records** (e.g., academic transcripts, attendance records, disciplinary records); **administrative records** (e.g., payroll records,

balance sheet; teacher attendance records); **school documents** (e.g., Annual Report, School Plan, Newsletters); and **feedback from teachers, school management, parents and/or students** (provided through questionnaires, interviews or consultation meetings). In external evaluation, an important source of information may be the school's **internal evaluation report**. In some countries, the review of a school's internal evaluation report is the starting point for conducting an external evaluation, and external evaluation includes an assessment of the internal evaluation processes in place and the reliability of the report produced by the school.

Recently, a few high-performing systems have invested in creating data systems that integrate school-level information on various variables. In England, for example, Ofsted (the national evaluating agency) provides access to RAISEonline, a database that provides school-level data on student attainment as well context variables such as students' socioeconomic background. RAISEonline can be accessed through the Internet by schools, local authorities, school improvement partners and inspectors, and is used for both external and internal evaluation. While putting in place an integrated data system of this sort is not easy, it does provide a valuable resource for schools to conduct internal evaluation using reliable and comparable data. Moreover, RAISEonline also provides tools that enable schools to analyze their data, and thus contribute to build schools' capacity (and culture) to use data in order to identify their strengths and areas for improvement. Scotland and the Netherlands have also made progress in developing an integrated data system that combines student attainment data with context variables within a single electronic school file. In general, electronic schools files are more prevalent among high-performing systems than among lower-performing ones, and high-performing systems themselves are currently facing the challenges of developing these integrated data systems –data collection, coordination, programming, updating, and other challenges.

While classroom observations are often used to inform the judgments made by external inspectors and evaluators, Van Bruggen (2010) notes that a key distinction is not whether classroom observations are part of the external evaluation process or not, but rather *how* those observations are carried out. For example, are teaching periods observed from beginning to end, are they observed partially (e.g., during 20-30 minutes), or do inspectors just take a quick glance at classroom learning as part of their walking around the school? Do classroom observations focus on student learning and teaching activities? Do they focus on behavior management? What criteria and methods are used to select the sample of classrooms that are observed (e.g., random selection vs. selection informed by the school head)? Is this sample representative of the different subjects and grades taught in the school? Does it provide an unbiased picture of the quality of teaching in the school? Are systematic frameworks, templates or methods applied in order to conduct classroom observations? Do inspectors discuss their observations with teachers, to provide them with feedback about the strengths and areas for improvement they have identified? These are all important issues on which each country must make a decision.

In addition, while classroom observations can provide valuable information about a school's teaching and learning processes, they may also have undesired side-effects. It is important to acknowledge and identify these effects, and attempt to minimize them. Negative side-effects typically include increased stress among teachers; putting on a show when classroom observation takes place; providing special instructions to students about how they should behave and who should participate while inspectors are

in the classroom; or other deviations from the way in which teachers and students usually behave and interact.

- **How much administrative work is required from schools in preparation for external evaluation, and to what extent does external evaluation interrupt schools' normal operation?**

External evaluation often, if not always, requires some administrative work from schools, such as gathering of student and teacher records and school documents; completion of forms and questionnaires specially designed for external evaluation; or circulation of questionnaires among teachers, parents or students on behalf of the external evaluating agency. This work may be required before external evaluators arrive at a school; while they are conducting the inspection/evaluation, or once the evaluators have left. The burden and timing of administrative work required by the school is likely to vary across countries. Accordingly, the extent to which this work and the evaluation process disrupt the school's normal operation will also vary considerably from one country to another. The challenge is to find the right balance between how much external evaluators benefit vis-à-vis how much teaching and learning suffer from the fact that schools need to do this administrative work.

Knowing what documents may be required from schools can contribute to reduce schools' stress and sense of burden. By anticipating what may be required from them, schools can prepare all necessary documents in advance, avoiding the stress that may occur when inspectors/evaluators request something "out of nowhere". For schools to be able to anticipate what may be required from them, evaluating agencies must be transparent and clear about what documents may be requested at the different stages of an evaluation process.

- **Do school-level evaluations include consultations with teachers, management staff, parents or students?**

External evaluation

In some countries, the processes and procedures for evaluating the quality of a school include consultations with a range of relevant stakeholders such as the school head, the school board or council, teachers, parents and/or students. Consultation processes can help evaluators gain a broader understanding of the school's strengths and weaknesses as perceived by the different actors affected by the school's quality. They can also contribute to build a good relationship and smooth communication between school staff and evaluators, especially in the case of external evaluation, thus contributing to build the legitimacy of evaluation processes and results. Moreover, consultation processes can be useful for capacity-building purposes – for example, by challenging the validity of some claims made by school staff, and encouraging them to use evidence in order to support their arguments.

While the potential benefits of consultation are many, carrying out genuine and effective consultations is not easy. In some countries, consultation processes are well-embedded in the country's history, politics and institutional culture, and transcend the education sector. For example, parliamentary systems tend to be more prone to reaching agreements through consultation processes than presidential ones – because in the former, agreement from a larger number of actors is usually necessary for making policy

decisions or approving reforms. When a country's political and institutional culture are not characterized by consultation, incorporating these as part of school evaluation processes is likely to require substantial investments in improving the communication with schools and building the interpersonal skills of evaluators; a clear definition of the extent of consultation; and mechanisms to monitor the quality of consultation processes. It is also important to acknowledge that it may take several years or even a few decades before consultations are conducted relatively smoothly.

Where consultation exists, it can take different forms, ranging from simply informing school actors about the results of evaluation, to gathering and incorporating these actors' feedback into the final evaluation report. The issue of whether schools are allowed to contest the results of evaluation is a delicate one, but enabling schools to contest results may be particularly important in systems where there is a suspicion that inspectors might be unreliable because of low capabilities or because of corrupt practices. On the other hand, if schools are allowed to contest the results of evaluation, a well-defined process for addressing their claims needs to be built in. For example, in some countries, when a school contests the results of an evaluation, the evaluation report corresponding to that school is sent for review to a higher-rank inspector/evaluator or panel. Some countries take note of schools' claims prior to publishing any final evaluation reports; other countries publish their original report but accompany it with another document which consists of the school's questioning of that report.

Internal evaluation

Internal evaluation can be more or less inclusive of the views of different school actors, depending on the style of whoever is in charge of conducting the evaluation, but also depending on the existing regulations. These may require or recommend that different school actors (e.g., students, parents, teachers, administrative staff, the school head, the school board, members of the local community) be included in the process of assessing the school's quality. Regulations may also stipulate the organizational arrangements through which these actors' participation takes place. Participatory arrangements may include, for example, membership of bodies involved in the internal evaluation of schools (e.g., the school council, school board, a body specifically constituted for evaluation, or teacher assemblies); bilateral discussions with the authority in charge of internal evaluation; questionnaires; consultation forums, etc.

As is the case with external evaluation, the school community's participation may take different forms such as consultation in the early stages of the evaluation vs. toward the very end, once most judgments about the school's quality have already been made. Indeed, through participatory processes, school actors may influence one or more of the following aspects of internal evaluation: determining criteria and procedures, gathering and analyzing data, making judgments about the school's quality, and deciding on improvement strategies.

- **Structure (phases) of the evaluation process: pre-evaluation, evaluation and follow-up**

School-level evaluations may be conducted in one or several phases. The three main phases that may exist are: **pre-evaluation**, which typically includes the gathering and analysis of school-level data and documents; **in-school phase**, which typically includes school visits and, as part of that, classroom

observations or other forms of "data collection" (e.g., talking with teachers, parents or students), and sometimes also includes the provision of feedback to the school head and/or other stakeholders (teachers, parents); and **follow-up**, which typically refers to the activities that are conducted to monitor a school's progress with respect to the evaluation's findings (e.g., the development of a school plan that outlines how the school will address the main weaknesses identified by the evaluation; the approval of this plan by the evaluators; and the subsequent monitoring of the school's progress implementing this plan).

These phases may not always exist in all countries. In a few countries, for example, the evaluation stage may not include school visits, and may be only based on school data and documents, more in line with the pre-evaluation stage of some countries. Also, in some countries there may be no follow-up stage, while in other countries the monitoring of a school's progress with respect to the evaluation results may only be done for schools in which a below-standard overall quality was observed during the evaluation. An arrangement in which follow-up activities are only conducted among low-quality schools is in line with **risk-based inspection/evaluation**[5], a trend that has recently emerged in several high-performing countries in an effort to optimize the use of resources.

<div style="background-color:#6699cc;">

5.   *With what* resources, expertise and capabilities are schools inspected and evaluated? **Capabilities.**

</div>

The previous sections discussed issues about the *design* of school inspection, evaluation and support policies. Sometimes, documenting the policies that have been enacted is sufficient to identify areas for improvement – especially when important functions have not been clearly assigned to any authority, when two or more authorities have overlapping responsibilities, when accountability mechanisms for the exercise of certain functions are weak or inexistent, or when regulations that should be well-coordinated with one another are inconsistent or come to conflict. However, in developing countries it is common to observe that a set of laws that are coherent, well-thought and in line with the policies of some high-performing countries are in place, and yet fail to produce the desired results. Often, the reason for this is not the in policies' design but in their implementation, which is far from what the laws and regulations stipulate.

Indeed, the proper implementation of school inspection, evaluation and support policies is just as important as the design of these policies. Moreover, when designing a country's school inspection, evaluation and support policies, it is important to take stock of the country's capabilities to implement these policies, and ensure that the policies that are enacted can be properly implemented. Putting in place policies for which there are insufficient implementation capabilities can be counterproductive – it can harm the relationship between the State and schools, the credibility of the evaluation process and the overall legitimacy of the policies in place. In particular, the adequate implementation of school

---

[5] A preliminary assessment is made of each school, taking into account a few key indicators of school quality. Then, if a school is suspected to provide below-standard educational services, it is evaluated more comprehensively (i.e., the duration of evaluation is longer). Moreover, if the results of the comprehensive evaluation show that in fact the school has serious deficiencies, then it is also evaluated more frequently, and its progress is monitored more closely.

inspection, evaluation and support policies requires the right human, information and financial resources. Each of these is discussed in the rest of this section.

- **Human resources and expertise**

Evaluating the quality of an individual school in a way that is conducive to educational improvement is not an easy task. After reviewing several empirical studies about the effects and side-effects of external evaluation, de Wolfe and Janssens (2007) conclude that one of the key factors that explain the extent to which evaluation is conducive to school improvement is **the relationship between school staff and the inspector/evaluator**. The latter's capacity to develop a strong relationship with teachers, school heads and the educational community affects the legitimacy of external evaluation, the extent to which schools view evaluation as an opportunity to learn and improve, and whether they are open to receive and incorporate feedback from external evaluators. Similarly, a strong relationship between the authority in charge of internal evaluation (e.g., the school head, the school board or council, or a teachers committee) and the rest of the school can be a key determinant of the extent to which evaluation results and proposals for improvement are viewed as valid and legitimate across the board.

External evaluation

A strong relationship is one where there is mutual trust, respect for others' views, and where schools are allowed and encouraged to question inspectors/evaluators. Developing this type of relationship requires that inspectors/evaluators have solid communication and interpersonal skills to encourage dialogue, openness and trust, as well as a very good understanding of the education sector that inspires respect among school staff. Anecdotally, in the Netherlands, when the Education Inspectorate was created over 200 years ago, inspectors were referred to as "wise men" because they brought to schools knowledge that they could apply and that was based on their observation of many other schools and the education system as a whole. In fact, that is a comparative advantage of external evaluation vis-à-vis internal evaluation: that it can bring to schools knowledge *from outside* of what other schools are doing to improve the quality of education – including best practices but also the practices of schools with better results and a similar context (e.g., students' socioeconomic background, geographic location).

In line with this role, **several years of work experience in the education sector** are usually required to become an inspector/evaluator, whether as a teacher, school head or administrative employee. Solid communication, interpersonal and analytical skills are also highly valued in the profession. In recent years, management skills have become a plus, and many evaluating agencies offer training to strengthen their employees' skills in this area.

High-performing systems use **competitive processes to recruit** individuals with this profile. Job vacancies are posted in several venues, such as newspapers and online job-search websites. Often, a committee consisting of the chief inspector and other inspectors select a long list of candidates from the pool of applicants. Interviews are then conducted among these by the selection committee, and a short list is selected. Then, a second committee is put together that includes the senior chief inspector, the chief inspector, inspectors and other staff. They interview short-listed candidates and make a decision after,

among other things, consulting the candidates' former employers and colleagues. Interviews and references are particularly important to make hiring decisions (The World Bank 2010a).

Once in the profession, three key factors can shape the quality of an inspector/evaluator's work: the existence and enforcement of a Code of Ethics; the availability of training opportunities through which inspectors/evaluators can continue to develop their skills; and the incentives to perform. The existence and enforcement of a **Code of Ethics that is specific to inspectors/evaluators** and not the same as the one that could apply to public sector employees is important to outline the behavior that is expected of inspectors/evaluators. In some high-performing systems, the evaluating agency makes sure that this Code of Ethics is publicly disseminated, so that schools also know what they should expect from inspectors in terms of personal integrity, objectivity, impartiality and other values that are essential to the creditworthiness of the evaluation and the relationship with schools. In fact, in New Zealand, there are two codes of ethics: one, more extensive, for use by inspectors and internally within the Education Review Office; and another one, more concise, that is available to the general public so that they know what they should expect from inspectors. Enforcement of the Code of Ethics can be done through a complaints procedure that enables schools to file claims against improper conduct by inspectors; or by assessing compliance with the Code as part of the individual performance evaluation process.

Initial training of new inspectors/evaluators can be useful to disseminate the organizational mission of the evaluating agency, the purpose of evaluations, how inspectors/evaluators are expected to contribute to the institutional mission, and the contents of the Code of Ethics. Like in the teaching profession, it may be advisable for new inspectors/evaluators to receive mentoring from a more experienced peer and assume their responsibilities gradually. **Continuous professional development** is also important, because evaluation policies and regulations, institutional values within the evaluating agency, the priorities in evaluation work and the skills required to conduct effective evaluations are all likely to change over time and accompany changes in the education system. Training activities for inspectors may include internal workshops, seminars and similar activities, but also exchanges with inspectors/evaluators from other education systems through participation in international workshops; visits to other inspectorates; online collaboration through a network of inspectors/evaluators. In Ireland, the Netherlands, New Zealand and Scotland, around 1 percent of Inspectorates' budget is allocated to initial training, and roughly 2 percent is allocated to continuous professional development (The World Bank 2010a).

Individual and group performance evaluations can be part of the incentives in place to encourage inspectors/evaluators to perform at their best. Performance evaluations were discussed in the treatment of question 1 (Who inspects and evaluates schools? Do (what) mechanisms exist to assure the quality of inspection and evaluation, and hold inspecting and evaluating authorities accountable?). In addition to this type of incentives, the **salary and non-salary benefits** perceived by inspectors/evaluators, as well as their employment conditions, can be an important determinant of the quality of applicants attracted to the profession and of those who remain in it. Ideally, to attract the best education practitioners into the profession, the benefits perceived should be comparable to what individuals with the same qualifications would obtain in the private sector.

The **employment status** of inspectors/evaluators can vary from being **civil servants** to being hired as **contract employees** for specific tasks. In some countries, all inspectors/evaluators are civil servants, and as such they enjoy the same benefits as other public employees (e.g., pension benefits, health insurance) and a considerable level of job stability. In other countries, the vast majority of inspectors are hired as contract employees, which means they are entitled to the same benefits as any person who works in the private sector, although they may be able to negotiate their terms of employment directly with their employer, and their position on the job is less secure and often expires after a given period or the fulfillment of specific tasks. Both the civil servant and the contract employee options have pros and cons. The former may be good in terms of attracting talented individuals into the profession, but may be less effective at making sure that only the good performers remain in it. Conversely, the latter may provide benefits (including job stability) that are less attractive for talented applicants, but also provide more flexibility to remove inspectors/evaluators who repeatedly fail to perform or whose conduct is in conflict with the Code of Ethics. A third group of countries has a mixed labor force, with a stable group of inspectors/evaluators who are usually hired as civil servants, and another group of temporary inspectors/evaluators who are hired as contract employees for specific tasks which depend on their background and skills. For example, one temporary inspector/evaluator may be hired specifically to participate in the observation of classrooms where mathematics is taught; another one may be hired specifically to evaluate how well schools address the differential needs of students, etc. In Scotland, for instance, a teacher from a different school than the one being evaluated may be hired to participate in the evaluation of the subject(s) he/she teaches. This adds another layer of objectivity and peer review in the evaluation process.

In a few countries, such as Scotland and England, the figure of the "**Lay inspector**" also exists. Lay inspectors form part of evaluation teams as representatives of the public. As such, they "come from different walks of life and must not have any significant experience in the management or provision of education. [Their role is to] take a particular interest in how the school and its students interact with parents and the local community, the children's comfort and their personal and social development" (Eurydice 2004, 113).

In addition to the policy decisions on how to recruit, retain and motivate individuals with the right skills for evaluating schools, a key issue is to determine the size of the inspector/evaluator labor force. This decision will vary from country to country, depending on the resources available; number of schools and how densely or sparsely they are located; tasks assigned to inspectors/evaluators (can range from conducting the in-school evaluation to also carrying out the pre-evaluation analysis and/or the follow-up); regulations regarding evaluators' working time; efficiency of the external evaluating agency (for example in terms of internal coordination and management of information resources) and other factors. Given that these factors will vary across countries, it is hard (not to say impossible) to determine an optimal number of evaluators per school that applies in to all countries. That said, it may be worth considering that there may be economies of scale in external evaluation – in other words, that in countries with a larger the number of schools, each inspector/evaluator may be able to evaluate a larger number of schools. Indeed, preliminary and limited data from European countries supports this hypothesis, which deserves further research (see Figure 5).

Figure 5. Positive relationship between number of schools evaluated each year and number of schools evaluated per inspector, selected countries
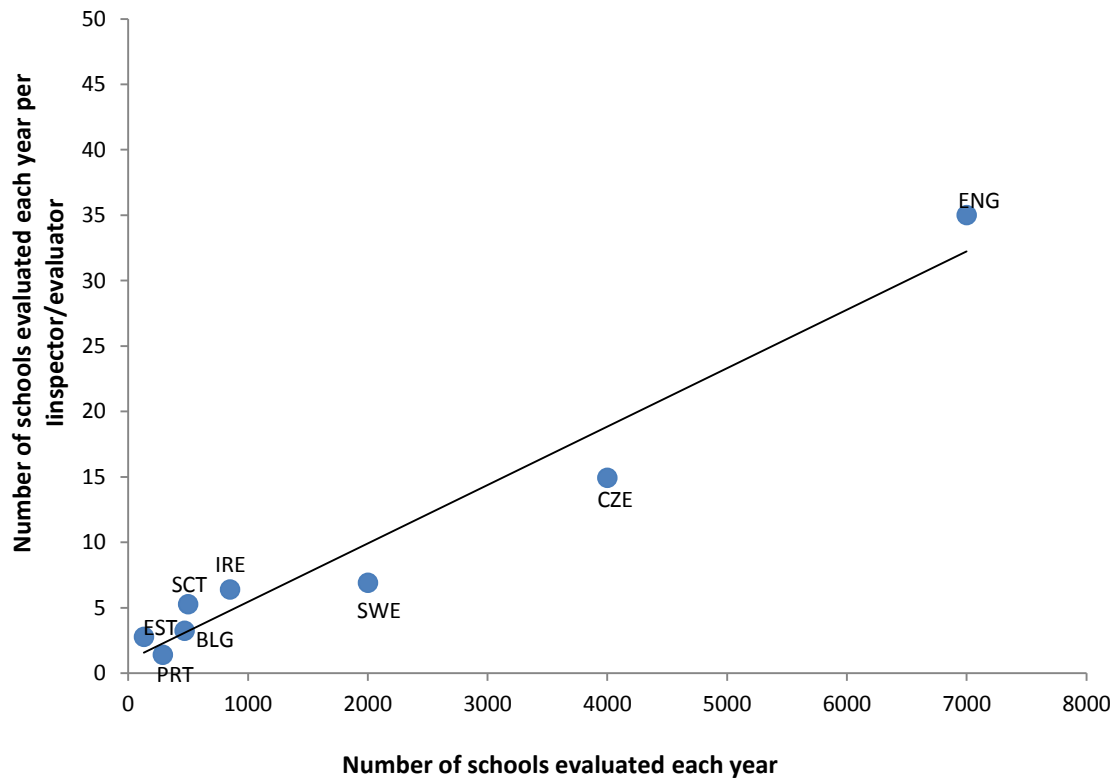
Source: Author based on Van Bruggen 2010.

In recent years, several external evaluating agencies have noticed the need to hire full-time analysts, web programmers and communication experts to complement the work of inspectors/evaluators. Analysts are particularly relevant in countries where there is a pre-evaluation phase in which many different sources and data are looked at in order to form a preliminary judgment about a school's quality on some areas/dimensions. While inspectors/evaluators may be involved in this phase as well, providing a big-picture view, the detailed analysis of all the data may be delegated to specialists who perhaps have relatively less experience and understanding of the education sector but sharper analytical skills, and particularly the statistical and computer skills to handle large databases. Moreover, the trend toward building digital school dossiers that contain all the information for one school and can facilitate national-level analyses about the quality of the education system requires not only the hiring of analysts but also of individuals with a solid training in information technologies who can build and maintain these databases. Other trends, such as the online dissemination of individual school reports produced after each evaluation, also require web programmers and designers. While some agencies choose to outsource some of these tasks, others choose to do them in-house.

Internal evaluation

When regulations are set to require or recommend that schools conduct their own evaluations, a long way remains to be traveled in terms of building schools' capacity to implement these requirements/recommendations. Schools will need to organize themselves to conduct internal evaluations; improve their administrative, student and teacher records; upgrade their information and systems and technological infrastructure; collect data they do not have; get acquainted with analyzing data and using it to inform their judgments about their strengths and weaknesses; build the capacity to leverage on their strengths in order to improve their weaknesses; design and implement strategies to address their weaknesses; gradually incorporate into the institutional culture the importance of regularly monitoring the school's progress; and improve the school's transparency and accountability to students, parents, local communities and governments.

In some countries, once regulations stipulating that schools should conduct internal evaluations are approved, this is followed by the flourishing of a private sector market that offers consulting services to support schools' newly-acquired evaluation responsibilities. These services may be very useful, especially at the beginning, when schools have not yet acquired the capacity to conduct their own evaluations. Ideally, the provision of these services should be regulated by the State (for example through an accreditation process) to ensure that schools receive quality support. In other countries, the public sector is directly in charge of providing support to schools as they transition to and implement their internal evaluation responsibilities.

Private consulting and support services may have an undesirable effect – they may create dependency by schools, especially when these delegate all or a majority of evaluation tasks to external experts, thus failing to contribute to the goal of modifying schools' culture so that they will regularly monitor themselves and devise strategies for their own improvement. Indeed, because they are usually profit-oriented, private providers may have an incentive to prevent schools from becoming independent. Moreover, reliance on private consulting firms that evaluate schools' quality may put in danger the legitimacy of external evaluations conducted by the State – in particular, schools may question the use of external evaluation given that their own evaluations are conducted by experts as well. All this means that the regulation of this private market should be carefully thought and certain adverse effects should be anticipated and try to be mitigated.

Whether provided by the public or private sectors, or by both, support services for internal evaluation may take many different forms such as: specialized training of the authority in charge of evaluation (e.g., the school head or school board) and of other participating individuals, such as teachers; development and provision of school evaluation frameworks, detailing the school quality areas/dimensions, indicators, variables, grading or classification scale, and standards to be used; guidelines and manuals on the processes and procedures of school evaluation; development and provision of databases (such as RAISEonline in England); comparative research on best practices; comparative analyses with schools that have a similar context (e.g., student socioeconomic background, geographic location); online resources, including access to school and teacher networks; and assignation of a resource person who is an expert in evaluation and school improvement and is available for consultations at any time.

- **Information resources and capabilities**

School-level evaluation requires the gathering of data at many different levels (e.g., students, teachers, schools, parents, local communities), which are obtained through many different sources (e.g., national standardized tests of student achievement, administrative records of the school, questionnaires administered among students/teachers/parents, interviews, school documents, classroom observations, etc.). To facilitate the analysis and interpretation of all these different types of data, it is convenient to have integrated data systems for each school, in which the pieces of information are organized in some purposeful way (e.g., according to their source or the quality area/dimension whose judgment they inform). Integrated information systems which include many different types of data for each school are convenient for both external and internal evaluation. In the case of external evaluation, these integrated systems can not only facilitate the analysis done prior to and after an individual school visit, but they can also facilitate aggregate analyses about the quality of education at the national or sub-national levels.

The technological advances observed over the past decade or so can contribute to facilitate analyses about the quality of a school or system. Software packages can be designed to enable comparisons across different schools, or across local or sub-national jurisdictions; to evaluate a school based on the parameters selected by the user; to identify the main strengths and weaknesses of a single school or a whole education system; to show a school's progress over time, or to perform other analytical tasks which are at the core of school-level evaluations.

Often times, integrated and automated information systems require the pulling of information from many different databases (e.g., the national evaluating agency, a separate agency that administers national standardized tests of student achievement, the Ministry of Education, local educational authorities, schools, etc.). This effort requires coordination with other agencies, so that the quality of information gathered by each is of acceptable level, the information is reported in an interpretable way, and the schedules for updating the different types of information are respected.

- **Financial resources**

External evaluation

A key question that policymakers may have concerns the amount of resources required to put in place and maintain an effective school evaluation system. Unfortunately, in order to answer this question, much more research is necessary on, at a minimum, two fronts: research on the effectiveness of different school inspection, evaluation and support systems; and documenting of the costs and financing mechanisms of these systems. How much countries spend on external evaluation will likely depend not only on their resources, policy priorities, and how much relative importance they give to external evaluation vis-à-vis internal evaluation.

Internal evaluation

A central question about schools' capabilities to conduct internal evaluation activities has to do with the support, technical as well as financial, that schools receive from the State to carry out their evaluation

responsibilities. Providing schools with adequate financial resources for internal evaluation may enable them to hire consulting services or on-site staff that can help with the evaluation process including data collection, data analysis, and the identification of strategies for improvement. Little is known, however, about what is an "adequate" level of financial resources.

6. *For what* are the results of inspection and evaluation used? **Uses and consequences.**

The use of the results of inspection and evaluation should be tightly linked to the original purposes of evaluation. However, it is possible that the implementation of policies is not consistent with their design and, therefore, that the actual use of the results of inspection and evaluation does not meet the goals for which these were introduced. For example, an external evaluation system may be introduced with the primary goal of collecting information that supports school improvement, yet lack the technical capabilities or resources to provide such support. Thus, in understanding an education system's inspection, evaluation and support policies, it is equally important to understand their purpose (a policy design issue) as well as the actual consequences that follow inspection/evaluation (a policy implementation issue).

- **Are evaluation results disseminated publicly?**

A first question that policymakers should ask themselves is, "Are the results of internal/external evaluation available to the general public?" When one of the purposes of evaluation is to hold schools and/or the education system as a whole accountable for their performance, then the answer to this question will likely be "yes" if policies are well-implemented. There are many different ways in which results can be disseminated, and therefore even among countries that have accountability purposes and consistent implementation of this purpose, differences in how results of inspection/evaluation are used are likely to be encountered. In particular, for both internal and external evaluation, it is possible that (i) the individual school report that results from the evaluation is publicly disseminated to everyone who wishes to access this report; (ii) the individual school report is disseminated among certain groups, such as the school's parents, the local community and the local educational authority; (iii) the individual school reports are used to produce a national report about the quality of the education system as whole, or to produce sub-national reports, or to produce aggregated reports but of some particular aspect of the education system (e.g., the quality of teaching and learning processes, or even more narrowly, the quality of teaching and learning mathematics). It is also possible that the results of internal/external evaluation are reported to the general public or to a sub-group even for non-accountability purposes. In particular, among systems where providing support for school improvement is a central purpose of inspection and evaluation, educational authorities may use the results of internal/external evaluation to identify and disseminate best practices at the school level.

The question of what to disseminate and to who presents a number of trade-offs. While at first sight one might think that providing a lot of information about individual school quality to as many actors as relevant will be most conducive to school improvements by imposing reputational incentives on schools, there is a risk in disseminating results to everyone, and in particular to the media. It is likely that the media will use the results to produce school rankings which might be weak methodologically, and which

might not serve the purpose of focusing schools on improvement efforts. For example, the weakest schools might be frustrated to see themselves at the bottom of the ranking, and that frustration and feelings of self-defeat might inhibit the introduction of comprehensive reforms. Frustration is not the only possible outcome among these schools; they might also feel that the ranking is unfair, for example if it does not account for differences across schools in the socio-economic background of their students, and that sense of unfairness might be projected on to the evaluation process in general, questioning its legitimacy.

This is not to say that the media should not have access to individual school reports. What is important is for policymakers to be aware of the potential problems this could bring, and seek ways to mitigate those problems. Mitigation mechanisms might include being very clear about the methodology used to evaluate schools during external evaluation, specifying that the results should not be used to produce school rankings (when that is not the purpose of the evaluation), and publishing press releases that question newspaper articles when these reflect an inappropriate interpretation of the school reports.

Aggregate reports about the quality of the education system as a whole or about certain aspects of it are also an interesting option to be considered. These reports, which often rely on the results of external evaluation in individual schools, can serve as an important mechanism for holding the Ministry of Education accountable its policies. This possibility, however, is more likely to be materialized in systems where the agency or department responsible for external evaluation is at least operationally independent from the Ministry of Education. In the Netherlands, for example, the Inspectorate of Education published an annual report about the quality of the education system as a whole, and the Minister of Education must respond to this report before Parliament.

- **What are the consequences of unsatisfactory evaluation for an individual school?**

In addition to the reputational consequences of publishing a school's unsatisfactory results in an evaluation, other consequences can also be considered. Examples of consequences that might results from external evaluation include: (i) giving more support to the weakest schools; (ii) evaluating the weakest schools more often; (iii) recommending or requiring specific changes to be introduced; (iv) requiring schools to develop their own improvement plans, to be approved by a higher authority; (v) withholding or withdrawing funds; (vi) intervening a school; and/or (vii) closing a school.

- **What are the consequences of highly satisfactory evaluation for an individual school?**

In addition to the reputational rewards of publishing a school's highly satisfactory results in an evaluation, other consequences could also be considered, such as giving the school's employees a monetary bonus, benefits to pursue advanced education courses, or more paid time to participate in professional development activities. Other rewards could include creating a network of high-quality schools that exchange practices with one another and that serve as a model for the other schools. Rewarding high quality might be just as powerful as sanctioning low quality to promote quality, an issue that deserved further exploration in the empirical research, especially given that from a political perspective it is often more feasible to introduce rewards than sanctions.

III.   Diagnostic Framework

Based on the conceptual framework discussed in Section II, this section summarizes the key issues that characterize and differentiate school inspection, evaluation and support systems. Answering these key questions and issues (Table 3) for a set of countries can yield comparable data about these systems, which can in turn inform a country's subsequent decisions in this area if the country has specific role models that it wants to follow; or, in the absence of role models, the exercise itself could help a country identify a broad set of alternative policies and practices and in turn inform the country's policy choices.

Table 3. Diagnostic framework for documenting countries' school inspection, evaluation and support policies

Phase	Key question	Key issues	Possible answers
Policy design	1. ***Who*** inspects and evaluates schools? **Governance arrangements**	1.1. Who is in charge of inspection and evaluation functions?	National, sub-national, local and/or school authorities
		1.2. What is the relationship and balance between different types of inspection and evaluation?	Independent; external evaluation based on internal evaluation and/or internal evaluation based on external evaluation
		1.3. To what extent are external evaluations and judgments independent from political interference?	Evaluating authority is legally independent from Ministry of Education; operationally though not legally independent from it; or not independent from the Ministry
		1.4. Do (what) mechanisms exist to assure the quality of inspection and evaluation, and hold inspecting and evaluating authorities accountable?	Feedback about inspectors from schools and/or colleagues and managers; and/or external auditing/evaluation
	2. ***Why*** are school inspections and evaluations conducted? **Purpose**	2.1. What level of decision-making does external evaluation seek to inform?	National, sub-national, local and/or school authorities
		2.2. What type of decisions does external evaluation seek to inform?	Support and/or accountability
		2.3. What level of decision-making does internal evaluation seek to inform?	National, sub-national, local and/or school authorities
		2.4. What type of decisions does internal evaluation seek to inform?	Support and/or accountability

Phase	Key question	Key issues	Possible answers
Policy design	3. *What* is evaluated? **School quality criteria**	3.1. Do common criteria for evaluating the quality of schools exist?	Yes/No
		3.2. What criteria are used to evaluate school quality?	
		3.2.1. What areas/dimensions and indicators are considered to evaluate school quality?	- Student outcomes: academic attainment, cognitive learning skills, social and emotional skills, repetition and dropout rates, destination of school leavers, attendance - Teaching and learning processes: Curriculum/subjects offered by the school, teaching processes, learning processes, addressing of students' individual needs - Management processes: School climate, leadership and vision, organization, administrative procedures, communication within the school, time management, financial management, building management, human resource management, internal evaluation and school development, relationship with wider community - Context/inputs: Local socio-economic, educational and geographic characteristics; students' socio-economic, health, linguistic/cultural background; teachers' socio-economic, educational, linguistic/cultural background; infrastructure (buildings, funding, science labs, ICT); balance of intake of students (by ethnicity, gender, language skills)
		3.2.2. How many values can areas/dimensions, indicators and variables of school quality take?	[Number]
		3.2.3. What are the names of the categories into which areas/dimensions, indicators and variables can fall?	[Categories]
		3.2.4. Do evaluators use common definitions of what constitutes the standard for a certain dimension/indicator?	Yes/No
		3.3. How were common school quality criteria determined?	Broad/narrow/no consultation process to agree on a definition of "educational quality"; Testing of evaluation framework and criteria; Evaluation of pilot experience; Revision of pilot framework and criteria before scale-up

Phase	Key question	Key issues	Possible answers
Policy design	4. *How* are evaluations conducted? **Processes and procedures**	4.1. How often is each school evaluated? How long does each evaluation last?	[Frequency and duration]
		4.2. What sources of information are used to inform judgments about a school's quality?	Classroom observations; standardized test scores; samples of students' work; student records; administrative records; school documents; feedback from teachers, school management, parents and/or students; the school's internal evaluation report.
		4.3. How much administrative work is required from schools in preparation for external evaluation, and to what extent does external evaluation interrupt schools' normal operation?	High/Moderate/Low
		4.4. Do school-level evaluations include consultations with teachers, management staff, parents or students?	Yes/No
		4.5. What is the structure of the evaluation process?	Pre-evaluation; Evaluation; Follow-up

Phase	Key question	Key issues	Possible answers
Policy implementation	5. *With what* resources, expertise and capabilities are schools inspected and evaluated? **Capabilities**	5.1. Human resources	
		5.1.1. What qualifications are required to become an inspector/evaluator?	Educational requirement; Special courses in evaluation; Passing a written exam; Teaching experience; Experience in the education sector; School management experience
		5.1.2. What kind of training is provided to inspectors/evaluators?	Induction/mentoring for new inspectors; In-service professional development for all staff
		5.1.3. Are inspectors evaluated regularly for their performance?	Yes/No
		5.1.4. Is there a Code of Ethics specifically for inspectors, different than the Code of Ethics that applies to public employees in general?	Yes/No
		5.1.5. What is the employment status of inspectors?	Civil servant/ Permanent contract but not civil servant/ Temporary employment contract with some benefits/ Temporary employment contract without benefits
		5.1.6. How many inspectors are there?	[Number]
		5.1.7. How many schools are evaluated by each inspector each year?	[Number]
		5.1.8. How many inspectors participate in an evaluation?	[Number]
		5.1.9. How does the State help internal evaluators build their capacities?	Training; Evaluation framework; Guidelines and manuals on procedures; Exchange of experiences through school networks; Appointment of a resource person
		5.2. Information resources	
		5.2.1. Do external evaluators have access to electronic school files?	Yes/No
		5.2.2. Do external evaluations rely on national data and statistics?	Yes/No
		5.2.3. Do internal evaluators have access to electronic databases?	Yes/No
		5.2.4. Are internal evaluators required to use national data and statistics?	Yes/No
		5.3. Financial resources	
		5.3.1. What is the annual budget for external evaluation?	[Dollars]
		5.3.1. Do schools have access to public funds specifically for internal evaluation?	Yes, all/ Yes, some/ No

Phase	Key question	Key issues	Possible answers
Policy impact	6. *For what* are the results of inspection and evaluation used? **Uses and consequences**	6.1. Are evaluation results disseminated publicly?	Yes/No
		6.1.1. Are individual school reports that result from external evaluation made publicly available?	Yes/No
		6.1.2. Are thematic reports based on the cross-sectional analysis of external evaluation school reports published?	Yes/No
		6.1.3. Are national reports based on the cross-sectional analysis of external evaluation school reports published?	Yes/No
		6.1.1. Are individual school reports that result from internal evaluation made publicly available?	Yes/No
		6.1.3. Are local reports based on the cross-sectional analysis of internal evaluation school reports published?	Yes/No
		6.1.3. Are national reports based on the cross-sectional analysis of internal evaluation school reports published?	Yes/No
		6.2. What are the consequences of unsatisfactory evaluation for an individual school?	Reputational sanction through publication of results; professional development requirements; salary sanctions; firing of teachers or school administrator; school closure; school fees; withdrawal or withholding of public funds
		6.3. What are the consequences of highly satisfactory evaluation for an individual school?	Reputational reward through publication of results; professional development or educational benefits for teachers; salary increases or bonuses

IV.   A Quick Look at Eighteen Selected Systems

This section applies the diagnostic framework proposed in the previous section to document and compare the school inspection, evaluation and support policies of eighteen education systems in Europe including above-, equivalent- and below-OECD performers in PISA 2009. The countries included are Finland, Netherlands, Estonia, Germany, Poland, Slovenia (above-OECD performers); Denmark, Ireland, Hungary, France, the Czech Republic and Sweden (equivalent-OECD performers); and Slovak Republic, Portugal, Spain, Lithuania, Italy and Romania (below-OECD performers). A more detailed categorization of these systems with respect to PISA 2009 is shown on Table 1. The focus on European systems was mostly driven by the fact that there have already been efforts to document the school inspection policies in these systems (see Van Bruggen 2010; Eurydice 2004; Standaert 2000). I therefore rely on the data collected by these previous efforts, namely Van Bruggen (2010) and Eurydice (2004), to illustrate how key issues can be addressed in different ways. Because I rely on the data that has already been collected by others, the framework is not applied perfectly – data on a few questions of the proposed framework have not been collected before, especially for questions concerning internal evaluations.

Table 4. Selected countries and categorization based on PISA 2009

Category	Color Reference	Country	Abbreviation	Performance in PISA 2009
Above-OECD systems		Finland	FIN	Above OECD in Reading, Math and Science
		Netherlands	NLD	Above OECD in Reading, Math and Science
		Estonia	EST	Above OECD in Reading, Math and Science
		Germany	DEU	Above OECD in Math and Science; equivalent in Reading
		Poland	POL	Above OECD in Reading and Science; equivalent in Math
		Slovenia	SVN	Above OECD in Math and Science; below in Reading
Equivalent-OECD systems		Denmark	DNK	Equivalent to OECD in Reading and Science; above in Math
		Ireland	IRL	Equivalent to OECD in Reading; above in Science; below in Math
		Hungary	HUN	Equivalent to OECD in Reading, Math and Science
		France	FRA	Equivalent to OECD in Reading, Math and Science
		Czeck Republic	CZE	Equivalent to OECD in Math and Science; below in Reading
		Sweden	SWE	Equivalent to OECD in Reading and Math; below in Science
Below-OECD systems		Slovak Republic	SVK	Below OECD in Reading and Science; equivalent in Math
		Portugal	PRT	Below OECD in Math and Science; equivalent in Reading
		Spain	ESP	Below OECD in Reading, Math and Science
		Lithuania	LTU	Below OECD in Reading, Math and Science
		Italy	ITA	Below OECD in Reading, Math and Science
		Romania	ROM	Below OECD in Reading, Math and Science

Source: Author based on PISA 2009.

1.   *Who* inspects and evaluates schools? **Governance arrangements.**

This section applies the proposed diagnostic framework to document the governance arrangements for school inspection, evaluation and support. The arrangements of the selected systems are summarized in Table 5.

As shown in Table 5, all selected European countries require or recommend school authorities to conduct their own assessments of the school's quality (internal evaluation). The large majority of countries also require that schools be evaluated by a national authority –with the exceptions being Finland, where external evaluation is conducted by local authorities, and Italy, where external evaluation procedures are not in place–, and some countries have both national and local external evaluations (Estonia, Hungary, Czech Republic, Slovak Republic, Lithuania, Denmark and Sweden). Overall, countries can be grouped into five categories: (i) those that require external evaluation by a national authority and internal evaluation by a school authority (Netherlands, Germany, Slovenia, France, Spain and Romania); (ii) those that require external evaluation by a national authority and by local authorities as well as internal evaluation (Estonia, Poland, Czech Republic, Slovak Republic and Sweden); (iii) those that require external evaluation by local authorities and internal evaluation (Lithuania); (iv) those that require national and/or local external evaluation and recommend but do not require internal evaluation (Finland, Denmark, Ireland, Hungary and Portugal); (v) and those that require internal evaluation but not external evaluation (Italy).

No clear distinctions exist between above-, below- and equivalent-OECD systems in terms of the distribution of school evaluation responsibilities. It does seem to be the case that the *number of required types of evaluations* (external national, external local, internal) is smaller among countries that perform below the OECD average in PISA 2009 (on average, 1.8 types of evaluations) than among countries that perform at levels comparable to or above the OECD average (on average, 2.2 types of evaluations). This does not mean that having different types of evaluations will lead to better educational outcomes; the relationship could just occur by chance, or it could be that certain characteristics of the education system such as the public-sector's capacity lead to good educational outcomes *and* enable countries to implement various different types of evaluations.

• **What is the relationship and balance between different types of inspection and evaluation?**

Among the selected systems, there are six countries in which external and internal evaluation are independent from one another (Estonia, Germany, Hungary, Slovak Republic, Portugal and Spain); in seven countries, external evaluators rely on the results of internal evaluation to inform their own judgments about a school's quality (Poland, Slovenia, Denmark, Ireland, France, Sweden, Lithuania and Romania); and in two countries external and internal evaluation rely on one another (Netherlands and Czech Republic). It is the case that internal and external evaluations are more likely to be independent from one another in below-OECD systems than in above- or equivalent-OECD systems. Again, this need not mean that interdependence between to internal and external evaluation leads to more educational improvement.

- **To what extent are external evaluations and judgments independent from political interference?**

In eight of the selected systems, external evaluation is conducted by an agency or unit that reports directly to the Minister of Education but has operational independence to decide how to carry out its functions (including how to evaluate schools' quality). This is the case in the Netherlands, Germany, Denmark, Ireland, Czech Republic, Sweden, Slovak Republic and Portugal. In only country (Slovak Republic) the external evaluation authority is not just operationally but also legally independent from the Ministry of Education. There are four countries in which external evaluation is carried out within the Ministry of Education with very little independence from political interference (Estonia, Poland, France and Spain). No clear patterns emerge that differentiate the level of political independence given to external evaluation authorities in above-, equivalent- and below-OECD performers.

- **Do (what) mechanisms exist to assure the quality of inspection and evaluation, and hold inspecting and evaluating authorities accountable?**

All the selected systems for which there are data, except for Denmark, have some mechanism to monitor the quality of external inspection and evaluation. In five systems, feedback about inspectors and evaluators' work is requested from schools and/or peers, and an external auditing agency or a special unit within the evaluation agency checks the reliability of evaluation results (Netherlands, Estonia, Czech Republic, Slovak Republic and Portugal). Three countries rely only on feedback from schools and peers to monitor the quality of external evaluation (Germany, Ireland and Sweden) and one country relies only on auditing of evaluations (Spain). No clear patterns emerge that differentiate the existence and type of accountability mechanisms for external evaluation in above-, equivalent- and below-OECD performers.

Table 5. Governance arrangements, selected systems

	1. *Who* inspects and evaluates schools? GOVERNANCE ARRANGEMENTS										
	1.1. Who is in charge of inspection and evaluation functions?			1.2. What is the relationship and balance between different types of inspection and evaluation?			1.3. To what extent is the external evaluation agency/unit independent from the Ministry of Education?			1.4. Do (what) mechanisms exist to assure the quality of inspection and evaluation, and hold inspecting and evaluating authorities accountable?	
	National	*Local*	*School*	*Independent*	*Interdependent*		*No independence*	*Operational independence*	*Legal independence*	*Feedback from schools and/or colleagues & managers*	*External auditing and/or evaluation*
	(external evaluation)		(internal eval.)		(external eval. based on internal)	(internal eval. based on external)					
FIN		•	•	…	…	…	na	na	na	na	na
NLD	•		•		•	•		•		•	•
EST	•	•	•	•			•			•	•
DEU	•		•	•				•		•	
POL	•	•	•		•		•			…	…
SVN	•		•		•		…	…	…	…	…
DNK	•	•	•		•			•			
IRL	•		•		•			•		•	
HUN	•	•	•	•			…	…	…	…	…
FRA	•		•		•		•			…	…
CZE	•	•	•		•	•		•		•	•
SWE	•	•	•		•			•		•	
SVK	•	•	•	•				•	•	•	•
PRT	•		•	•				•		•	•
ESP	•		•	•			•				•
LTU	•	•	•		•		…	…	…	…	…
ITA			•	na	na	na	na	na	na	na	na
ROM	•		•		•		…	…	…	…	…

Reference Key:
- • Yes
- Evaluation is Required
- Evaluation is Recommended
- … Not available
- na: Not applicable

Sources: Author based on PISA 2009; (1)-(2) Eurydice 2004; (3)-(4) SICI Blue Book 2010, Van Bruggen 2010 and Eurydice Country Reports 2001.

2. *Why* are school inspections and evaluations conducted? Purpose.

This section discusses the purpose of inspection and evaluation in the selected systems, as documented in Table 6.

- **What level of decision-making, and what type of decisions, does external evaluation seek to inform?**

In all the selected systems except for Hungary, external evaluation seeks to inform at least some decisions at both the school and system levels. Systems can be grouped into seven categories depending on what decisions seek to be informed by external evaluation: (i) systems where it seeks to inform both support and accountability decisions at both the school and system-wide levels (Netherlands); (ii) systems where it seeks to inform only accountability decisions at the school level, and both support and accountability decisions at the system level (Estonia and Portugal); (iii) systems where it seeks to inform both support and accountability decisions at the school level and only accountability at the system level (Slovenia, Czech Republic, Sweden and Lithuania); (iv) systems where it seeks to inform only support decisions at the school level and only accountability at the system level (Germany, Poland and France); (v) systems where it seeks to inform only support decisions at the school level and both support and accountability decisions at the system level (Slovak Republic and Spain); (vi) systems where it seeks to inform only accountability decisions at both the school and system level (Denmark and Romania); and (vii) systems where it seeks to inform both support and accountability decisions at the school level and only support decisions at the system level (Ireland). In Hungary, external evaluation only seeks to inform support decisions at the school level. Put differently, in most systems external evaluation has, at a minimum, school support and system-wide accountability purposes. In addition, in several countries, external evaluation also has school accountability purposes. Less common is the scenario where external evaluation seeks to inform support decisions at the system-wide level.

Interestingly, in above- and equivalent-OECD education systems there is a tendency to balance both goals, with external evaluation serving both school support and school accountability purposes. In contrast, among below-OECD systems, there is a tendency to concentrate on one of these purposes, either supporting schools or holding them accountable. A different pattern is observed at the system-level: in both above- and below-OECD systems the results of school-level evaluations are used for both policy advice (support) and to provide information to the public about the state of education (accountability); whereas in equivalent-OECD systems, school-level evaluations are rarely used for producing policy advice, and more often are used for reporting about the state of the education system.

- **What level of decision-making, and what type of decisions, does internal evaluation seek to inform?**

No clear patterns emerge about differences in the purpose of internal evaluation in above-, equivalent- and below-OECD systems. In most systems, internal evaluation has both school support and school accountability purposes, as well as both support and accountability purposes at the local educational authority level. Internal evaluation also usually has an accountability purpose in terms of informing national-level external evaluations, but not a purpose of informing policy decisions.

44

Table 6. Purpose of inspection and evaluation, selected systems

	2. *Why* are school inspections and evaluations conducted? PURPOSE									
	2.1. & 2.2. What level of decision-making, and what type of decisions, does external evaluation seek to inform?				2.1. & 2.2. What level of decision-making, and what type of decisions, does compulsory internal evaluation seek to inform?					
	School-level		System-level		School-level		Local-level		National-level	
	Support decisions	Accountability decisions	Support decisions	Accountability decisions	Support decisions	Accountability decisions	Support decisions	Accountability decisions	Support decisions	Accountability decisions
FIN	…	…	…	…	na	na	na	na	na	na
NLD	•	•	•	•	•	•				•
EST		•	•	•	•	•	•	•		
DEU	•			•	•	•	•			•
POL	•		…	•	…	…	•	•		
SVN	•	•	…	•	…	…	…	…		•
DNK		•		•	na	na	na	na	na	na
IRL	•	•	•		na	na	na	na	na	na
HUN	•				•	•	na	na	na	na
FRA	•		…		…	…	…	…		•
CZE	•	•		•	•	•	•	•		•
SWE	•	•		•	•	•	•	•		•
SVK	•		•	•	•	•	•	•		
PRT		•	•	•	na	na	na	na	na	na
ESP	•		•	•	•		…	…		•
LTU	•	•	…	•	…	…	…	…		•
ITA	na	na	na	na	…	…	…	…		
ROM		•	…	•	…	…	•			•

Reference Key:
- • Yes
- … Not available
- na: Not applicable

Sources: Author based on PISA 2009, Eurydice 2004 and Van Bruggen 2010.

3. *What* is evaluated? **School quality criteria.**

This section discusses the criteria that are used to evaluate school quality in the selected systems, whenever these have come to agreement on a common set of criteria for conducting evaluations. The criteria used in the selected systems are summarized in Table 7.

- **Do common criteria for evaluating the quality of schools exist?**

Ten of the sixteen selected systems that have national external evaluations have defined a common set of criteria for conducting these evaluations of school quality – such common criteria exist in Netherlands, Estonia, Ireland, Czech Republic, Sweden, Slovak Republic, Portugal, Spain, Lithuania and Romania. The review of selected systems suggests that, if anything, national frameworks that outline the dimensions and indicators that should be used to evaluate schools, whether recommended or required, are more likely to be found among lower-performing systems than among higher-performing ones. This is a surprising result that could reflect the use of sub-national frameworks in higher-performing systems, which is the case for example in Germany.

- **What criteria are used to evaluate school quality?**

Quality areas/dimension and indicators

The review of selected countries suggests several patterns about the school quality areas/dimensions and indicators used to evaluate the quality of education. First, all the systems that have a national framework for conducting external evaluations take into account both educational outcomes and processes when evaluating school quality.

Second, among the selected systems' national frameworks, the most common indicators of educational outcomes used are academic attainment (as measured by tests or examinations), and repetition and dropout rates. Cognitive learning skills, as measured by qualitative assessments; social skills; attendance; and the destination of school leavers are less often included in national frameworks. No obvious patterns emerge in terms of what outcome variables are taken into account in above-, equivalent- and below-OECD systems. The main difference lies in that social skills are not taken into account in any of the selected above-OECD systems, whereas they are taken into account in the majority of below-OECD systems that have national frameworks.

Third, among educational processes, all the national frameworks for which there are data include indicators of both teaching and learning processes on one hand, and school management processes on the other. Among the teaching and learning processes, all national frameworks look at the curriculum and in-classroom processes, and a majority but not all of them take into account whether schools and teachers address students' individual and differential needs. Among the management processes, all national frameworks look at schools' leadership and vision, financial management and relations with the local community. The majority of systems also take into account the school climate, the organizational structure and functioning, the internal communication, human resources management, the management of school buildings, and the policies and procedures for internal evaluation and school

improvement. The majority of systems, but not all above-OECD ones, also take into account schools' administrative procedures and the management of school time.

Fourth, school quality evaluations in high-performing systems seem to be more exhaustive and comprehensive than those in lower-performing systems. The average number of school quality indicators considered among high-performing systems is greater than the average number among lower-performing systems (18.5 school quality indicators in above-OECD systems, 17.5 in equivalent-OECD systems and 15.8 in below-OECD systems, out of a total of 21 possible indicators in Table 7).

Data have not been systematically collected on whether countries take into account the inputs and context in which education is provided. These are presumably important to take into account in order to provide a fair judgment of a school's quality, as educational outcomes may reflect not only factors that the school controls, such as teaching and management processes, but also factors that are beyond the school's direct control, such as students' linguistic background and their parents' education. Although I know of some high-performing systems which take into account schools' inputs and context when evaluating their quality (e.g., New Zealand, the Netherlands and Ireland), more data and research are needed in order to identify whether any patterns exist in terms of the inclusion of this area/dimension in high-performing vis-à-vis lower-performing systems.

Quality variables

Two systems can care about similar areas/dimensions and indicators of quality, yet still have significant differences in how they think about educational quality. While comparable data about specific indicators has not been collected, the point can be illustrated by comparing, for example, the national frameworks developed in New Zealand and Ireland. In both cases, the national framework includes teaching and learning processes among the set of key dimensions considered when evaluating schools, and within this dimension, both countries look specifically at the quality of teaching processes among other indicators. However, when looking at the specific variables used to judge the quality of teaching processes, important difference between countries can be observed. A distinctive feature of New Zealand is that many of the variables of the quality of teaching processes focus on whether teachers set high standards and expectations for their students, and whether they use data to identify student needs, design learning strategies and monitor progress[6]. In Ireland, many of the variables used to assess the quality of teaching processes focus on how teachers plan and structure their work[7]. Similarly, the national frameworks used in Romania and Lithuania include variables about the quality of teaching and learning processes. However, in Romania, these variables are more comprehensive, including measures of teaching methods, student assessment methods, the use of homework, students' participation in class, their attitude toward learning, and the interaction between teachers and student; while in Lithuania, the variables focus on teaching methods and students' participation only.

---

[6] Education Review Office. June 2010 (Draft). *Evaluation Indicators for School Quality Reviews*. Available from: http://www.ero.govt.nz/Review-Process/Frameworks-and-Evaluation-Indicators-for-ERO-Reviews.
[7] Department of Education and Science, Inspectorate. 2003. *Looking at our school. An aid to self-evaluation in second-level schools*. Available from: http://ppds.ie/sdparchive/LAOS/Contents%20-%20Looking%20at%20our%20School%20%20Primary.htm.

Possible values the variables can take

Dimensions and indicators are often classified using an even number of categories. For example, in Ireland each dimension and indicator can fall into four different categories: significant strengths (uniformly strong); strengths outweigh weaknesses (more strengths than weaknesses); weaknesses outweigh strengths (more weaknesses than strengths); or significant/major weaknesses (uniformly weak)[8]. In Scotland each dimension and indicator can be classified under one of six categories: excellent; very good; good; satisfactory; unsatisfactory; or weak[9]. Four categories are also used to classify schools in other high-performing systems such as New Zealand, the Netherlands[10] and England[11], as well as in Portugal[12]. Still, some countries use an odd number of categories to classify schools. The Czech Republic uses three categories: below standard, standard and above standard[13]; the Slovak Republic uses five: very good, good, average, less satisfactory, unsatisfactory[14].

Strengths and weaknesses are explicitly identified in all the selected systems that have a national framework for assessing school quality. The identification of strengths and weaknesses is also a common result of school-level evaluations in several systems that do not have a national framework, such as Finland, Germany and Denmark.

Standards, norms or descriptors

In the selected systems that have national frameworks and for which there are data, almost all frameworks include standards and descriptors that enable evaluators to classify a school's dimensions or indicators into one of the plausible categories[15]. Usually, descriptions are not provided for all the possible categories. For example, if four possible categories exist (e.g., very good, good, fair and unsatisfactory), descriptions of what each dimension and indicator would look like may be provided for only two categories (e.g., good and unsatisfactory). As already noted in the conceptual discussion, the standards that are made publicly available are those used to classify the quality of a certain dimension or indicator; not those used to assign values to each specific variable.

- **How were common school quality criteria determined?**

Comparable data about the process by which school quality criteria were determined have not been collected – an important area for further work.

---

[8] Department of Education and Science, Inspectorate. 2003. *Looking at our school. An aid to self-evaluation in second-level schools*. Available from:
http://ppds.ie/sdparchive/LAOS/Contents%20-%20Looking%20at%20our%20School%20%20Primary.htm.
[9] Her Majesty's Inspectorate of Education. 2007. *How good is our school? The journey to excellence.*
Part I and II available from: http://www.hmie.gov.uk/documents/publication/hgiosjte.pdf.
Part III available from: http://www.hmie.gov.uk/documents/publication/hgiosjte3.pdf.
[10] SICI. 2009. *The Inspectorate of the Netherlands*. Blue Book Series.
[11] OFSTED. 2009. *Criteria for making judgments about the quality of provision*.
[12] SICI. 2009. *The Inspectorate of Education of Portugal*. Blue Book Series.
[13] SICI. 2009. *The Czech School Inspectorate*. Blue Book Series.
[14] SICI. 2009. *The Inspectorate of Education of the Slovak Republic*. Blue Book Series.
[15] The only exception is Spain, which does not classify schools in any way and therefore does not have standards.

Table 7. School quality criteria, selected systems

		3.1. Do common criteria for conducting national external evaluations of schools exist?	3.2. According to national frameworks for external evaluation, what criteria are used to evaluate school quality?									
			Students' outcomes						Teaching and learning processes			
			Academic attainment	Cognitive learning skills	Social skills	Repetition and dropout rates	Destination of school leavers	Attendance	Curriculum/ Subjects	Teaching processes	Learning processes	Addressing of students' individual needs
	FIN	na	na	na	na	na	na	na	na	na	na	na
	NLD	•	•	•		•	•	•	•	•	•	•
	EST	•	•			•	•		•	•	•	
	DEU		na	na	na	na	na	na	na	na	na	na
	POL		na	na	na	na	na	na	na	na	na	na
	SVN		na	na	na	na	na	na	na	na	na	na
	DNK		na	na	na	na	na	na	na	na	na	na
	IRL	•	•			•			•	•	•	•
	HUN	…	…	…	…	…	…	…	…	…	…	…
	FRA		na	na	na	na	na	na	na	na	na	na
	CZE	•		•	•				•	•	•	•
	SWE	•	•						•	•	•	•
	SVK	•	•	•	•	•			•	•	•	
	PRT	•	•			•			•	•	•	•
	ESP	•	•	•	•	•		•	•	•	•	•
	LTU	•	•			•	•		•	•	•	
	ITA	na	na	na	na	na	na	na	na	na	na	na
	ROM	•	•	•	•		•		•	•	•	•

Reference Key:
- Yes
… Not available
na: Not applicable

Sources: Author based on PISA 2009 and Eurydice 2004.

Table 7 (continued)

		3. *What* is evaluated? SCHOOL QUALITY CRITERIA 3.2. According to national frameworks for external evaluation, what criteria are used to evaluate school quality?										
		Management processes										
		School climate	Leadership and vision building	Organiza-tion	Adminin-istrative procedures	Communica-tion in the school	Time managemt.	Financial managemt.	Building managemt.	HR managemt.	Internal evaluation and school development	Relations with wider community
	FIN	na	na	na	na	na	na	na	na	na	na	na
	NLD	•	•	•	•	•	•	•	•	•	•	•
	EST	•	•	•	•	•		•	•	•	•	•
	DEU	na	na	na	na	na	na	na	na	na	na	na
	POL	na	na	na	na	na	na	na	na	na	na	na
	SVN	na	na	na	na	na	na	na	na	na	na	na
	DNK	na	na	na	na	na	na	na	na	na	na	na
	IRL	•	•	•	•	•	•	•	•	•	•	•
	HUN	…	…	…	…	…	…	…	…	…	…	…
	FRA	na	na	na	na	na	na	na	na	na	na	na
	CZE	•	•	•	•	•	•	•	•	•	•	•
	SWE	•	•	•			•				•	•
	SVK	•	•	•	•	•	•	•	•	•	•	•
	PRT	•	•	•	•	•	•	•	•	•	•	•
	ESP	•	•	•	•	•	•	•	•	•	•	•
	LTU		•	•				•				•
	ITA	na	na	na	na	na	na	na	na	na	na	na
	ROM	•	•		•			•	•	•		•

Reference Key:

•   Yes

…   Not available

na: Not applicable

Sources: Author based on PISA 2009 and Eurydice 2004.

4. *How* are evaluations conducted? Processes and procedures.

A summary of the processes and procedures used to evaluate schools is shown on Table 8. Most of the information presented refers to the processes and procedures of external evaluation, because those of internal evaluation have not been very well documented by the prior studies on which this note relies.

- **How often is each school evaluated? How long does each evaluation last?**

Among the selected systems shown on Table 8, the average frequency of external evaluation is every 5.3 years among above-OECD systems, every 4.7 years among equivalent-OECD performers, and every 4.5 years among below-OECD systems. Some high-performing systems (Estonia, Belgium and Scotland) conduct an external evaluation of each school every 7 or 8 years; among lower-performing systems, Sweden conducts evaluations every 6 years, and the rest of the countries conduct evaluations more frequently, every 3 to 5 years for each school.

A similar pattern is observed at the within-system level: in some systems, higher-performing schools are evaluated less frequently than lower-performing ones. This is true for the Netherlands, Estonia, Denmark, Ireland, England, Sweden and Spain (Van Bruggen 2010).

- **What sources of information are used to inform judgments about a school's quality?**

Among the selected systems, high-performing systems use a more varied set of sources of information for national external evaluations than lower-performing ones. Among the selected lower-performing systems, a majority of countries use information contained in school documents, and to a lesser extent, standardized test scores; only a minority uses the information provided by student and administrative records, administers questionnaires to gather feedback from school stakeholders (teachers, management, parents, students), or takes into account the results of schools' internal evaluation. In contrast, almost all the selected high-performing systems take into account not only school documents and standardized test scores but also student and administrative records, and the reports produced by the school as a result of self-evaluation. These observations are not surprising – they are consistent with previous observations, for example, that national external evaluation in high-performing systems is more likely to incorporate the results of internal evaluation as well as take into account a more comprehensive list of school quality dimensions, indicators and variables. These observations –including the fact that external evaluators in high-performing systems look at a more varied set of sources of information– may simply reflect the fact that the implementation capacity and expertise in these countries is greater.

Table 8 also shows that the vast majority of selected countries for which there are data (all except Denmark) use classroom observations to inform their judgments about a school's quality –in particular, the quality of teaching and learning processes. As discussed previously, the difference between countries may lie in how these observations are conducted (e.g., how long each observation lasts, how the sample of classrooms that is observed is selected, whether a common methodology is used to observe classrooms, etc.).

- **How much administrative work is required from schools in preparation for external evaluation, and to what extent does external evaluation interrupt schools' normal operation?**

In European countries, most administrative work required from schools must be done before external evaluators arrive at the school (Van Bruggen 2010). In fact, the documents, questionnaires and forms submitted by schools are used to prepare for the inspection/evaluation visits. In some countries, external evaluating agencies have analysts who look at these documents and provide a summary of observations to evaluators. In other countries, evaluators themselves conduct this analysis. Observations can be made on, for example, student and teacher attendance; dropout and repetition rates; compliance with certain laws and regulations; the quality of record and data management systems at the school; the socioeconomic background of students and teachers; or teachers, parents and students' level of satisfaction with different aspects of the school.

- **Do school-level evaluations include consultations with teachers, management staff, parents or students?**

In seven of the thirteen systems with national external evaluation for which there are data, the external evaluation process includes consultations with teachers, management staff, parents or students. Consultations are included in two above-OECD systems (Netherlands and Estonia); two equivalent-OECD systems (Ireland and Czech Republic) and three below-OECD systems (Slovak Republic, Spain and Romania).

Nearly all systems require or recommend that teachers be consulted in the process of carrying out internal evaluations (the exception is Portugal) and all require or recommend that school heads be consulted too. In above- and equivalent-OECD systems, the school board is also usually included in the internal evaluation process.

- **Structure (phases) of the evaluation process: pre-evaluation, evaluation and follow-up**

Among the selected systems, the large majority of countries have pre-evaluation and in-school phases as part of the external evaluation process. However, follow-up stages for some or all schools are more prevalent among above- and equivalent-OECD systems than among below-OECD ones.

Table 8. Processes and procedures of evaluation, selected countries

		4. *How* are external evaluations conducted? PROCESSES AND PROCEDURES									
		4.1. How often is each school evaluated externally? How long does each external evaluation last?			4.2. What sources of information are used to inform judgments about a school's quality during external evaluation?						
		Frequency of evaluation	*Number of schools evaluated every year*	*Duration (in days)*	*Classroom observation*	*Standardized test scores*	*Student and administrative records*	*School documents*	*Feedback from teachers, management, parents or students*	*Internal evaluation report*	*National data and statistics about the school*
	FIN	na	na	na	na	na	na	na	na	na	na
	NLD	every 1-4 years	…	1-4	•	•	•	•	•	•	•
	EST	every 8 years	130	1-5	•	•	•	•	•		•
	DEU	every 5-6 years	1080	1-4	•	…	…	…	…	…	•
	POL	…	…	…	…			•		•	…
	SVN	…		…	…	•	•	•		•	…
	DNK	…	…	1-10		…	…	…	…	…	•
	IRL	every 4 years	850	2-6	•		•	•	•	•	•
	HUN	…	…	…	…	…	…	…	…	…	…
	FRA	…	…	…	…	•	•				…
	CZE	every 4 years	4000	2-5	•		•	•	•	•	•
	SWE	every 6 years	2000	2	•	•	•	•		•	•
	SVK	every 5 years	…	2-5	•		•	•	•	•	•
	PRT	every 4 years	287	1-2	•	•		•			•
	ESP	every 3-6 years	…	3-4	•	•		•	•		•
	LTU	…	…	…	…			•			…
	ITA	na	na	na	na	na	na	na	na	na	na
	ROM	…	…	…	…	•		•	•	•	…

Reference Key:
• Yes
… Not available
na: Not applicable

Sources: Author based on PISA 2009; (4.1.) Van Bruggen 2010; and (4.2.) Eurydice 2004, except for first and last columns which are based on Van Bruggen 2010.

Table 8 (continued)

	4. *How* are external evaluations conducted? PROCESSES AND PROCEDURES								
	4.4. Do school-level evaluations include consultations with teachers, management staff, parents or students?					4.5. What is the structure of the external evaluation process?			
	Internal evaluations				External evaluations	Pre-evaluation	Evaluation	Follow-up	
	Teachers	*Management team*	*School board*	*Parents and/or Students*				*In all schools*	*In some schools*
FIN	na	na	na	na	na	na	na	na	na
NLD	•	•	•		•	•	•		•
EST	•	•			•	•	•		•
DEU	•	•	•		...		•		
POL	•	•	•	•		•	...	•	
SVN	•	•	•	•		•	...	•	
DNK	•	•	•		...	...		...	...
IRL	•	•	•	•	•	•	•		•
HUN	•	•			...		...		
FRA	•	•	•			•	...		
CZE	•	•	•		•	•	•		•
SWE	•	•	•	•		•	•		•
SVK	•	•	•		•	•	•	•	
PRT		•	•			•	•		
ESP	•	•	•		•	•	•	Δ	Δ
LTU	•	•				•	...		
ITA	na	na	na	na	na	na	na	na	na
ROM	•	•			•	•	...		

Reference Key:
- • Yes
- Δ Varies by region
- (blue) Required
- (tan) Recommended
- ... Not available
- na: Not applicable

Sources: Author based on PISA 2009 and Eurydice 2004.

5. *With what* resources, expertise and capabilities are schools inspected and evaluated? **Capabilities.**

This section documents the capabilities for conducting internal and external evaluations in the selected systems, with emphasis on human resources and expertise and, to a lesser extent, information and financial resources. The selected systems' capabilities are summarized in .

- **Human resources and expertise**

External evaluation

Among the selected European systems, the qualifications required to become an external evaluator vary. In several countries, a teaching qualification and teaching experience are required, as a way to ensure that inspectors/evaluators are knowledgeable of the education sector. In other countries, teaching experience is also required[16] but any relevant education degree is accepted, including, but not limited to, a teaching qualification. In addition, a few countries require that applicants take special courses in evaluation and/or pass a written exam. In some countries, teaching experience can be substituted for other forms of relevant experience in the education sector –for example, experience as a school head, manager, administrative staff, or student counselor. This is the case in the Netherlands, Slovenia, England, Scotland, Ireland, Sweden and Portugal. In other countries, teaching experience is always required, and management experience is either required or preferred –for example, in Poland, Latvia, Slovakia and Spain.

In almost all the systems for which there are data, new inspectors/evaluators participate in an induction or mentoring program, and professional development opportunities for all staff are also available. Inspectors/evaluators receive systematic feedback on their performance in all above-OECD systems, and in some equivalent- and below-OECD systems (e.g., Ireland, Czech Republic, Sweden and Slovakia) but not in others (e.g., Denmark and Spain).

Permanent contracts for inspectors/evaluators prevail in a majority of countries, high- and lower-performing ones. In some cases, evaluators also enjoy civil servant status. There are, however, some exceptions. In England, all inspectors/evaluators are hired under temporary contracts, that is, contracts that have an expiration date. Also, in a few countries (Scotland, Portugal and Romania) some inspectors are hired under permanent contracts and others are hired under temporary ones.

Among the selected systems, the total number of inspectors/evaluators ranges from 47 to 1,400, with the average being 340; and the number of schools evaluated every year per inspector/evaluator ranges from less than 2 to 15, with the average being 6.5.

Internal evaluation

All of these forms of support for internal evaluation activities are provided in Europe, in different combinations depending on the country. Among the selected systems, the most common forms of support are the provision of training on evaluation, the assignment of a resource person, the

---

[16] France is the only country in which teaching experience is not a requirement to become an external evaluator.

development and provision of guidelines and manuals on the processes and procedures of evaluation and, to a lesser extent, the development and provision of school evaluation frameworks. Other forms of support, such as online resources, sharing of best practices, exchange of experiences and provision of databases are less common, found in less than a third of the selected countries. Also, in some countries there are six or more different forms of support available to schools (e.g., in England, Scotland and Sweden, as well as Finland and Norway, in which internal evaluation takes a central role, given the absence of national external evaluations); in other countries, mostly lower-performing and/or former communist ones, there are just a few (two or less) forms of support (e.g., in Estonia, Slovenia, the Czech Republic, Hungary, France, Slovakia, Spain, Italy, Portugal and Spain); and then there is a third group of countries in which there is a moderate provision of support, with three to five different forms of support (e.g., in the Netherlands, Germany, Austria, Belgium, Ireland, Poland, Iceland and Lithuania). Research that sheds light on what forms of support are most effective (and cost-effective) could be very useful to inform policy decisions on how to build schools' capacity to conduct internal evaluations.

- **Information resources and capabilities**

In all selected systems for which there are data, external evaluations rely in part on national data and statistics. A few systems, mostly above-OECD performers, have developed or are developing electronic school files which integrate all the information of a school (test scores, student socioeconomic background, characteristics of teaching staff, results from classroom observations, notes from consultations with school stakeholders, administrative data, etc.) into a single electronic file that is easy to analyze and to integrate with other school files for conducting national and topic analyses.

- **Financial resources**

External evaluation

According to data collected by SICI, European systems spend as little as 0.2% of their total education budget in external evaluation (e.g., Estonia, Ireland, Czech Republic and Portugal) and as much as 1.63% (e.g., Sweden). Much more work is needed to document the costs of external evaluation, an issue of central relevance to policymakers.

Internal evaluation

Among European countries, only Finland, the Netherlands, Ireland and England provide financial resources to schools specifically to support internal evaluation activities (Eurydice 2004).

Table 9. Capabilities for inspection and evaluation, selected systems

| | 5. *With what* resources, expertise and capabilities are schools inspected and evaluated? **CAPABILITIES** — 5.1. Human resources | | | | | | | | | | | | | | |
| | 5.1.1. What qualifications are required to become an inspector/evaluator? | | | | | | | 5.1.2. What kind of training is provided to inspectors/evaluators? | | 5.1.3. Are inspectors evaluated regularly for their performance? | 5.1.5. What is the employment status of inspectors? | | 5.1.6. How many inspectors are there? | 5.1.7. How many schools are evaluated per inspector each year? | 5.1.8. How many inspectors participate in an evaluation? |
	Teaching qualification	*Any relevant degree*	*Special course in evaluation*	*Passing a written exam*	*Teaching experience*	*Experience in education sector*	*School management experience*	*New inspectors have an induction/ mentoring period*	*There are in-service professional development facilities for all staff*		*Civil servant/ Permanent employment contract*	*Temporary employment contract*			
FIN	na	na	na		na	na	na	na	na	na	na	na	na	na	na
NLD		•			•	•	•	•	•	•	•		181	…	1-5
EST	•				•		•	•	•	•	•		47	2.77	1-2
DEU	•				•		•	•	•	•	•		Δ	Δ	2-4
POL	•	•	•	•	•			…	…	…	•		…	…	…
SVN	•		•	•	•	•	•	…	…	…	•		…	…	…
DNK	Δ	Δ	Δ	Δ	Δ	Δ	Δ		•		•		…	…	2-5
IRL	•	•			•		•	•	•	•	•		133	6.39	1-6
HUN	…	…	…	…	…	…	…	…	…	…	…		…	…	…
FRA		•	•			•		…	…	…	•		…	…	…
CZE	•				•		•	•	•	•	•		268	14.93	1-9
SWE		•			•	•	•	•	•	•	•		290	6.90	2
SVK	•				•		•	•	•	•	•		200	…	2-5
PRT	•		•		•	•		…	…	…	•	•	206	1.39	3
ESP	•		•	•	•			•	•		•		1400	…	2-3
LTU	•				•		•	…	…	…	•		…	…	…
ITA	na	na	na		na	na	na	na	na	na	na	na	na	na	na
ROM	•	•			•		•	…	…	…	•	•	…	…	…

Reference Key:
- • Yes
- Δ Varies by region
- … Not available
- na: Not applicable

<u>Sources:</u> Author based on PISA 2009; (5.1) Eurydice 2004 and Van Bruggen 2010; (5.2) & (5.3) Van Bruggen 2010.

Table 9 (continued)

		5. *With what* resources, expertise and capabilities are schools inspected and evaluated? CAPABILITIES								5.2. Information resources		5.3. Financial resources
		5.1. Human resources								5.2. Information resources		5.3. Financial resources
		5.1.9. How does the State help internal evaluators build their capacities?								5.2.1. Do external evaluators have access to electronic school files, or are these under development?	5.2.2. Do external evaluations rely on national data and statistics?	*External evaluation budget (as a % of total education budget)*
		Training	*Evaluation framework*	*Guidelines and manuals on procedures*	*Databases*	*Best practices*	*Website*	*Exchange of experience/ networks*	*Resource person*			
	FIN	•	•		•	•	•	•	•	na	na	
	NLD		•			•			•	•	•	0.80
	EST			•					•	•	•	0.20
	DEU	•			•				•	•	•	...
	POL	•		•		•				...	...	...
	SVN	•							•	...	...	...
	DNK	...	...	...	...	...	...	...	...	•	•	...
	IRL	•	•	•			•		•		•	0.20
	HUN	•								...	...	...
	FRA			•	•					...	...	...
	CZE		•				•				•	0.22
	SWE		•	•	•		•	•	•		•	1.63
	SVK			•							•	0.36
	PRT		•								•	0.24
	ESP	•	•							•	•	...
	LTU	•			•			•		...	...	...
	ITA								•	na	na	na
	ROM	•		•						...	...	...

Reference Key:
- • Yes
- Δ Varies by region
- ... Not available
- na: Not applicable

Sources: Author based on PISA 2009; (5.1.9.) Eurydice 2004; and (5.2) & (5.3.) Van Bruggen 2010.

6. *For what* are the results of inspection and evaluation used? **Uses and consequences.**

The uses and consequences of inspection and evaluation, which ideally should be consistent with the purpose of inspection and evaluation, are documents in Table 10 and discussed below.

- **Are evaluation results disseminated publicly?**

Among the selected systems, only five of eighteen countries make the individual school reports of external evaluation publicly available to everyone: the Netherlands, Ireland, Czech Republic, Sweden and Portugal. Although a minority, there has been a trend toward increasing dissemination of the results of external evaluation, and it is likely that more countries will join this group in years to come. In contrast, almost all systems publish a national report about the quality of education based on the individual-school results of external evaluation. The results of internal evaluation are rarely used for such a purpose; only Spain and Romania publish national reports based on the results of schools' internal evaluation. Finally, several countries identify and disseminate best practices as a result of external evaluation, a habit more common among higher-performing systems than among lower-performing ones.

- **What are the consequences of unsatisfactory evaluation for an individual school?**

In the large majority of selected systems, external evaluation leads to recommendations or instructions for improvement. The only exceptions are Hungary and Spain. In contrast, a small minority (Poland, Sweden and Romania) require schools to draft a school improvement plan based on the results of external evaluation. That said, many of the selected countries require schools to draft a school plan that includes strategies for improvement, even if such a plan need not be based on the results of external evaluation. This is the case in, for example, the Czech Republic, Denmark, Estonia, Spain, Hungary, Ireland, the Netherlands, Portugal and Slovak Republic.

Disciplinary action as a result of external evaluation (e.g. withholding or withdrawing funds, intervention or closure of a school) also occurs in a minority of the selected systems (Slovenia, Czech Republic, Slovak Republic and Lithuania). In countries where schools enjoy relatively high autonomy, such as the Netherlands or Ireland, disciplinary actions are certainly applied if a school fails to comply with educational laws or regulations – but the framework used for external evaluation of schools' quality is viewed as a guide for improvement and not as a legal document, as it is assumed that schools should be given relatively high autonomy to define quality in their own terms.

- **What are the consequences of highly satisfactory evaluation for an individual school?**

While it is possible that rewarding is as powerful as sanctioning low quality to promote quality improvements, this issue deserves further exploration in the empirical research. Unfortunately, previous studies have not systematically documented the consequences for schools of obtaining highly satisfactory results in an evaluation.

Table 10. Uses and consequences of evaluation, selected systems

		6. *For what* are the results of inspection and evaluation used? USES AND CONSEQUENCES									
		What are the consequences of external evaluation?						What are the consequences of compulsory internal evaluation?			
		School-level				System-level		School-level		System-level	
		Recommendations/ instructions for improvement	Requirement to draft a plan for improvement	Publication of individual school report	Disciplinary action	Dissemination of best practices	Disseminaton of national report	Management decisions	Reporting to higher authorities	Informing external evaluation	Evaluation of national system
	FIN	…	…	…	…	…	…	na	na	na	na
	NLD	•	•			•	•	•	•	•	
	EST	•				•	•	•			
	DEU	•					•	•		•	
	POL	•	•			…	•	•			
	SVN	•			•	…	•	•		•	
	DNK	…	…	…	…	•	•	na	na	na	na
	IRL	•	•			•	•	na	na	na	na
	HUN					…		na	na	na	na
	FRA	•				…		•		•	
	CZE	•		•	•		•	•		•	
	SWE	•	•	•		•	•	•	•	•	
	SVK	•			•		•	•			
	PRT	•		•			•	na	na	na	na
	ESP					•	•	•			•
	LTU	•			•	…		•	•	•	
	ITA	na	na	na	na	na	na	•			
	ROM	•	•			…	•	•		•	•

Reference Key:
- • Yes
… Not available
na: Not applicable

Sources: Author based on PISA 2009, Eurydice 2004 and Van Bruggen 2010.

V. Useful Tools and Materials

This section includes links to useful materials such as school evaluation frameworks, information about inspection and evaluation for schools and parents, inspection-related instructions for schools, individual school quality reports, the specific Codes of Ethics that apply to inspectors, and consultation processes for policy decisions. The section builds and expands on a list of useful resources published in The World Bank (2010a).

School evaluation frameworks

http://www.education.ie/servlet/blobservlet/insp_wse_pprimary_guide.pdf A guide on Ireland's Whole School Evaluation processes, procedures and framework for post-primary schools.

http://www.see-educoop.net/education_in/pdf/workshop/reconstructing_ml_org/pdf/netherlands_reconstructing_ml_org.pdf The 2002 Dutch framework for inspection/evaluation of secondary schools. This is not the most updated version, but has been included here because it is in English.

http://www.onderwijsinspectie.nl/site/binaries/content/assets/Actueel_publicaties/2010/0210+analyse+waarderingen+opbrengsten.pdf A document published by the Dutch Inspectorate of Education in February of 2010, explaining the methodology used to inspect and evaluate schools, including the indicators used, how a school's quality is measured against those indicators, and how assessments of a school's quality take into account special circumstances and context, such as a student population with special education needs. This document is in Dutch.

http://www.hmie.gov.uk/documents/publication/hgiosjte3.html *How Good is Our School? The Journey to Excellence* is Scotland's framework for inspection/evaluation. It includes the school quality indicators as well as the 6-point assessment scale used by Her Majesty's inspectors of education. In addition, this framework must be used by schools themselves to produce the self-evaluation report.

http://www.ero.govt.nz/Review-Process/Frameworks-and-Evaluation-Indicators-for-ERO-Reviews A link to the framework and indicators used by New Zealand's Education Review Office to evaluate school quality.

Information about inspection and evaluation for schools and parents

http://www.education.ie/servlet/blobservlet/info_note_pp_bom_wse.pdf?language=EN Summary note on Ireland's post-primary Whole School Evaluations, aimed at schools' Boards of Management.

http://www.education.ie/servlet/blobservlet/info_note_pp_parents_assoc_wse.pdf?language=EN Summary note on Ireland's post-primary Whole School Evaluations, aimed at parents.

http://www.ero.govt.nz/Review-Process Information for parents, schools and early childhood service providers about New Zealand's external evaluation process, such as the review cycle/periodicity, and an explanation of all the different steps involved in the evaluation process, from notifying the school when it will be reviewed to publishing the evaluation report on the ERO's website.

http://www.ero.govt.nz/Review-Process/Criteria-for-Timing-Decisions Detailed information about the timing of external evaluation in New Zealand.

Post-evaluation questionnaires and complaints forms

http://www.surveymonkey.com/s/CJN659J A questionnaire designed by New Zealand's Education Review Office to collect feedback from schools about their experience of the review process.

http://www.education.ie/servlet/blobservlet/des_insp_rpt_sch_request_review_form.pdf Form that enables schools in Ireland to raise concerns about the external evaluation that was most recently conducted.

School report cards and other evaluation reports

http://www.ero.govt.nz/National-Reports/(year)/2010 Links to several thematic reports published by New Zealand's Education Review Office, as well as "best practices" reports.

http://www.ero.govt.nz/ero/publishing.nsf/Content/Search New Zealand ERO's search tool to find specific school reports. Note also that a search tool is available upfront on the ERO's home page.

http://www.hmie.gov.uk/documents/inspection/Bo'ness%20Public%20School.pdf Example of a school review report published by Scotland's HMIE in 2006, also available on HTML: http://www.hmie.gov.uk/documents/inspection/Bo'ness%20Public%20School.html

http://www.hmie.gov.uk/documents/inspection/CanberraPS8530629.pdf Example of a school review report published by Scotland's HMIE in 2009, also available on HTML: http://www.hmie.gov.uk/documents/inspection/CanberraPS8530629.html

http://www.education.ie/home/home.jsp?pcategory=32818&ecategory=36092&language=EN Ireland's search tool to find specific school reports, including those from Whole School Evaluations as well as Subject Evaluations. Note also that a search tool is available upfront on the ERO's home page.

http://www.education.ie/servlet/blobservlet/report1_71750U.pdf?language=EN Example of a Whole School Evaluation report published by Ireland's Inspectorate in 2006, also available on HTML: http://www.education.ie/servlet/blobservlet/report1_71750U.htm?language=EN

Consultations

http://www.minedu.govt.nz/ The home page of New Zealand's Ministry of Education. Note that within the home page, under "The Ministry", a specific link is included to "Consultation", which includes information about ongoing and recently completed consultation processes.

http://www.scotland.gov.uk/Consultations/Current All consultations in Scotland, including those from the education sector and other sectors as well, are publicized on this page. The page lists current, forthcoming and closed consultations, and allows users to search for archived consultations. For each consultation, the page lists the period of consultation, the contact information, and a document that sets the stage for discussions and on which feedback is sought. For consultations that have been completed, it also provides an analysis of consultation responses. The "About consultations" link, http://www.scotland.gov.uk/Consultations/About, highlights the importance of consultations in policymaking, summarizes the rules and procedures that apply to consultation and explains how responses and feedback received from consultations may be used. The Government of Scotland has

an email alert system for consultations, SEConsult. It allows individuals and organizations to register and receive a weekly email containing details of new consultations (including web links) and those about to end (http://www.scotland.gov.uk/Consultations/seConsult).

http://www.ero.govt.nz/Review-Process/Frameworks-and-Evaluation-Indicators-for-ERO-Reviews Includes a link to the *Feedback Form for School Framework and Evaluation Indicators*, designed to collect inputs from schools in order to inform the process of revising the national framework used for external evaluation in New Zealand.

Code of Ethics for inspectors

http://www.ero.govt.nz/Review-Process/Expectations-of-Review-Officers Includes a link to the Code of Conduct that applies to staff of New Zealand's Education Review Office.

http://www.education.ie/servlet/blobservlet/inspector_code_practice.pdf The Professional Code of Practice that applies to inspectors in Ireland.

Appendix 1. Quality areas and indicators proposed by Anne Deketelaere (1999), Johann Van Bruggen (2010) and Eurydice (2004)

	Dekeletelaere		Van Bruggen	Eurydice	
	Quality Areas or Dimensions	**Quality Indicators**	**Quality Areas or Dimensions**	**Quality Areas or Dimensions**	**Quality Indicators**
OUTCOMES	**I. Output**	Attainment and achievement	**I. Academic results and general outcomes**	**I. Outcomes**	Results in tests or examinations (cognitive and social skills)
		Added value of the school			
		Ability to apply learning skills			Cognitive skills
		Ability to apply social and life skills,			Social skills
		Satisfaction of pupils/parents/next schools			
		Destination of school leavers	**II. Efficiency of the school**		Repetition rates
					Drop-out rates
					Completion rates
					Proportion of students who qualify for special education
					Pathway following compulsory education
		Attendance and participation			Pupil absenteeism
PROCESSES	**II. Teaching-learning processes**	Curriculum offered by the school	**III. Curriculum**	**II. Classroom teaching/ learning**	Curriculum/ Subject
		Content			
		Teaching	**IV. The process of teaching and learning**		Teaching methods
					Pupil assessment
					Use of homework
					Use of ICT
					Treatment of the differing needs of students
		Learning			Pupil participation

					Attitude of pupils toward learning
		Classroom management			Teacher/pupil interaction
		School/pedagogical climate	**V. Organization of the school to meet the differences among pupils**		Treatment of the differing needs of students
			VI. Care for students with special learning or educational needs	**III. Guidance and support of pupils**	Pupil counseling
					Remedial activity
					Pastoral care
					Personal development of pupils
					Development of pupils' social skills
					Measures to improve the behavior of pupils
					Treatment of pupils with special needs
					Fighting absenteeism
			VII. Pedagogical climate	**IV. School atmosphere**	Relationships between pupils, teachers, and pupils/teachers
	III. Management	Leadership	**VIII. Leadership and management**	**V. Leadership**	Action of the head/ management team
					Pursuit and achievement of common aims
					Compliance with internal rules
					Improvements since last inspection
					Coordination, internal communication and/or participatory processes
					Conflict management
					Common school values
		Vision building		**VI. General/ educational policy of the school**	Provision of lessons
					Definition of aims
					Homework policy
					School plan
					School rules
					Annual report
		Organization		**VII. Functioning of the bodies/ organization of the**	School board/ council
					Teachers' council
					Class council

				school	Advisory bodies
					Pupil unions
				VIII. School time management	Length of lessons
					Frequency of holidays
					School subject time allocation
					Time management during classes
				IX. Building management	Use of space and infrastructure
					State of classrooms/ buildings
				X. Management of financial and material resources	Use of operational resources
					Use/ state of science facilities
					Use/ state of ICT facilities
					Management of salaries/ budget
					Management of operational budget
					Management of overall budget
					Management of additional and special resources
				XI. Administrative procedures	Quality of the administrative management of: pupil enrollment; registers of attendance; pupil records; pupil achievement reports; staff records; staff evaluation files
					Complaints procedure
					Compliance with regulation for certifying pupil achievement
		Human resources management		**XII. Human resources management**	Professional development
					Support for new entrants
					Criteria for allocation of pupils and/or teachers
					Allocation of tasks unrelated to teaching
					Number of teachers and their qualifications
					Remuneration policy and incentives
					Policy when posts are vacant
		Communication in the school	**IX. The process of school development**	**XIII. Internal evaluation**	Internal evaluation of: teaching; the school; the teachers and school head
		Quality assurance			Measures introduces following internal

					evaluation
					Staff participation in internal evaluation
		Links with the wider community		**IV. Relations between the school and local community/ external relations**	Relations with: parents; other schools; local institutions; educational providers; the local community
					Making available documents on education provision, activities and/or pupil results
					Participation in local programs; activities in the local environment
					Participation in international programs
INPUTS	**IV. Context and Inputs**	Environment characteristics			
		Characteristics of students			
		Staff characteristics			
		Infrastructure			
		Balance of intake of students			

Sources: Deketelaere (1999) as cited in Van Bruggen (2000); Van Bruggen (2010); and Eurydice (2004).

References

Butterworth, Graham and Susan Butterworth. 1998. *Reforming Education: The New Zealand Experience 1984-1996*. Dunmore Press and the New Zealand Ministry of Education.

Chapman, Christopher. 2002. Ofsted and school improvement: teachers' perceptions of the inspection process in schools facing challenging circumstances. *School Leadership and Management* 22 (3): 257-272.

de Wolfe, Inge F. and Frans J.G. Janssens. 2007. Effects and side-effects of inspections and accountability in education: an overview of empirical studies. *Oxford Review of Education* 33 (3): 379-396.

Department of Education and Science, Inspectorate. 2003. *Looking at our school. An aid to self-evaluation in second-level schools*. Available from: http://ppds.ie/sdparchive/LAOS/Contents%20-%20Looking%20at%20our%20School%20%20Primary.htm.

Dresscher, W.T.G., F.H.J.G. Brekelmans and C.M.C Verlinden. 2007. *The Role of Inspection in the Dutch Educational System*. Available from: http://books.google.com/books?hl=en&lr=&id=8WUxYkwlH9EC&oi=fnd&pg=PA133&dq=school+inspection+Netherlands&ots=9o_qzy262S&sig=GadtPelmNvh2YfCK0HyIqNqVW3w

Education Council. 1999. *A portrayal of school quality. Proposals for the responsible publication of data on the quality of schools*. Available from: http://www.onderwijsraad.nl/upload/english/publications/quality.pdf.

Education Council. 2006. *Vigorous supervision of education*. Available from: http://www.onderwijsraad.nl/upload/english/publications/engelse_samenvatting_doortastend_onderwijstoezicht.pdf.

Education Review Office. June 2010 (Draft). *Evaluation Indicators for School Quality Reviews*. Available from: http://www.ero.govt.nz/Review-Process/Frameworks-and-Evaluation-Indicators-for-ERO-Reviews.

Ehren, Melanie, Frans Leeuw and Jaap Scheerens. 2005. On the Impact of the Dutch Educational Supervision Act. Analyzing Assumptions Concerning the Inspection of Primary Education. *American Journal of Evaluation* 26 (1): 60-76. Available from: http://ics.uda.ub.rug.nl/FILES/root/Articles/2005/LeeuwFL-Ontheimpact/LeeuwFL-Ontheimpact-2005.pdf.

Ehren, Melanie, and Adrie Visscher. 2006. Towards a theory on the impact of school inspections. *British Journal of Educational Studies* 54 (1): 51-72.

Ehren, Melanie, and Adrie Visscher. 2008. The relationships between school inspection, school characteristics and school improvement. *British Journal of Educational Studies* 56 (2): 205-227.

Elmore, Richard. 2003. *Knowing the right thing to do: School improvement and performance-based accountability*. Washington, DC: NGA Center for Best Practices. Available from:

http://www.ecs.org/html/offsite.asp?document=http%3A%2F%2Fwww%2Enga%2Eorg%2Fcda%2Ffiles%2F0803KNOWING%2Epdf+.

Eurydice. 2004. Evaluation of schools providing compulsory education in Europe. Available from: http://eacea.ec.europa.eu/portal/page/portal/Eurydice/showPresentation?pubid=042EN

Her Majesty's Inspectorate of Education. 2007. *How good is our school? The journey to excellence*. Part I and II available from: http://www.hmie.gov.uk/documents/publication/hgiosjte.pdf. Part III available from: http://www.hmie.gov.uk/documents/publication/hgiosjte3.pdf.

Jaap Scheerens, H.W.C. Gonnie van Amelsvoort and Carol Donoughue. 1999. Aspects of the organizational and political context of school evaluation in four European countries. *Studies in Educational Evaluation* 25: 79-108.

Janssens, Frans. 2007. *Supervising the Quality of Education*. Available from: http://books.google.com/books?hl=en&lr=&id=8WUxYkwlH9EC&oi=fnd&pg=PA133&dq=school+inspection+Netherlands&ots=9o_qzy262S&sig=GadtPelmNvh2YfCK0HyIqNqVW3w

Janssens, Frans, and Gonnie van Amelsvoort. 2008. School self-evaluations and school inspections in Europe: an exploratory study. *Studies in Educational Evaluation* 34: 15-23.

Mac Beath, John, Denis Meuret, Michael Schratz, and Lars Bo Jakobsen. 1999. *Evaluating quality in school education. A European pilot project*. European Commission, Brussels. Available from: http://www.mp.gov.rs/resursi/dokumenti/dok297-eng-Evaluating%20quality.pdf

OFSTED. 2009. *Criteria for making judgments about the quality of provision*.

Omona, Julius. 2003. Design that supports service quality: New Zealand's model. *Quality Assurance in Education* 11 (3): 157-171.

Sakura, Fubuki. 2007. School monitoring and quality assurance in the New Zealand school system. *Education Research Policy and Practice* 6: 229-234.

SICI. 2009. *Blue Book Series*. Individual country profiles available from: http://www.sici-inspectorates.org/web/guest/blueBook

Standaert, Roger. 2000. *Inspectorates of Education in Europe: A critical analysis*.

van Bruggen, Johan C. 2010 (Draft). *Inspectorates of Education in Europe; some comparative remarks about their tasks and work*. SICI Secretariat. Available from: available from: http://www.sici-inspectorates.org/web/guest/blueBook

van Bruggen, Johan C. 2002. "Evaluating and reporting educational performance and achievement. International trends, main themes and approaches". SICI Secretariat.

van Bruggen, Johan C. October 2001. "The Netherlands Inspectorate: Where we are now and what are the current and future challenges". Speech delivered at the "Inspecting in a New Age" International Meeting, on the occasion of the 200th anniversary of the Dutch Inspectorate of Education.

van Bruggen, Johan C. 1999. "The Netherlands Inspectorate of Education: Old but Young". Transcript of speech delivered at the 1998 National Conference of the Portuguese Inspectorate.

The World Bank. 2009. *Strengthening the quality assurance system for basic and secondary education: Comparison of education quality assurance systems and institutions in selected countries*. Washington, DC: World Bank.

The World Bank. 2010a. *Implementation of education quality assurance reforms: Lessons for Chile from selected OECD systems and local reforms in other sectors*. Washington, DC: World Bank.

The World Bank. 2010b. *Design of a new institutional arrangement for education quality assurance. Lessons for Chile from international and local experiences*. Washington, DC: World Bank.

The World Bank. 2010c. *Supervision of Primary and Secondary Education: A Policy Note for the Government of Poland*. Washington, DC: World Bank.