

Impact of school inspections on improvement of schools—describing assumptions on causal mechanisms in six European countries

M. C. M. Ehren · H. Altrichter · G. McNamara · J. O'Hara

Received: 31 January 2012 / Accepted: 17 December 2012 /
Published online: 4 January 2013
© Springer Science+Business Media New York 2012

Abstract School inspection is used by most European education systems as a major instrument for controlling and promoting the quality of schools. Surprisingly, there is little research knowledge about how school inspections drive the improvement of schools and which types of approaches are most effective and cause the least unintended consequences. The study presented in this paper uses interviews with inspection officials and a document analysis to reconstruct the “program theories” (i.e. the assumptions on causal mechanisms, linking school inspections to their intended outcomes of improved teaching and learning) of Inspectorates of Education in six European countries. The results section of the paper starts with a summary of the commonalities and differences of these six national inspection models with respect to standards and thresholds used, to types of feedback and reporting, and to the sanctions, rewards and interventions applied to motivate schools to improve. Next, the intermediate processes through which these inspection models are expected to promote good education (e.g. through actions of stakeholders) are explained. In the concluding section, these assumptions are critically discussed in the light of research knowledge.

M. C. M. Ehren (✉)
Institute of Education, 20 Bedford Way, WC1H 0AL London, UK
e-mail: m.ehren@ioe.ac.uk

H. Altrichter
Johannes Kepler Universität Linz, 4040 Linz-Auhof Österreich, Austria
e-mail: herbert.altrichter@jku.at

G. McNamara · J. O'Hara
Dublin City University, School of Education Studies, Dublin 9, Ireland

G. McNamara
e-mail: gerry.mcnamara@dcu.ie

J. O'Hara
e-mail: Joe.OHara@dcu.ie

Keywords School inspections · Accountability · External evaluation · Comparative research · School improvement · School effectiveness

1 Introduction

Almost all countries in Europe arrange for evaluation of their schools in order to improve the quality of education (Eurydice 2004, p. 1). The quest for quality has become even more focussed in the last decade as economic globalisation has increased the significance of quality education and international assessments of student performance provide measures for comparative appreciation of education results. Many education systems have attempted to modernize their governance by establishing some variety of an “*evidence-based governance regime*” which may be characterized by the following features (see Altrichter and Maag Merki 2010):

1. They *set expectations* for the performance of the education system and communicate them more clearly than before.
2. Evaluation and accountability are considered to be key issues in ensuring quality provision for all. Evaluation measures are to *produce evidence* as to whether or not expectations have been met by the practical operation of the system units.
3. “Evidence” will stimulate and orientate *system development*. Actors at all levels of the system—education politicians, administrators, school leaders, teachers, students, parents, members of school; boards, etc.—will use evaluation information to make more rational choices in developing their contribution to the education system and in improving their performance.

In Europe, “evidence-based governance systems” are built on two dominant arrangements which often exist side by side. The first one includes *performance standards* (which set expectations) and tests of student performance (which produce evidence of system performance; see Altrichter and Maag Merki 2010). Student achievement results on national standardized tests are aggregated to evaluate the performance of schools and in some cases to publish league tables of schools.

The second arrangement for educational accountability is *school inspections*. Inspectorates of Education set expectations through their inspection standards and procedures. They assess the quality of education by using existing data (e.g. statistics, data on student performance) and by collecting additional information (e.g. by interviews with stakeholders, classroom observation). As a consequence, they produce reports which hold schools accountable for a broad range of goals related to student achievement, teaching, organization and leadership.

The specific approaches towards the inspection of schools differ greatly across education systems. An OECD commissioned literature review (Faubert 2009) for example shows differences in the scope, methods, standards, data and instruments of inspection between countries. The scope of school inspections varies according to the level of power and influence granted to each inspectorate and the inspection style used to evaluate schools. Hughes et al. (1997) suggest an inspection continuum, ranging from ‘non-punitive’ inspections based on peer review controlled by the profession to ‘punitive’ regimes characterized by managerial approaches subject to

direct central government control. McGarvey and Stoker (1999) describe how regulators, such as Inspectorates of Education, may range from collegial and emancipatory (in which self-responsibility, self-evaluation and self-regulation is emphasized) to a bureaucratic and technicist approach in which rules, procedures, accountability, compliance and sanctions are central. This distinction is often characterised as that between a policing style in which enforcement of rules is the main task of the inspectorate as opposed to a consultancy approach in which persuasion, advice and education are the main functions of the Inspectorate.

Inspection standards can vary from being based on indicators on teaching and learning, which are often inspired by school effectiveness research, to ensuring the conformity of schools with particular statutory requirements. Indicators used to judge teaching and learning for example include the pedagogical and didactical behaviour of teachers and the quality of the school curriculum which is often evaluated through observations of lessons and analysis of textbooks and lesson plans of the school. The conformity of schools to regulations typically involves checking the availability and use of procedures, policies and protocols concerning for example, admission policies or safety regulations and increasingly the satisfactory completion of school self-evaluation documents. The body which conducts school inspections may be located on different levels within the education system, ranging from the national centralized level to a more decentralized level of a province, region or municipality (Whitby 2010).

Although many of these distinctions seem to be mutually exclusive in a logical sense, there are also instances where inspectorates seem to combine different, sometimes even contradictory, approaches and emphases in the various aspects of their work. This may be due to the fact that inspections in some countries have been entrusted not just with evaluative tasks but also with other functions (e.g. personnel management). It is also related to the fact that the increased emphasis on school inspection in recent decades is closely connected to the apparently contradictory policy of making schools more autonomous and self-governing. High levels of school autonomy are counterbalanced in some countries by systematic evaluations of schools to assure the quality and effectiveness of school level decisions. Declining student achievement results, as measured in international surveys such as PISA and TIMSS, have also often spurred an increase in evaluation and control of schools even in supposedly decentralized education systems. Gustafsson and Myrberg (*in prep*) for example describe how the declining results in national evaluations and international comparative studies prompted the Swedish government to separate the tasks of steering and support of schools from those of evaluation and control. School inspections are now, as a result, largely concerned with stricter control of quality and ensuring compliance of schools with the law. Overall though it can still be argued that, even though Inspectorates of Education in Europe vary in their approaches and in their origins, they generally share a common purpose of improving teaching and learning (Faubert 2009).

A recent literature review by Klerks (*submitted*; see also Ehren and Visscher 2008) summarizes the effects of school inspections on behavioural change among teachers, school improvement and student achievement results. Her systematic study of peer-reviewed articles that were published after 2000 and include empirical research (preferably with a high score on the Maryland Scientific Methods Scale, including quasi-experimental research designs) shows plausible connections between

inspection and school improvement and behavioural change among teachers. In addition, Gray (cited in Visscher and Coe 2002, p. 2), Kogan and Maden (1999) and Ehren and Visscher (2008) describe how schools use school inspections, and the feedback provided during inspections, to implement improvements such as rules of conduct for students, strategies for raising examination results, changes in monitoring and assessment of students and changes in management styles and structures.

However, the overall results of inspection research are, at present, far from conclusive. Klerks' review shows small (positive and negative) causal effects of school inspections on student achievement results. Luginbuhl et al. (2009) found that test scores of Dutch primary students improved by 2 to 3 % of a standard deviation in the 2 years following an inspection visit. In contrast, Rosenthal (2004) reports a decrease in examination results in English secondary schools in the year of the inspection visit. He explains this result by arguing that the extensive preparation by schools for the visit may take time and energy away from the teaching and learning process.

These reported declines in student achievement results may point to unintended negative consequences of school inspections. Such negative consequences may occur when schools implement procedures and protocols that have no effect on primary processes in the school but just aim at receiving a positive inspection evaluation. In effect, this means a (usually unintended) undue emphasis on the elements that are assessed. Schools focus for example on programming a large number of lesson hours instead of trying to improve the quality of lessons offered (when the inspection rubric only measures the number of lesson hours) or they pursue short-term targets at the expense of legitimate long-term objectives. Or again schools may construct self-evaluation instruments to score positively on inspection indicators used for measuring quality assurance, instead of implementing such instruments to improve the quality of their education. These types of behaviours may negatively affect student achievement.

School inspections are widely used, and they are charged with a key role in the quest for quality. It is therefore of great importance to gain more knowledge about the in-school processes which may take place between the inspection and the ultimate goal, namely, the improvement of student performance (Husfeldt 2011, p. 1). We do not know how school inspections drive improvement of schools and which types of approaches are most effective and cause the least unintended consequences. The study presented in this paper intends to expand this knowledge base by describing the assumptions regarding causal mechanisms, linking school inspections to their intended outcomes of improved teaching and learning in the cases of six different European Inspectorates. Insight into these assumed causal mechanisms will help us identify those types of school inspections that are considered to be most effective across Europe and the mechanisms through which they are expected to affect school improvement. The following research questions will be answered in this paper:

1. What impact do different school inspection regimes in Europe intend to have on schools?
2. How do school inspections in Europe intend to contribute to the improvement of schools?

2 A framework for describing Inspectorates of Education in Europe

We start our analysis by describing a number of characteristics to differentiate between different types of school inspections. These characteristics help us identify the specific aspects of school inspections that are expected to be effective in promoting school improvement. Several authors propose categories that may be used to describe relevant characteristics of school inspections.

Van Bruggen (2010) for example distinguishes 51 characteristics to summarize the features of Inspectorates in 18 European countries. He describes the characteristics of inspection processes, the characteristics of the inspection report and the follow-up, the characteristics of the observation of teaching and learning as one element of the inspection. He analyses the characteristics of the system of inspection in a more general way, including specific types of inspections and inspection products such as the regime for failing schools or the publication of good practices. Other studies (such as Eurydice 2004 and Whitby 2010) use a similar taxonomy of characteristics, but add the consequences of these evaluations and the ways in which evaluation results are used (e.g. to monitor the education system and/or recommendations are provided to schools).

As this study focuses on the impact of school inspections, the characteristics of and differences between school inspections that have an impact on schools must be described in our model. The literature review of Klerks (submitted) provides a systematic summary of the inspection characteristics that have been found to be effective in empirical studies. Her review shows that no specific characteristic of school inspections in itself leads to improvement, but effects arise as a result of a complex interaction between inspection characteristics and students, teachers and the school management.

Generally, the type and frequency of school inspections (e.g. full/thematic school inspections), the standards and thresholds used to assess and provide feedback to schools during inspection visits and the sanctions, rewards and interventions used to motivate schools to improve (including the public report of the Inspectorate) seem to be the dominant aspects of school inspections affecting change in schools (Ehren and Visscher 2006; De Wolf and Janssens 2007). Standards are expected to influence actions in schools, particularly when schools face consequences for failing to meet these standards and thresholds. School officials can be expected to select the improvement actions that they perceive to have the highest yield, given their planning horizon, budget and appetite for risk. The feedback provided by school inspectors on the school's performance in relation to the standards, working in conjunction with sanctions and/or rewards, should support and motivate improvement. This set of relationships between inspection inputs and change in schools provides, in theory, an explanation of the way in which inspectorates hope to influence school improvement.

2.1 Types and frequency of inspection visits

Inspectorates of Education choose different methods to collect information on schools. These methods may be part of regular cycles of full inspections of all schools or of differentiated/proportional inspections of schools. According to Whitby (2010), in most countries the frequency of external inspections depends on an analysis of documents (including self-evaluation documentation) that the school submits to the external Inspectorate, and schools are then visited 'proportional to the need'. Within

these different models, a wide variety of evaluation methods are used to collect information, including desk research, visits to schools, interviews, questionnaires, classroom observation and analysis of documents produced by the school such as the results of self-evaluations.

2.2 Standards and thresholds

Standards present the details of what is expected of schools; they create boundaries or domains for attention with respect to educational quality. The standards and criteria may, according to Eurydice (2004), include educational responsibilities such as the teaching/learning of skills and knowledge, teaching/learning about appropriate patterns of social behaviour and personal development, and other administrative responsibilities such as the management of resources, external relations and partnerships.

According to Whitby (2010), standards may emphasize input expectations (such as opportunity to learn, class size, teacher training, etc.) and/or output standards as measured by the performance of schools. In addition, the “quality frameworks” which have been recently produced by many German *lander* education systems to guide school inspection also distinguish context and process standards (Kotthoff and Böttcher 2010). Context standards are used to take into account the school-specific location, history, identity and student population, whereas process standards emphasize compliance with legislation or principles and practices of good education.

The type of standards and thresholds developed will invariably influence which improvement actions schools take and how effective these are in improving student achievement. From this perspective, Scheerens et al. (2005) describe the process indicators in the (Dutch) inspection framework and use educational effectiveness research to evaluate their likely positive association with learning outcomes. Indicators on the quality of learning and instruction (such as learning time, and clear and structured teaching) have received empirical support in the literature as being associated with relatively high performance.

The thresholds used to identify schools that are to be judged as failing, overall, to meet the standards should motivate schools to alter their behaviour. Hanushek and Raymond (2002) for example describe how schools that have scores close to a performance target change their behaviour more than schools further away from that target.

2.3 Sanctions, rewards and interventions

Schools that are evaluated as failing often face consequences, such as sanctions or interventions (Van Bruggen 2010). Sanctions may include fines or closure of a school. School inspectors may intervene in such schools by means of increased monitoring of specific improvement plans which the schools are required to implement to address their weak points. Consequences of school inspections can also include rewards for high-performing schools which may receive awards or financial bonuses.

Several studies suggest that sanctions and rewards have a positive effect on educational quality in schools. The operating assumption in these studies is that schools work harder to perform well when something valuable is to be gained or lost; information and feedback alone are seen as insufficient to motivate schools to perform to high standards (Malen 1999; Elmore and Fuhrman 2001; Nichols et al.

2006). Heubert and Hauser (1999) found a significant relationship between the level of incentives for schools and students and the extent to which the curriculum and teaching in schools improves. Responses to inspection tend to be most focused and effective where funding is at stake or exposure is higher, according to Matthews and Sammons (2004). Formal sanctions like forced reconstitution of consistently low performing schools were more likely to promote responses than just “informal embarrassment” arising from the grading of schools and the reporting of results publicly. In summary, the desired responses by schools are likely to be governed by, on the one hand, a greater awareness of the importance of the standards and on the other, sanctions and rewards which force schools to comply with the standards.

Importantly, however, high stakes (test-based) accountability systems have also been shown to produce harmful consequences (Heubert and Hauser 1999; Koretz 2003; Stecher 2002). Sanctions and rewards may discourage desirable behaviour or may stimulate unintended and undesirable behaviour. Kerr (1975) describes how organisms seek information concerning what activities are rewarded, and then seek to do (or at least pretend to do) those things, often to the virtual exclusion of activities not rewarded. The extent to which this occurs depends, according to Kerr (1975), on the perceived attractiveness of the rewards offered. According to Elmore and Fuhrman (2001), schools operating under severe sanctions such as reconstitution and probation do not appear to be making fundamental changes in their core processes. Instead, they may place considerable emphasis entirely on the elements in the organization of education that are assessed as part of school inspections. These quick fix solutions often lead to more rapid improvement on the measures of the Inspectorate than to genuine long-term improvements. Some of these schools may incorporate structural changes but few appear to be making extensive or deep efforts to rethink their instructional programs.

2.4 Feedback during inspection visits

Inspectors normally assess schools with respect to standards, usually defined within a wider “quality framework”, and give feedback on the strong and weak points of the performance of schools based on these standards. Some Inspectorates also give schools advice on how to improve, while others are required to limit themselves to their evaluative role and to refrain from remedial action.

In this research, feedback is expected to emerge as an important inspection characteristic for improvement of schools. Ehren and Visscher (2008) for example found in their case study that all schools use the feedback received from the school inspectors to improve their functioning, and after 6 months, all schools were still carrying out improvement plans. Theories on schools as learning organisations and school improvement support the role of performance feedback in effecting change. However, not all types of feedback may be useful for schools (see Hattie and Timperley 2007). Research on the use of feedback on student performance shows that many schools have difficulties in taking appropriate action in response to feedback data (see Altrichter 2010). Only feedback that is relevant, understandable, clear, constructive, specific, accurate and useful will lead to actual improvement (Doolaard and Karstanje 2001; Brimblecombe et al. 1996; Ilgen et al. 1979; Kluger and DeNisi 1996).

2.5 Public reporting

Inspectorates of Education generally publish inspection reports in which the functioning of the school with respect to the inspection standards is described and areas of improvement are identified. Public reporting by Inspectorates of Education may also include lists of failing schools or tables of schools summarizing their performance in relation to inspection standards. These lists, tables and reports are published to inform stakeholders about the quality of the school. Public reporting is expected to promote a ‘market mechanism’ where improvement of schools is motivated through informed school choice and the voice of parents (Ehren et al. 2005).

Research on the impact of these public reports on school improvement particularly describes how schools respond to being ‘named and shamed’ in the media and how parents make only very limited use of the inspection reports in order to motivate schools to improve. Dronkers and Veenstra (2001) and Karsten and Visscher (2001) for example show that parents are primarily concerned about the atmosphere, pedagogical climate, working methods, safety, clarity of regulations, waiting lists for special education, reputation of the school and about decisions concerning the promotion of pupils to the next class. Parents mostly suggest possible improvements on conditional matters such as timetables. In general, they do not interfere in matters relating to educational quality, nor are their school choices based on educational quality. This research looks again at the relationship between school improvement and the reaction of the stakeholders, if any, to the outcomes of inspection since most inspection regimes perceive public and parental pressure as significant driving forces of change.

3 Method

Our study includes a documentary analysis and interviews with inspection officials in six European countries to identify the causal mechanisms of how school inspections are supposed to lead to the improvement of schools. The descriptions present inspection models that were in place in these countries in 2009/2010.

3.1 Selection of European Inspectorates of Education

The Inspectorates of Education selected for this study are the Netherlands, England, Sweden, Ireland, Austria (Styria¹) and the Czech Republic. These Inspectorates differ in the characteristics described in the previous paragraphs and therefore represent a broad variety of types of school inspections. They vary in using a low stakes capacity-building inspection approach (e.g. Ireland), to test-based early warning inspections to control schools (e.g. the Netherlands), and range from very centralized national Inspectorates of Education (e.g. England) to inspection agencies that operate at the level of the provinces (Austria).

¹ Austria is represented in this study by the regional education authority of the province of Styria which uses some room for maneuver in the Austrian centralist education legislation for developing a specific type of “team inspection”.

3.2 Data collection and analysis²

In each country, an analysis of relevant documents (such as inspection frameworks, legislation and documents describing rationales for inspection methods) was used to reconstruct assumptions on the causal mechanisms underlying intended effects of school inspections (Table 1). Additional interviews with inspection officials and policy makers were scheduled in each country to validate and clarify the reconstructed assumptions. For this purpose, inspection officials and policy makers were asked to indicate whether the assumptions in their reconstructed program theories provide an accurate description of the current inspection methods, their intended effects and of the intermediate mechanisms explaining these effects. Their comments and clarifications were used to revise the assumptions. The following table provides an overview of the data collection in each country (Table 1):

The assumptions were summarized in country-specific program theories. The “*program theory*” (see Ehren et al. 2005) served to describe and organize the intended effects of inspections in each country, the characteristics of inspection and the mechanisms through which effects are expected to occur. Only assumptions related to the characteristics of school inspections and expected effects as described in the previous section were taken into account. A policy scientific approach, consisting of the following steps, was used to reconstruct these assumptions (Leeuw 2003, p. 7):

1. Identify the social and behavioural mechanisms that are expected to solve the social, organizational or policy problem in question; search formal and informal documents for statements indicating the necessity of solving these problems, the goals of the proposed policy or program and how they are to be achieved. These latter statements refer to mechanisms (or “engines”) that drive the policies or programs and are believed to make them effective. Examples of this methodology in use in other fields are manifold. They include determinants of innovation diffusion, mechanisms underlying Prisoner’s Dilemma games, processes producing social capital, cognitive dissonance, different types of learning behaviour and many more. Statements having the following form are especially relevant for detecting these mechanisms:
 - It is evident that x will work.
 - In our opinion, the best way to address this problem is to . . .
 - The only way to solve this problem is to . . .
 - Our institution’s x years of experience tell us that . . .
2. Compile a survey of these statements and link the mechanisms to the goals of the program under review.
3. Reformulate the statements into conditional “if–then” propositions or propositions of a similar structure (e.g., “the more x , the less y ”).
4. Search for warrants that will identify disconnects in or among different propositions using argumentation analysis. This analysis, founded in part on Toulmin’s

² Data collection was carried out by a team of researchers, including (in addition to the authors of this paper) Peter Tymms, Karen Jones, Jan-Eric Gustafsson, Eva Myrberg, Gerry Conyngham, David Kemethofer and David Greger.

Table 1 Overview of data collection and data analysis

Country	Documents analysed	Inspection officials interviewed
The Netherlands	Documents outlining rationale and specifications for re-enactment of Dutch Supervision Act (e.g. whitepaper on supervision and regulation, 2005; ‘Governance letter’, 2005, policy document ‘trustworthy inspections’, 2006; evaluation of the Supervision Act in 2007; Smeets and Verkroost, 2011; De Wolf and Verkroost, 2010; Janssens and De Wolf, 2009;) newsletters of the Inspectorate of Education; minutes of discussions in parliament about the re-enactment of the Supervision Act, the final outline of the revisions of the Supervision Act and the memorandum of understanding, explaining the rationale for the re-enactment of the Supervision Act)	Interviews and/or e-mail correspondence with five employees of the Inspectorate Focus group discussion with four policymakers of the Department of Education
England	Education Act (2005), Education and Inspections Act (2006), Framework for School Inspection (Ofsted 2009a), Ofsted Inspects (Ofsted 2009b), Ofsted, Raising standards, improving lives (Ofsted 2009c), Evaluation schedule for schools (Ofsted 2010)	Meeting with Acting Ofsted’s Divisional Manager for Schools, and Ofsted’s Principal Officer, Knowledge Strategy Directorate Follow-up conference call with Ofsted’s Divisional Manager for Challenge and Analysis, Ofsted’s Divisional Manager for Schools, and Ofsted’s Principal Officer, Knowledge Strategy Directorate
Sweden	Swedish Official Report Series (SOU, 2007:11), Swedish Educational Act (Svensk Författningssamling (2010), Svensk Författningssamling (2009), “Letters of Regulation” (“Regleringsbrev”) (Skolinspektionen, 2009, 2010, 2011), Skolinspektionen (2011). Yearly report 2010	Interviews with Senior Inspection Director and undervisningsråd and written comments on drafts of two undervisningsråd (advisors on education within the Inspectorate)
Ireland	Key source documents describing LAOS and WSE-MLL	Three focus groups with school leaders
Austria/ Styria	Analysis of legal documents (Entwurf zur Änderung des §18 Bundes-Schulaufsichtsgesetz, 11. November 2010; Erläuterungen zum Entwurf zur Änderung des §18 Bundes-Schulaufsichtsgesetz, 11. November 2010) Analysis of documents outlining the rationale of Styrian team inspection (Teaminspektionen 2007/08 in Hauptschulen, Polytechnischen Schulen und Realschulen in der Steiermark; Beobachtungsformular für Unterrichtssituationen Sekundarbereich I; Schulinspektion–Besprechung mit Klassenelternvertreter/innen; Schulinspektion–Besprechung mit	Interviews with two district school inspectors and central official in the Ministry of Education

Table 1 (continued)

Country	Documents analysed	Inspection officials interviewed
Czech Republic	Klassensprecher/innen; Standort- und Perspektivengespräch mit dem Schulleiter; Schulentwicklungsplan; Brief Bezirksschulrat N.N. an die Hauptschule/PTS; LehrerInnengespräch; Zoller, 2005; 2008) White paper on education reform (2001), Education Act regulating tasks and functioning of Inspectorate (2004), documents and reports of Inspectorate, describing their work and outcomes.	Interviews with head of regional school inspectorate and the deputy chief school inspector

(1964) ‘The Use of Argument’, refers to a model for analysing chains of arguments and helps to reconstruct and “fill in” argumentations. A central concept is the warrant, which, according to Toulmin (1958) and Mason and Mitroff (1981), is the “because” part of an argument. A warrant says that B follows from A because of a (generally) accepted principle. For example, “the organization’s performance will not improve next year” follows from “the performance of this organization has not improved during the past 5 years” because of the principle that past performance is the best predictor of future performance. The “because” part of such an argument is often left implicit, with the consequence that warrants must be inferred by the person performing the analysis.

5. Reformulate these warrants in terms of conditional “if–then” (or similar) propositions and draw a chart of the (mostly causal) links.

The following example clarifies the first three steps by showing how two statements from documents can be used to reconstruct an assumption.

Example

- Statement 1: ‘In cases where the Inspectorate of Education identifies failures in schools, schools are monitored more intensively until problems are solved.’
- Statement 2: ‘Whenever a school is assessed as failing, the Inspectorate of Education monitors the school intensively. The goal of this intervention phase is to make sure educational quality in the school is improved to an acceptable level as soon as possible.’

These two statements lead to the following assumption:

‘If failing schools are inspected more often, then they will improve more quickly’.

3.3 Cross case analysis

In another step, we compared the program theories in the six countries to distinguish commonalities and differences in their assumptions on causal mechanisms of the

respective inspection systems. The assumptions that are similar in most program theories are considered to be the ‘conceptual core of inspection’, representing those potential causal mechanisms explaining how school inspections may lead to school improvement which all national inspection systems in our study seem to subscribe to. The ‘conceptual model’ of the processes and effects of school inspection resulting from this procedure will be empirically tested in another stage of this international project (which is not reported in this paper). The findings of this procedure are summarized in Fig. 1 and explained in the following section.

However, there are also differences between the national inspection systems which empirical research must attend to. The Appendix includes a summary of the major features and mechanisms identified in the country program theories. It functions as our account of how different national inspection approaches vary in inspection elements and assumptions on effect.

4 Intended effects of European Inspectorates of Education

The intended outcomes of the Inspectorates of Education in our study seem similar: all Inspectorates of Education aim for *good education* in individual schools and/or the education system as a whole. They want to pursue this aim through some specific combination of measures of accountability and control on one hand, and by stimulating development on the other. Inspectorates of Education, however, *differ* in the extent to which they specify these intended effects and how they define good education.

The Irish Inspectorate for example identifies very broad objectives, such as contributing to self-evaluation, to school development and to the improvement of the education system. The Swedish Inspectorate of Education aims to ensure the right of all students to a good education in a safe environment. It expects to improve inspected schools and to contribute to the improvement of the whole education system. Ofsted (the English Inspectorate of Education) emphasizes promoting

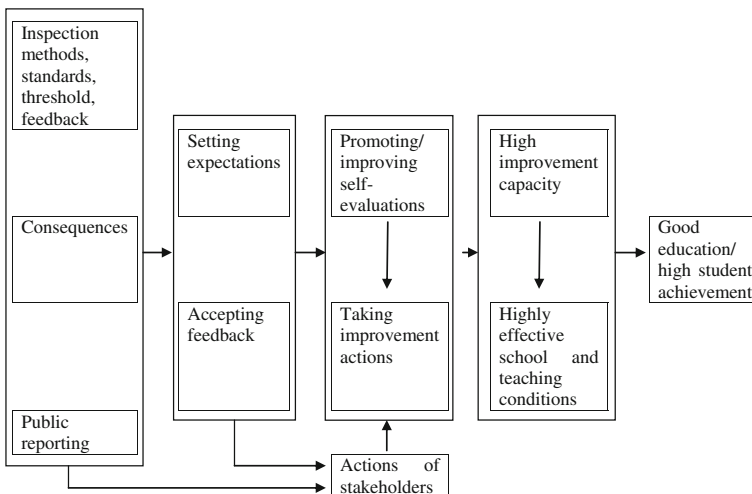


Fig. 1 Intended effects of school inspections—proposed conceptual model

improvement of schools as well as ensuring services are user-focused and provide value for money. School inspections should also encourage improvement of the education system as a whole. The goal of the Styrian school inspections in Austria is also very broadly described as supporting schools to constantly improve their quality and to promote the educational effectiveness and quality of the individual school, to ensure legal and administrative compliance and equivalence and comparability of various educational provisions within the system. The expected effects of the Dutch Inspectorate are more specifically described as improvement of schools towards good education, where the standards in the inspection framework are used to detail ‘good education’.

5 Inspection of schools

The Inspectorates of Education in our study use different models and methods for evaluation of schools to achieve these intended effects. This section first includes a summary of how they inspect schools, the standards and thresholds they use to classify schools as failing or performing well, the feedback they provide to schools, how they report on evaluation findings and the sanctions, rewards and interventions (consequences) used to motivate schools to improve. In the next section, we describe how these inspection characteristics are expected to promote intended effects.

5.1 Types of inspection

The Inspectorates of Education in our study use “*cyclical*” school inspections of every school, and *differentiated* inspections of particularly weak schools, to promote good education. Differentiated school inspections are generally implemented to increase the efficiency of school inspections by targeting inspection resources to potentially weak schools. Also, the increasing scale and professionalism of some schools allow for differentiated inspections when Inspectorates of Education can use results of self-evaluations to target inspection visits and potential areas for improvement.

Both types of school inspections are however also often used as parallel methods by the same Inspectorate of Education. The Netherlands, for example, uses early warning analyses to schedule inspection visits in potentially failing schools, while each school also receives at least one inspection visit every 4 years in which specific areas of concern or national targets are evaluated. Similarly, Ofsted conducts regular inspection visits to all schools, while 40 % of schools graded as satisfactory, and all schools graded as inadequate, receive monitoring inspections. In Sweden, regular supervision includes basic inspection visits to all schools once every 4 to 5 years, while schools that are evaluated as weak receive more elaborate ‘widened’ school inspections. The selection of schools for widened inspection is based on grades and results on national tests, observations made in previous inspections, complaints and questionnaire responses from students, parents and teachers.

The Irish, Czech and Styrian Inspectorate of Education on the other hand only carry out whole school evaluations of all schools. In Ireland, these inspections are generally scheduled once every 5 years. However, recent policy changes in Ireland have seen the introduction of ‘incidental inspections’. Unannounced and truncated in

terms of what they examine, there is a suggestion that this new category may be used to assist weak schools, although it is not clear if re-inspection will be part of this process. New policy directions in Sweden also imply a move from regular “cyclical inspections” to a more differentiated model in which schools are deliberately targeted for inspections. In Styria and the Czech Republic, the original schedule of full cyclical inspections proved too ambitious, and school inspectors currently use additional criteria and strategies for the selection of schools, such as, for example, choosing schools with young/new head teachers. In the Czech Republic, the number of school inspectors and inspection days per visit is decreased when small schools are inspected or when the preparation phase included extensive information.

In addition, the Inspectorates of Education in Sweden, England, Ireland, the Netherlands and the Czech Republic often also implement *thematic school inspections* of, for example, the teaching in specific subject areas or the use of ICT in schools. These inspections are added to the regular cycle of school inspections. Sweden even uses a different (research-based) inspection team and procedure to undertake these types of school inspections. Thematic school inspections are scheduled in a selection of schools, or information is additionally collected during regular inspection visits. Results are reported to individual schools, and a general report summarizing the main findings is also published.

Self-evaluations conducted by schools are an important part of all school inspection systems in our study. A focus on self-evaluation is often prompted by a tradition of autonomy and decentralization of educational policy. Inspectorates of Education align their inspections to the school’s priorities for improvement, and self-evaluation at the local level is expected to guarantee school development (e.g. in the Netherlands). The option of a greater emphasis on school autonomy and self-evaluation (as opposed to accountability and measuring student and teacher performance) is sometimes made when there is little agreement on the inspection criteria and methods of evaluation (Ireland). The results of these self-evaluations are in theory used by Inspectorates of Education as a relevant source of information in targeting inspection visits and/or forming a judgement on the quality of the school. In the Netherlands, self-evaluations of schools are for example part of the early warning analyses to identify potentially failing schools and schedule these schools for inspection visits. The role of school inspections is also to ensure that internal systems of evaluation and self-review are implemented effectively. Inspectorates in the Netherlands, England, Sweden and Ireland evaluate the quality of self-evaluation, and the feedback they provide on the product and process of self-evaluation should improve such systems. In these countries, schools are required to formally report on self-evaluation activities and are invited to make statements regarding areas that are covered in the inspection framework to assure alignment of self-evaluation results to inspection activities. In Ireland, the Netherlands and England schools are also required to survey stakeholders (parents and students) as part of their self-evaluation and gain insight in their opinions on the functioning and performance of the school.

In England, strong support is provided to schools for self-evaluation through accompanying self-evaluation forms and by detailed (external) back-up data, gathered over a period of time. Schools can use these data to support self-generated claims and satisfaction ratings. In practice, the actual importance accorded to self-evaluation varies greatly, depending on the skills and resources schools have to gather evidence and make judgements on their own functioning.

5.2 Standards, thresholds and feedback

The Inspectorates of Education in our study address three types of standards: *legal standards* (indicating the extent to which schools comply with regulations), standards relating to the *context and process quality* of education and finally standards defining the *performance or results* of schools. The Inspectorate of Education in Styria also includes the goals of the specific school (as documented in a school program) in the inspection framework and evaluation.

School inspections on legal standards for example check the extent to which schools provide equal access to education for all students (Sweden), or offer a minimum number of lesson hours (the Netherlands, Ireland). Standards related to the quality of education are often inspired by educational research on school improvement and/or school effectiveness. The indicators of school effectiveness as described by Scheerens (2009) are to varying degrees covered by inspection frameworks in each country. Indicators on opportunity to learn and learning time, achievement orientation, clear and structured teaching, challenging teaching approaches and orderly learning environment are part of the inspection frameworks in at least four countries of our study:

- *Opportunity to learn and learning time*: offering pupils a range of subjects and tasks that cover educational goals; classroom exercises corresponding with the content of the tests for monitoring performance. Opportunity to learn has different aspects: the quality of the curriculum (i.e. the textbooks and methods), the amount of time offered to reach the targets of the curriculum, and last but not least, the amount of students that lag behind so far that they do not even have a chance to reach the minimum targets. (Scheerens 2009, p. 33; van der Grift/ICALT study, 2007, p. 26³).
- *Achievement orientation*: clear focus on mastering basic concepts; high expectations of student achievement and record keeping of student achievement. Also, explicit or implicit standards are used as targets and as assessment norms. The core idea is the determination to obtain the best possible performance from pupils. Standards are set in such a way that pupils are challenged, but not demotivated by either too high or too low expectations (Scheerens 2009, pp. 46 and 63).
- *Clear and structured teaching*: structured, direct teaching and a clear goal-directed teaching approach (Scheerens 2009, p. 41).
- *Challenging teaching approach*: stimulating motivation through cognitive challenge.
- *Orderly learning environment*: emotionally supportive climate, order and discipline in classroom climate, achievement pressure, mastery and performance orientation (Scheerens 2009, p. 45).

No Inspectorate of Education in our study includes teacher characteristics in their framework of inspection standards, such as whether teachers teach ‘learning to learn strategies’ or how they group their students in its inspection rubrics. The quality of the school’s management is explicitly evaluated in five of the six countries. All

³ http://www.onderwijsinspectie.nl/binaries/content/assets/Actueel_publicaties/2009/ICALT.pdf

countries also include additional indicators in their frameworks which are not specifically related to school effectiveness or school improvement research. Ofsted for example evaluates students' well-being, Sweden assesses the education of head teachers and teachers, and the Czech Republic and Austria include an indicator on partnerships of the school with external stakeholders.

The Inspectorates of Education in England, the Netherlands and the Czech Republic also evaluate outcomes of schools. These outcomes may include cognitive outcomes on a number of subjects (generally mathematics, reading and writing), but sometimes other outcomes (such as cultural or social development of students) are also taken into account.

The Inspectorates of Education in the Netherlands, England and the Czech Republic use a threshold, including a minimum score on the inspection standards, to award grades to schools that often range from “failing” to “well developed”. Ofsted, for example, grades schools as “outstanding”, “good”, “satisfactory” or “inadequate”. The Swedish, Irish and Styrian Inspectorates of Education on the other hand do not grade school as failing or satisfactory; they only identify how schools are functioning on all the inspection standards. A Styrian school inspector explains the lack of threshold by stating that it is impossible to clearly identify weak schools; there are schools which are excellent in some areas but which are deficient in other areas.

Schools receive feedback on their strong and weak points with respect to the inspection standards and thresholds. Feedback sometimes also includes recommendations on how to improve, or examples of good practices in other schools. Feedback is generally provided to the head teacher and/or the entire school staff during meetings at the end of an inspection visit and in inspection reports that are drawn up after the visit. In the Netherlands, however, feedback is specifically directed to the school board. Inspectorates of Education generally do not provide feedback to individual teachers (an exception is post subject inspections in Ireland), even though teachers often ask for individual feedback by inspectors who have observed their classrooms. School inspectors however feel that they should not give feedback to individual teachers, but to the whole school; feedback to individual teachers is considered to be the task of headpersons and principals. School inspectors sometimes interpret requests of individual teachers for inspection feedback as an indication of a lack of a feedback culture in schools.

5.3 Reporting feedback to stakeholders

The Inspectorates of Education in the Netherlands, England and Ireland publish *inspection reports* in which the functioning of individual schools according to the inspection standards is described and areas of improvement are identified. Inspectorates of Education aim to make the reports easily accessible to stakeholders (particularly parents) by using plain language and by making the format of the reports as similar as possible in order to aid comparison. The Netherlands publish (in addition to reports on individual schools) lists of failing schools and summaries of the inspection assessments of all schools.

A number of countries also have arrangements in place which assure active notification of stakeholders about the outcomes of school inspections and/or the inspection report. As part of the publication process, school staff, management and

parents' associations in Ireland are informed in advance that the report will be published. In Styria, head teachers have the duty to "demonstrably inform school partners" (parents, students and teachers) and the school maintaining body (mostly communities) about the inspection results. Inspectors usually do not check whether or not the inspection report was on the agenda of the parent–teacher meeting, but, however, would do so if the problems persisted or the parents complained. The purpose of requiring schools to actively inform their stakeholders is to motivate schools to set out how the report will be used in the context of the school's ongoing program of planning, improvement and self-evaluation (Styria, Ireland). This provides a first step in encouraging schools to use the findings of the evaluation for school improvement.

Inspectorates of Education in Sweden, the Netherlands, England and Ireland also publish *more general reports*, for example, on the state of education in the country in a particular year or on the performance of (a subset of) schools in a particular area (e.g. teaching in a specific subject area). The general reports are often disseminated through media, conferences and seminars (Sweden, England) and are expected to highlight good practice that may inspire individual schools to improve. General reports are also published to inform local or national governments on policy for the improvement of all schools in a country (Sweden, England). Ofsted, for example, through the collection of a wealth of data in all areas of care and learning, can isolate in-depth evidence on virtually every aspect of learning in schools. This evidence is disseminated through a program of survey reports to influence policy and providers. Less formal types of feedback may also be used. In Ireland, for example, inspection teams provide feedback to management, subject departments and other stakeholders. Research indicates that this feedback may be more clear, useful and indeed robust than the published inspection reports, which are often couched in anodyne language (McNamara and O'Hara 2008).

5.4 Consequences

The consequences of school inspections include rewards, sanctions or interventions in schools resulting from a judgement on the quality of the school. The Inspectorates of Education in our study are not in a position to *sanction* schools, except for the Swedish Schools Inspectorate which may withdraw the license and funding of independent schools and may temporarily close down public schools. In the Netherlands, England and the Czech Republic, Inspectorates of Education however may advise the Minister of Education to impose sanctions on failing schools (e.g. to remove the school from the Register of Educational Facilities, or to impose administrative or financial sanctions).

Inspectorates of Education (except for the Irish Inspectorate and policy here may be changing) do *intervene* in schools that are judged to be failing. Ofsted, for example, categorizes schools as being in 'special measures' if the school is evaluated as inadequate and does not have the capacity to improve; or it gives a school 'a notice to improve' when it is performing less well than expected. Schools in special measures are required to work with the local authority, and if no improvement follows, they are under threat of closure by the Secretary of State. In the Netherlands, schools are monitored intensively when they are not performing up to the required

standard. School boards are expected to develop an improvement plan in which they address the weaknesses that have been identified in the inspection visit. The Dutch Inspectorate monitors the implementation of this plan. In Sweden, the schools are given a certain amount of time to amend identified shortcomings, and the implementation of improvements is inspected at a follow-up visit. In Styria and Ireland, all schools have to develop an improvement plan for the next time period, even if they are not considered to be failing. This plan serves as a target agreement between the principal and the inspector, and school inspectors check the implementation of these targets after 1 or 2 years. However, there are no consequences in place for schools that fail to implement these targets. In the Czech Republic, failing schools are monitored more frequently and are obliged to implement corrections that have been identified by the Inspectorate. Table 2 provides a summary of inspection characteristics of the Inspectorates of Education in our study.

6 Causal mechanisms to promote intended effects

Appendix provides a summary of the causal mechanisms and processes that are expected to occur when school inspections have an impact on schools. These processes are often part of a wider range of quality assurance measures in the education system. In Ireland, these measures include for example promotion of school self-evaluation and extensive support for school development planning; teacher in-

Table 2 Summary of inspection characteristics

	The Netherlands	England	Sweden	Ireland	Austria (Styria)	Czech Republic
Inspection methods						
Cyclical inspections of all school	Every 4 years	Every 5 years	Every 4–5 years	Every 5 years	Every 2–4 years	Every 3 years
Differentiated inspections	*	*	*			
Thematic school inspections	*	*	*	*		*
Standards						
Legal aspects	*	*	*	*	*	*
Context and process quality	*	*	*	*	*	*
Outcomes	*	*				*
Threshold for distinguishing failing schools	*	*				*
Consequences						
(Advising on) sanctions	*	*	*			*
Interventions	*	*	*	*	*	*
Reporting:						
General/thematic reports	*	*	*	*		
Reports on individual schools to the general public	*	*		*		

Asterisk indicates presence of characteristics/mechanisms in each country

career development and support in the context of curriculum change; school-designed assessment and reporting to parents; use of standardized assessment and state certificate examinations; and program evaluations focusing on aspects of curriculum provision and system evaluation through international surveys of attainment. In the Netherlands, the national government has implemented legislation to promote good governance of schools; schools are required to have separate mechanisms in place for the administration and internal supervision in the school. The Inspectorate of Education aligns its inspections to the activities of school boards and internal supervisors. In Styria, school inspectors are even responsible for implementing some of these additional quality assurance measures as they have additional tasks in leading, planning and coordinating schools. These mechanisms are designed to govern the chain of processes that link school inspections to their intended outcomes of promoting good education. In this paper, however, we focus only on assumptions about explicit and direct connections that can be made between the inspection characteristics and outcomes.

In summary, school inspections, their criteria and procedures in general, and the specific feedback given during inspection visits are expected to influence schools and their stakeholders to align their views/beliefs and expectations of what constitutes good education and good schools to the standards in the inspection framework. In particular, alignment is expected with respect to those standards the school failed to meet during the latest inspection visit.

Schools are expected to act on these views and expectations and to use the inspection feedback when conducting self-evaluations and when taking improvement actions. Likewise, stakeholders are expected to use the inspection standards and the inspection assessment of the school's functioning (as publicly reported) to take actions that will motivate the school to adapt their expectations and to improve.

Self-evaluations by schools are expected to build capacity to improve that will lead to more effective teaching and learning conditions. Likewise, improvement actions will (when successfully implemented) lead to more effective school and teaching conditions. In turn, this process should, logically, result in higher student achievement. The model in Fig. 1 below summarizes these mechanisms and links them to a proposed four-phase conceptual model of how inspection is designed to impact on school improvement.

6.1 Setting expectations and accepting feedback

The standards and thresholds are first of all supposed to promote the improvement of schools through the expectations they set. The criteria and descriptors set out in the inspection frameworks are designed to illustrate standards of performance and effectiveness expected of schools; they identify what is meant by a 'good school'. These expectations are often illustrated by national publications of examples of good practice that are expected to serve as a 'national knowledge base'. Schools should use these standards to guide their work. Acceptance and internalisation of them by key stakeholders (schools, teacher unions, local authorities or the wider public) is also considered key to their use in schools. The Inspectorate of Education in the Netherlands promotes acceptance of standards through an intensive process of collaboration and consultation with stakeholders in developing the standards. In Styria and Ireland, inspectors meet with schools before an inspection to outline the inspection process and standards. The procedures of

data collection, interpretation and feedback used during inspection can be an equally important method of communicating to schools what it means to monitor quality and what elements in-school quality management systems ought to include. Ofsted relies on good relationships between inspector and head teacher (enhanced through joint observations of lessons) to increase acceptance of standards and feedback in schools. A similar assumption can be found in Ireland where schools are explicitly invited to respond to the inspection report and feedback as a means to promote acceptance of the feedback in the report. In Styria, schools' acceptance of the feedback is made visible in an obligatory 'school development plan; a feedback conference with school staff is also held at the end of an inspection visit to promote acceptance of feedback.

The consequences schools face for not meeting the standards will also motivate them to align their conception of what constitutes good education to those standards. Schools face a strong impetus to incorporate these standards into their own definition of good education and to meet the inspection standards in order to avoid sanctions to receive rewards or simply the opportunity to manage improvement with as little outside interference as possible. The threat of bad practice triggering an inspection visit is also expected to keep schools 'on their toes' and accepting of the inspection standards. Ofsted also illustrates the power of incentives for schools to be judged as 'outstanding' on the standards in the inspection framework as this enables them to apply for benefits or special status, e.g. for 'academy status', 'teaching schools' or head teachers may apply to become 'national leaders of education'.

The standards also more explicitly guide improvement of schools when used by the Inspectorates of Education to evaluate schools and to provide feedback on those areas of work that need to be improved upon. The use of differentiated school inspections is considered to enhance this mechanism as weak schools are visited more frequently and are given more feedback on areas that need to be improved. As strong schools have a track record of delivering good education and are expected to have incorporated inspection standards into their daily operations, inspection resources and feedback in the Netherlands, England, Sweden and the Czech Republic are allocated to, or intensified for weak schools to make sure these schools use the standards to guide their work. The feedback includes recommendations on priorities for future action by the school with respect to the standards and offers a challenge and impetus to act where improvement is needed. Interventions in, and feedback to, failing schools are aimed at repairing any shortcomings as soon as possible. Inspectorates of Education expect schools in general to be willing and able to use the feedback to improve.

Inspectorates of Education in Ireland and England also expect to promote acceptance of recommendations and feedback by giving the school a chance to respond to them through a formal complaint process or as part of a professional dialogue between the inspector and the school staff during the inspection visit; such a dialogue takes place in some countries as part of a feedback conference in which the entire school staff participates (e.g. Styria).

6.2 Promoting/improving self-evaluations

Inspectorates of Education in the Netherlands, England, Ireland, the Czech Republic and Styria intend to promote self-evaluation in schools and to ensure that internal systems of evaluation and self-review are implemented effectively in schools. In

Ireland, school self-evaluation is seen as an important aspect of quality assurance in schools, and inspections areas are informed by the results of self-evaluations. Inspections in turn are expected to complement the school's quality assurance. The Czech Inspectorate of Education has recently been assigned a more active role in guiding and consulting schools in their internal evaluations which aims at developing an 'evaluation culture' in schools.

Schools are expected to implement self-evaluations because they know school inspectors will return, and they want to be prepared. They will use the inspection standards and expectations of adequate self-evaluation to conduct self-evaluations. In most countries, schools are also required by law to implement self-evaluations and to provide the Inspectorate of Education with results of self-evaluations. Also, Inspectorates of Education analyse the quality of school self-evaluations and provide feedback on how schools may improve their self-evaluations and internal quality assurance. Schools are expected to use this feedback to improve their internal systems of evaluation and self-review.

High-quality self-evaluation is considered to be a critical element in improvement of schools as schools identify and correct problems in the quality of their school. These internal quality assurance mechanisms, together with inspection, are seen as inseparable and integral parts of an improvement and accountability cycle.

6.3 Taking improvement actions

All Inspectorates of Education expect schools to act on the feedback given during inspection visits and in the inspection reports by developing and implementing improvement actions to meet the standards in the inspection framework.

Inspectorates of Education in Sweden, the Netherlands and the Czech Republic also oblige or motivate schools to implement the feedback by using some kind of follow-up to check on progress. In Styria and Ireland, schools have to write a 'school development plan' which identifies the areas for development (as indicated in the feedback) and which is checked by the Inspectorate during subsequent inspections. In the Netherlands, the Inspectorate of Education instructs the school board of failing schools to formulate a plan of approach aimed at improving quality. The Inspectorate analyses the plan and lays down performance agreements in an inspection plan. This plan specifies when the quality should be up to standard again and what (interim) results the school must attain. It also specifies the indicators the Inspectorate of Education will assess in (interim) inspection visits. The school board must commit to the inspection plan.

Mutual respect and trust between school inspectors and schools is considered to be a foundation for the development of improvement actions in schools. Also, the fact that results of school inspections are communicated to stakeholders creates a sense of obligation for schools to take improvement actions.

The Inspectorates of Education expect these improvement actions to be related to the standards in the inspection frameworks and the feedback given during the inspection visit and ultimately to lead to more effective teaching and learning in schools.

6.4 Actions of stakeholders

Stakeholders (such as parents, local policymakers or school boards) are expected to use the standards and the feedback of the Inspectorate of Education (as described in

the public report) to voice their opinion to the school, to choose a school or even to move students from a poor school. Parents for example are assumed to use the inspection reports to pressure schools to improve by questioning the school board and/or principal about improvements. Ofsted and the Dutch Inspectorate intend parents to use the information in the inspection report to inform their decisions about choosing a school for their child. Parents may also use the report to vote with their feet. Both types of parental action are expected to lead to the growth of strong schools and to a decline in student numbers in weak schools. These ‘market mechanisms’ should motivate schools to improve as school admissions can have a big effect on the job security of school staff including potential loss of jobs, or enforced redeployment to other schools.

Also, local authorities or school boards are expected to use the reports to stimulate their schools to improve. In the Netherlands and in Styria, the Inspectorate explicitly communicates the inspection findings to school boards, for example through a ‘Schulforum’ in which the report is discussed with the governing body of the school.

6.5 High improvement capacity

As a result of school inspections promoting self-evaluation in schools, the school’s capacity to improve is expected to be enhanced. The internal quality assurance systems that schools put in place should enable them to internally evaluate their weaknesses and build their capacity in improving these weaknesses. England and Styria also describe how communication between head teachers and principals during inspection visits may enhance the school’s capacity to improve. The head will often accompany the inspectors as they observe lessons. A dialogue will occur between the head and the inspector in order to gauge whether the two of them make the same judgments as each other. This can amount to mentoring the head by explaining how, for example, the head may use the inspection classroom observation scheme to supervise and observe teachers.

The Inspectorates of Education in our study do not give a very detailed description of what they consider to be capacity building in schools, but in general references are made to the capacity of schools to self-evaluate or to leadership in the school. Ofsted includes some examples and descriptions on their website such as ‘leadership which establishes a clear direction for change, takes tough decisions, sets the right pace for sustained progress and earns the commitment of staff’ or ‘an ethos that expects success and promotes ambition, combined with leadership which sets out what “outstanding” looks like’.

6.6 Highly effective schools and teaching conditions

To improve performance relative to the standards in the inspection frameworks, schools are expected to increase the effectiveness of their school and teaching conditions. Most inspection frameworks identify effective practices on the school level; some inspection frameworks also identify effective practices on the instructional level but aggregate these to evaluate the school as a whole. *Effective school level conditions* are presumed to include school processes and conditions related to school organization and management, including educational leadership, a productive

climate and culture and achievement-oriented school policy. These conditions are expected to contribute to and facilitate effective teaching and instruction and as a result lead to higher student achievement. *Teaching/instruction conditions* include what a teacher does to create effective learning environments and to boost learning (Scheerens 2009). The Dutch Inspectorate of Education for example includes indicators stating that ‘the didactical and pedagogical behaviour of teachers meets the basic requirements’. Ofsted evaluates indicators concerning ‘the quality of teaching and learning’. The indicators in the inspection rubrics are often less detailed when describing teaching practices; professional competences of teachers and subject-specific quality of instruction are generally not part of inspection indicators, although the latter is the focus of a great deal of inspection efforts in the case of Ireland, in the form of what are referred to as ‘subject inspections’. Table 3 provides a summary of assumptions on mechanisms and processes of school inspections in the various countries in our study.

7 Conclusion and discussion

The study presented in this paper intends to expand our knowledge base on the intended effects of school inspections across Europe and the mechanisms through which these effects are expected to occur. We used interviews with inspection officials and a document analysis to reconstruct the program theories and summarize the working assumptions of the Inspectorates of Education in six European countries. The commonalities and differences in these assumptions were presented in order to learn about the expected impact of inspection and particularly the characteristics of school inspections that are expected to be effective. We will conclude by discussing

Table 3 Summary of assumptions on mechanisms and outcomes of school inspections

	The Netherlands	England	Sweden	Ireland	Austria (Styria)	Czech Republic
Mechanisms						
Setting expectations for schools	*	*	*	*	*	*
Setting expectations for stakeholders	*	*		*		
Accepting feedback about improvement of inspection standards	*	*	*	*	*	*
Promoting self-evaluations	*	*		*	*	*
Taking improvement actions	*	*	*	*	*	*
Actions of stakeholders	*	*		*	*	
Outcomes						
High improvement capacity	*	*		*	*	*
Highly effective schools and teaching conditions	*	*	*		*	*
Improved student achievement	*	*	*			*

Asterisk indicates presence of characteristics/mechanisms in each country

these assumptions in the light of existing research on school inspections to identify which of them are most valid.

7.1 Intended effects of the program theories of European Inspectorates of Education

The intended effects across all European Inspectorates in our study are described under a common denominator of ‘aiming for good education’. The specific outline of what constitutes ‘good education’ however varies from country to country, ranging from equity-related perspectives of providing equal opportunities to students, to indicators on the quality of teaching and learning inspired by school effectiveness research to administrative/legislative input-related requirements. The descriptions of ‘good education’ also vary from a system-level perspective (focusing on for example on the quality of the whole education system) to describing good education in individual schools.

Regarding the improvement of student achievement, research indicates that it is reasonable to expect that inspection regimes which include standards on teaching and learning, derived from school effectiveness research, will be the most effective. Hattie (2009) summarizes 800 meta-analyses (albeit not always of randomised controlled trials) of factors contributing to student achievement and identifies a number of relevant teaching factors such as ‘teacher expectations’, ‘relationship between teacher and learner’ and ‘professional development of teachers’. According to Scheerens (2009) learning to learn strategies, teacher characteristics such as high expectations, a challenging teaching approach, an orderly learning environment, clear and structured teaching, and an activating choice of teaching arrangements and student grouping are relevant factors contributing to high student achievement.

These meta-analyses suggest that teaching/instruction level conditions are more important than school level conditions in improving student achievement. Most inspectorates however do not explicitly evaluate teaching or teachers on a classroom level, preferring instead to promote improvement through evaluating school level conditions and motivating principals to (build capacity to) improve the teaching and learning in the school. Most inspection frameworks focus on general instruction characteristics or teaching patterns such as described by Scheerens (2009) (e.g. learning time, classroom organization and classroom climate). This leaves an obvious gap when trying to assess the impact of the mode of inspection on actual achievement.

A further issue that needs to be borne in mind when applying the conceptual model described above relates to actual as opposed to intended effects of inspection. While improvement, increased collegiality, enhancement of core capacity and re-balancing of activities in order to take account of national priorities are often cited as the expected outcomes of the inspection process, almost without fail there are other unintended but nonetheless equally important results that have a significant impact on the quality of education provided. While the particularities of each inspection system will lead to a range of consequences peculiar to the national context, there are a number of general themes that emerge. In addition to the issues raised earlier in the paper—for example concentration on short-term objectives to the detriment of long-term planning and changing practices in order to meet the perceived requirements of

the inspection process rather than developing real capacity to enhance the quality of teaching and learning etc—inspections can also result in increased levels of stress throughout the school leading to absenteeism and in more extreme cases, resignation from the profession (Brunsdon et al 2006). House and Howe (2000) and Cook-Sather (2002) also argue that a failure to effectively include the parent and student voice largely disenfranchises these key stakeholders and can result in a disengagement from any long-term improvement strategies developed by schools.

7.2 Intermediate effects: Capacity building and improvement actions

Inspection policy documents in our study routinely refer to capacity building and improvement actions as important objectives of the inspection process. They function as what we describe as intermediate effects of school inspections in that high improvement capacity and improvement actions schools take to address their weaknesses are expected to lead to the eventual intended effect of good education as described in the previous section.

Most inspection documents do not provide a detailed description of what capacity building entails. Geijsel et al. (2009) however describe capacity building as the capacity of schools to enhance the professional learning of teachers and to transform large-scale reform into accountable student-oriented teaching practices. A school is thought to be capable of change if it is competent to implement innovations initiated either by the government (or the Inspectorate of Education) or by the school itself (as a result of the outcomes of self-evaluation), and if necessary, to match both types of innovations. In this context, Geijsel et al. (1999) stress the importance of the school as a learning organization which consciously attempts to expand its learning capacity to optimize its effectiveness (Geijsel et al. 2009). Participation of teachers in decision making, cooperation between teachers and transformational leadership are considered to be important conditions of learning organizations. In summary, the Inspectorates of Education in our study expect to promote capacity building, improvement of school conditions and the ultimate outcome of good education through three interlinked causal mechanisms, namely setting standards and expectations, providing feedback and informing stakeholders.

7.3 Promoting intended effects through setting expectations

The first causal mechanism relies on a process of setting expectations defining good education for schools and their stakeholders. The standards and procedures of the inspection are assumed to infuse schools and their stakeholders with notions and practices of what is considered ‘good education’ in each country. These notions will stimulate schools to embark on processes of adaptation to these norms and of institutionalizing them into their work structures and culture. The idea is that schools will attend to the information included in inspection standards and procedures; they will reflect on it, process it and adapt their goals and their practical ways of working in such a way that they come closer to the normative image of schools communicated by the inspection. Eventually, these notions should drive the school’s own planning, self-evaluations and daily practices.

However, this idea of a reflected and profound in-school process of institutionalizing the expectations of inspections may not always be met in reality as we want to clarify by using DiMaggio and Powell's (1983) account of organizational change. In their view, changes in the organization and daily practices in schools emerge out of the structuration of organizational fields, such as the education and evaluation system in which schools function. These organizational fields provide a context in which the efforts of individual schools to organize their teaching and learning are constrained which, in turn, leads to some homogeneity in structure, culture and output. In an inspection and evaluation context, such constraints typically derive from external actors, such as an Inspectorate of Education and other stakeholders of the school. Schools may be coerced into meeting inspection expectations when stakeholders or organizations upon which they are dependent exert formal or informal pressure on them to meet these expectations. However, schools may also actively seek out examples of peers who successfully meet the expectations of the Inspectorate and mimic their responses to proactively prevent potential action and pressure by these external actors. Examples of how schools try to meet such expectations include the implementation of curricula that are 'inspection-approved', using similar lesson plans with pre-set hours of teaching in specific subjects, and purchasing and implementing quality assurance systems that incorporate the inspection standards.

Although this process may lead to some adaptation to the inspection standards, it may be problematic in three ways. First of all, the changes that are sometimes made in this process of institutionalization may be *ceremonial or cosmetic* (DiMaggio and Powell 1983). McNamara and O'Hara (2009) for example describe how self-evaluations, rather than improving the learning capacity of schools, are often undertaken only as a part of an inspection process and can tend to become an end in themselves for schools used to facilitate the inspection 'dialogue' (Nevo 2006).

Second, even though these changes are ceremonial, *nevertheless they may be consequential* as they may change power relations or structures within schools over the long run, and make schools more homogenous on the inspection indicators, but less responsive to the specific context of the student population and community in which they function (DiMaggio and Powell 1983). Particularly when consequences for not meeting the expectations are feared, schools will seek for standard operating procedures and legitimated rules and structures (such as inspection-approved curricula or quality assurance systems) to show their conformity to the inspection expectations and standards. These expectations may therefore limit flexibility and adaptation of schools to their context. This may prove to be effective in improving student achievement when institutionalization leads schools to implement models and procedures of effective teaching. Standard procedures and structures may also be ineffective in some contexts. Thus, the content of expectations, and particularly whether they include understandable specifications of effective teaching conditions, together with the way they are communicated seem to be decisive factors.

Third, the strategy of promoting quality education through setting standards on a high level of generality may also *run counter to the quality strategy of increased school autonomy* which was promoted by educational policy in many European countries (see Eurydice 2007; Altrichter and Rürup 2010). The goal

of increased school autonomy is to give schools more responsibility for school-based management and to make them more responsive to the needs and potentials of their specific environment. In order to do so, Inspectorates of Education should allow some diversity and should stimulate schools to develop their specific versions and “profiles” of good education. This seems contradictory with a general and common framework of expectations. Many Inspectorates of Education try to solve this conflict by incorporating self-evaluations into their evaluation of schools and by promoting self-evaluations in which school also set their own goals. Both approaches may however lead to exposing schools to contradictory demands where the inspection expectations are likely to be the more dominant as these are the ones that are accompanied by consequences.

7.4 Promoting intended effects through feedback

The second causal mechanism to promote intended effects is through the feedback that is provided to schools during inspection visits and/or in inspection reports (and the consequences for not following up on the feedback). Schools are assumed to use the feedback to improve, and stakeholders are expected to take note of the feedback and hold schools accountable for their use of the feedback for improvement.

On an interpersonal level, feedback may be a powerful instrument for promoting and guiding learning and development. Research has shown that feedback may have a positive effect on the performance of learners, but not under all circumstances (see Kluger and DeNisi 1996; Hattie and Timperley 2007). Kluger and DeNisi (1996) point to some features which are associated with effective feedback:

1. *Feedback cues*: a feedback message contains cues which guide the recipient’s attention. Particularly helpful are cues which make learners turn their attention to their work on task, to the learning process and to the learners’ self-regulation. Feedback cues which induce learners to focus on their personal qualities may be distracting (see Visscher and Coe 2003, p. 327). This may also be the case with feedback with respect to the quality of results if the feedback does not contain additional information about productive ways of focusing the work on task (see Hattie and Timperley 2007, p. 90).
2. *Task characteristics*: Kluger and DeNisi (1996) found that it was easier to give effective feedback for simple tasks.
3. *Situational and personal variables*: clear, specific and challenging goals are moderately associated with useful feedback. If feedback is experienced as threatening, it will be associated with fewer effects (see Visscher and Coe 2003, p. 328).

It is intuitively plausible to also use this powerful interpersonal mechanism on the level of organisations, such as schools, and even on the system level, as is proposed by evidence-based governance models. In essence, this involves producing better data about the processes and results of schools and using it for improving schools. However, it seems to be more difficult than expected to tap this plausible potential in the reality of school improvement; international research results are not conclusive as to whether or not schools and teachers can use feedback for developing classrooms and schools and, consequently, for

improving results (see Coe 2002; van Ackeren 2003). Ehren and Visscher (2008) found that all ten schools in a Dutch case study had derived consequences from the inspection and put them into practice according to the school principals' reports (similar to Gärtner et al. 2009, for German schools). According to Matthews and Sammons (2004), clear and explicit reports and feedback to schools are successful in informing the improvement plan after school inspections, and this, in turn, results in more effective school action.

On the other hand, many teachers are not willing to change their teaching after an inspection (see Chapman 2001). Gärtner et al. (2009, p. 10) found in the German state of Brandenburg that 28 % of all schools which had been inspected during the first 2 years of a newly introduced inspection scheme ($N=170$) reacted actively to the inspection report, while 34 % behaved rather passively and 38 % showed medium activity by discussing the inspection report in a staff meeting. Moreover, teachers who are, in principle, willing to read and learn from feedback find it difficult to develop ideas for improvement from it and to put them into practice. Numerous studies show that teachers have problems in deriving conclusions from the feedback of assessment data (see Altrichter 2010; p. 234; Bensen and Gathen 2004, p. 228; Maier 2006; Peek 2004; Grabensberger et al. 2008; Steffens 2009, p. 1). If there is classroom development as a consequence of data feedback, then it rarely includes thorough innovation; rather there is repetition of content and tasks (Groß Ophoff et al. 2006, p. 8; Maier 2007; Hosenfeld et al. 2007), assessment task formats are adopted for teaching (see Leutner et al. 2007; Maier 2007; Hosenfeld et al. 2007) or slight changes in classroom interaction take place (see Schildkamp et al. 2009, p. 86).

An explanation for these difficulties to capitalize on data feedback was put forward by Visscher and Coe (2002, p. 247). They pointed to the fact that many feedback arrangements at school and system level fail to meet the conditions for effective feedback as stated by Kluger and DeNisi (1996). Comparative feedback of student performance results and feedback by inspection reports may direct the recipients' attention to comparison with other schools and to personal shortcomings, but may include little concrete information on how to improve performance. The request to develop corrective action from data feedback may rightly be seen as a rather complex and—particularly in the case of high-stakes systems—threatening task (see Visscher and Coe 2003, pp. 247 and 328).

Another explanation emphasizes that feedback on the organizational and system level is just one element in a more complex arena in which actors have to attend to many more pieces of information and other considerations. In the reality of a multi-level system, evaluative feedback is an “unspecific impulse” (Kuper 2005, S. 101 f.), not a road map for innovation. Feedback is another contribution which may ‘support negotiation between actors’ (ibid.); however, it cannot be expected to mitigate the ambiguity of organizational decision making and to replace the micro-political negotiation processes between actors with divergent interests.

Lacking competences with respect to data reading and analysis (see Schwippert 2004, p. 77; Peek 2006, p. 354), and knowledge with respect to teaching alternatives (Dubs 2006; Arnold 2007; Darling-Hammond 2004; Steffens 2009) might be other reasons for limited use of feedback data for

classroom and school development. These explanations seem to underline the fact that feedback will not automatically produce development; careful design of content, format and communication of feedback is essential as is implementation and support for teachers and schools who are to build up competences for making use of feedback.

7.5 Promoting intended effects through stakeholders

The third mechanism to promote intended effects of school inspections in our study is the provision of information on the inspection process to a broad range of stakeholders. While Inspectorates choose to interpret differently who is to be considered a relevant stakeholder, each of the systems sees stakeholder involvement as being essential for improvement. Stakeholders, such as parents and school boards, are expected to facilitate school improvement through actions of ‘voice’, ‘choice’ and ‘exit’. These stakeholders are seen as key change agents in a post-inspection school environment. In this understanding of stakeholder involvement, the informed parent is expected to hold the school to account by using the feedback provided by the inspection to demand improvement in specific areas. In the event of this improvement not taking place, ‘market mechanisms’ will induce parents to move their children to different schools which, in turn, will stimulate schools to improve their performance (Hoxby 2003).

Attractive though this theory may be, in practice it ignores a range of social, historical and other realities that often see parents as being either unable or unwilling to engage in the type of behaviour expected of them. Studies on school choice indicate that there are more factors than “school quality” which affect parents’ decision and that power and willingness to use choice options are unevenly distributed among different social groups (Belfield and Levin 2009; Buckley and Schneider 2003; Reback 2005). Thus, according to Wood (2003), teachers are still to be seen as the primary agents for school improvement. Other national studies (Ofsted 2009; Dillon 2011) suggest that despite the provision of copies of inspection reports in a public forum, parents still feel disempowered when faced with those who they feel are the ‘professionals’ in the field of school improvement.

What has been argued above does not invalidate the conceptual model. Rather it demonstrates the range of variables that need to be taken into account when seeking to research it on a transnational level and equally, when seeking to implement it in specific localities. The model still has an internal coherence and an operational logic that will allow an evaluation of the causal mechanisms of effective school inspections across a range of divergent inspection systems. There will be, however, issues relating to inherent contradictions in how systems are applied and gaps in the logic of that application that will need to be acknowledged when assessing the impact of inspection across a range of areas. In particular, the definition of school capacity, the setting of expectations for effective teaching conditions and the gaps in the causal chain between the actions of stakeholders and the improvement of schools leading to high student achievement all need to be addressed when designing and evaluating research in this field.

Appendix

Table 4 Summary of program theories of Inspectorates of Education in six European countries

Type of evaluation	Standards	Thresholds	Consequences (sanctions, rewards, interventions)	Mechanisms	Expected effects
<p>The Netherlands</p> <ul style="list-style-type: none"> - Risk-based school inspections: only potentially failing schools are scheduled for inspection visits in which nine standards are evaluated (e.g. student achievement, curriculum, climate etc.) - Additional (monitoring) inspection visits to failing schools - All schools receive at least one (thematic) visit once every 4 years - Communication of inspection activities to school boards 	<ul style="list-style-type: none"> - The students' results reach a level that may be expected (taking the characteristics of the student population into account) - The school has a system for assuring the quality of its education - The subject matter offered to pupils prepares them for continuing education - The students get enough lesson time to learn the subject matter - The school systematically assesses the progress of students - The school climate is safe and stimulating 	<ul style="list-style-type: none"> - The outcome of the risk analysis of a school may be green (no risk), orange (potential failing, further analysis during a visit), red (likely failing, further analysis during visit). - A school is failing when the student achievement results at the end of primary education have been insufficient for more than 3 years. - A school is assessed as highly underdeveloped when the results at the end of primary education have been insufficient for more than 3 years and the teaching and learning or the pupil care is insufficient. 	<ul style="list-style-type: none"> - Consequences: school boards of failing schools have to develop an action plan that is monitored by the Inspectorate, eventually administrative and financial sanctions, failing schools are put on a 'black list' that is published on the internet. 	<ul style="list-style-type: none"> - Intensive monitoring/ inspection of weak schools (no/limited inspection of strong schools) - Informing and empowering stakeholders and internal supervisors - Setting expectations particularly on student achievement results for schools and for internal supervisors, school boards. - Promoting self-evaluations is part of broader requirements 	<ul style="list-style-type: none"> - Preventing (further decline of) failing schools - Improvement of schools - Well developed schools use autonomy to make decisions - Good education (as defined by inspection standards)

Table 4 (continued)

Type of evaluation	Standards	Thresholds	Consequences (sanctions, rewards, interventions)	Mechanisms	Expected effects
	<ul style="list-style-type: none"> - The pedagogical behavior of teachers meets the basic requirements - De didactical behavior of teachers meets the basic requirements - Children with specific educational needs receive the care they need 			<p>on governance and quality assurance in/of schools, self evaluations are used in inspections.</p> <ul style="list-style-type: none"> - Internal supervisors, school boards, parents; schools are expected to actively inform stakeholders 	
England	<ul style="list-style-type: none"> - School inspections (section 5 inspections) of all maintained schools, typically every 5 years, using standards on pupils' outcomes, effectiveness of the provision and effectiveness of the leadership and management of the school to evaluate schools 	<p>Schools are evaluated as "outstanding", "good", "satisfactory" or "inadequate", with "inadequate" schools being subdivided into "notice to improve" and "special measures, and around 40 % of "satisfactory" schools being categorised as requiring a monitoring inspection</p>	<p>Consequences: special measures (schools are required to work with local authority to improve and a small minority are under threat of closure) choice to improve and weaker satisfactory schools undergo further monitoring inspections. Schools judged as outstanding can apply for special benefits/ special status e.g. for "academy status", 'teaching schools' or head teachers may apply to become "national leaders of education".</p>	<p><i>Individual schools:</i></p>	<ul style="list-style-type: none"> - Promoting improvement of schools
	<ul style="list-style-type: none"> - Pupils' personal development and well-being 		<p>Published inspection reports are expected to create market forces leading to increased/decreased demand for school places. This in</p>	<p>- Standards set expectations and (when school's expectations meet the standards) increase the school's confidence or</p>	<ul style="list-style-type: none"> - Ensuring services are user-focused and provide value for money
	<ul style="list-style-type: none"> - Monitoring inspections (section 8) of around 40 % of schools graded satisfactory and all school graded as inadequate. 				

Table 4 (continued)

Type of evaluation	Standards	Thresholds	Consequences (sanctions, rewards, interventions)	Mechanisms	Expected effects
Monitoring inspections focus on progress in making improvements.	<ul style="list-style-type: none"> - Quality of provision, including teaching and learning, curriculum and care, guidance and support for pupils - Leadership and management 		<ul style="list-style-type: none"> - tum could lead to increased/decreased job security, respectively. 	<ul style="list-style-type: none"> - challenge (when school sets too low expectations) the school's view of effectiveness. - Recommending priorities for future action and checking progress on these recommendations - Constructive dialogue between inspectors and senior school staff - Complementing school's self evaluation - Increased monitoring of weak schools - Informing policy with data and evidence in all areas of care and learning - Highlighting good practice - Market mechanisms: voice and exit - Reward of being an outstanding school - Avoiding sanctions (job loss and school closure) - Setting expectations for schools on specific criteria and descriptors, national examples of good practice, also for the 'wider public'. 	<ul style="list-style-type: none"> - Encouraging improvement of the education system as a whole
	<ul style="list-style-type: none"> - The overall effectiveness of the school <p>The school's capacity to improve</p>				

Table 4 (continued)

Type of evaluation	Standards	Thresholds	Consequences (sanctions, rewards, interventions)	Mechanisms	Expected effects
Sweden	<p>- Inspections of all schools once every 4/5 years, evaluating fulfilment of national objectives, student achievement, safety, orderly classrooms, equal access to education and legal rights for all students. For schools which on the basis of initially available knowledge are judged to be well-functioning a "Basic" inspection is conducted. For the other schools a more elaborated and intense "Widened" inspection is conducted.</p> <p>- Thematic quality evaluations of specific areas (such as teaching of specific subject matter), using an approach supported by research- and experience-based</p>	<p>- Attainment of objectives and results (knowledge)</p>	<p>- Consequences: withdraw the license and funding of independent schools that do not comply with rules and regulations.</p>	<p>- Promoting acceptance of feedback through good relationship between head teacher and principal.</p> <p>- Schools are required to use self evaluation forms that should inform school development plans.</p> <p>- Stakeholders are parents and local authorities.</p> <p><i>All schools/system:</i></p> <p>- Informing and advising the government about aspects relating to the quality of schools</p> <p><i>Individual schools:</i></p>	<p>- Ensure the right of all students to good education in a safe environment</p> <p>- Improvement of inspected schools</p> <p>- Reporting/feedback on quality of individual schools and (in the case of weak schools) instructions to correct problems within a specific time frame; Inspectorate</p>
	<p>- Educational leadership and school development (focus on the school's tasks; teachers' assessment and evaluation; the principal's</p>		<p>- Instructions to schools to correct weak aspects within a specific time frame, monitoring of implementation of corrections</p>	<p>Reporting/feedback on quality of individual schools and (in the case of weak schools) instructions to correct problems within a specific time frame; Inspectorate</p>	

Table 4 (continued)

Type of evaluation	Standards	Thresholds	Consequences (sanctions, rewards, interventions)	Mechanisms	Expected effects
knowledge of relevance to the area..	responsibility for monitoring and evaluation) - The learning environment of the school (focus on learning; trust in the student's ability; orderly and safe environment) - The rights of the individual student (grading and assessment; instructional time and offering of elective subjects; special support; study and vocational guidance; fees; the decision making of the principal)			monitors implementation of corrections. - Thematic inspections set knowledge base that will help set expectations and will provide examples of good practice to use when improving. All schools/system.	- Improvement of education system
Ireland	- Whole school evaluations of all schools at unspecified intervals (generally once every 5 years), evaluating 5 standards for each subject department in a school and for the entire school	No threshold	- Inspection reports are published on the website of the Department of Education and Science. Except in the most extreme cases the outcome of an evaluation does not have consequences for schools.	- General reports of thematic quality evaluations contain good practices that inform individual schools on potential improvement. - The general reports inform national government on potential improvement of all schools. Complementing internal continuous quality assurance/evaluation and improvement activities in schools through setting standards for self-evaluation, facilitating, full participation of the whole-school community in the process; to contribute to school	- Promote continuing improvement in the quality of education

Table 4 (continued)

Type of evaluation	Standards	Thresholds	Consequences (sanctions, rewards, interventions)	Mechanisms	Expected effects
- Subject inspections at post-primary level	- At post-primary level: the quality of the school management, the quality of school planning, the quality of curriculum provision, the quality of teaching and learning, the quality of support for pupils	A single overall rating is not applied to the evaluation at present		development; to ensure school accountability; to enable teachers and schools to use] the evaluation criteria for school self-review and improvement/ - Inspection themes may guide school self evaluations	- Promote self evaluation and continuous development by schools
- Thematic evaluations				- Teacher unions are relevant stakeholders to accept inspection standards	- Contributing to development and to the support of the education system
- Focused inspections				- Accepting feedback by schools is promoted by inviting schools to respond to inspection report - Promoting self-evaluations is part of a broader system to promote quality assurance of schools, self evaluations are used in inspections - Parents are expected to act on the inspection report - Strong focus on quality assurance should lead to improving evaluation capacity in schools)	

Table 4 (continued)

Type of evaluation	Standards	Thresholds	Consequences (sanctions, rewards, interventions)	Mechanisms	Expected effects
Austria	<ul style="list-style-type: none"> - Regular school inspection visits (every 2–4 years) of all compulsory schools in primary and secondary education by a team of two inspectors 	<ul style="list-style-type: none"> - Teaching and learning 	<ul style="list-style-type: none"> - Target agreements in a school development plan which are followed up after approximately 2 years - No sanctions or rewards 	<ul style="list-style-type: none"> - School use relevant criteria and procedures of school quality in their self evaluation (as communicated to them by the Inspectorate) - Inspectorate communicates obligation for quality development and supports head teachers in school development - Acceptance of inspection feedback is motivated by requirement to include feedback in school development plan - Inspections should strengthen commitment for school development - School is required to actively inform stakeholders on inspection report, such as parents, community, teachers, school maintaining body. 	<ul style="list-style-type: none"> - Educational effectiveness and quality of the individual school - Legal and administrative compliance - Equivalence and comparability of various educational provisions within the system
Czech Republic	<ul style="list-style-type: none"> - Equal opportunities in education, <p>Inspectorate schools regularly in a 3-year cycle, also additional thematic evaluations (e.g. ICT). Regular inspections include observations of school activities, documents and interviews.</p>	<ul style="list-style-type: none"> - Classroom and schools as a space for living - Partnerships in schools and external relationships - School management - Professionalism and personnel development 	<ul style="list-style-type: none"> - An unsatisfying result of one of the two checks can lead to a fine, collected by the CSI, but transferred to the state budget. 	<ul style="list-style-type: none"> - The development of the personality of the child, pupil and student as well as achievement of educational aims in specific schools and school facilities. 	

Table 4 (continued)

Type of evaluation	Standards	Thresholds	Consequences (sanctions, rewards, interventions)	Mechanisms	Expected effects
<p>The role of inspection is divided into evaluation—institutional inspection activities and checking—state checks and public—legal audits. There are two distinct outcomes of each task—inspection reports for evaluation and protocols for checks.</p>	<p>- Management of school</p>			<p>prevailing controlling activities will be supplemented by guidance and consulting work aimed to help schools with their internal evaluation and to introduce a new evaluation culture ... The Czech School Inspectorate is expected not only to ascertain the state of education activity in schools and the conditions in which education is taking place, and to identify shortcomings but also to motivate school heads and teachers to work towards improvements. This means not only indicating areas in which schools are not making full use of their opportunities, but also informing them about positive examples and models.</p>	<p>Schools should develop evaluation culture</p>
	<p>- Assumptions for the proper functioning of schools</p>		<p>- A bad “inspection report” can lead – after a period for improvement and repeated inspection – to the proposal of the Senior Chief Inspector to the authorities to remove the school from the register of funded schools.</p> <p>- Schools which have been inspected are obliged to adopt measures in order to correct, without undue delay,</p>		

Table 4 (continued)

Type of evaluation	Standards	Thresholds	Consequences (sanctions, rewards, interventions)	Mechanisms	Expected effects
			<p>shortcomings identified during the inspection, however not later than within the period specified by the CSI. On the basis of the inspection results the founders takes, without undue delay, needed measures at their schools.</p> <p>- In the event that it is ascertained that the school in question failed to take necessary steps or if significant shortcomings are identified in the school, the Chief School Inspector may submit to the Body of the Register of Educational Facilities a proposal for removal of the school or the field of education concerned from the Register of Educational Facilities.</p> <p>- The follow-up inspections usually focus on how the school has remedied the inadequacies found during the last inspection.</p>		
	- Course of education				
	- Partnership				
	- Demonstration of achieved level of key competencies through the educational content				
	- Educational outcomes of pupils on the school level				
	Overall evaluation of school				

References

- Altrichter, H. (2010). Schul- und Unterrichtsentwicklung durch Datenrückmeldung. In H. Altrichter & K. Maag Merki (Eds.), *Handbuch Neue Steuerung im Schulsystem* (pp. S. 219–S. 254). Wiesbaden: VS Verlag.
- Altrichter, H., & Maag Merki, K. (Eds.). (2010). *Handbuch Neue Steuerung im Schulsystem* (pp. S. 219–S. 254). Wiesbaden: VS Verlag.
- Altrichter, H., & Rürup, M. (2010). Schulautonomie und die Folgen. In H. Altrichter & K. Maag Merki (Eds.), *Handbuch neue Steuerung im Schulwesen* (pp. 111–144). Wiesbaden: VS.
- Arnold, K.-H. (2007). Bildungspolitische, diagnostische und didaktische Bedingungen von Schulleistungs Evaluationen. *Empirische Pädagogik*, 21, 448–457.
- Belfield, C. R., & Levin, H. N. (2009). Market reforms in education. In G. Sykes, B. Schneider, D. N. Plank, & T. G. Ford (Eds.), *Handbook of education policy research* (pp. 513–527). New York: Routledge.
- Bonsen, M., & Gathen, J. (2004). Schulentwicklung und Testdaten. In H. G. Holtappels, K. Klemm, H. Pfeiffer, H.-G. Rolff, & R. Schulz-Zander (Eds.), *Jahrbuch der Schulentwicklung. Band 13* (pp. S. 225–S. 252). Weinheim: Juventa.
- Brimblecombe, N., Shaw, M., & Ormston, M. (1996). Teachers' intention to change practice as a result of OFSTED school inspections. *Educational Management & Administration*, 24(4), 339–354.
- Brunsdon, V., Davies, M., & Shevlin, M. (2006). Anxiety and stress in educational professionals in relation to Ofsted. *Education Today*, 56(1), 24–31.
- Buckley, J., & Schneider, M. (2003). Making the grade: comparing DC charter schools to other DC public schools. *Educational Evaluation and Policy Analysis*, 25, 203–216.
- Chapman, C. (2001). Changing classrooms through inspections. *School Leadership Management*, 1(1), 59–73.
- Coe, R. (2002). Evidence on the role and impact of performance feedback in schools. In A. J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. S. 3–S. 26). Lisse: Swets & Zeitlinger.
- Cook-Sather, A. (2002). Authorizing student's perspectives: towards trust. *Dialogue and Change in Education Educational Researcher*, 19(5), 2–14.
- Darling-Hammond, L. (2004). Standards, accountability and school reform. *Teachers College Record*, 106(6), 1047–1085.
- De Wolf, I. F., & Janssens, F. J. G. (2007). Effects and side effects of inspections and accountability in education: an overview of empirical studies. *Oxford Review of Education*, 33(3), 379–396.
- Dillon, S. (2011). *The way I see it is...Whole-school evaluation in Irish post-primary schools from the perspectives of principals, teachers, parents and students*. Unpublished Doctoral Dissertation. [Online] www.dcu.ie/doras.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(1), 147–160.
- Doolaard, S., & Karstanje, P. (2001). Gebruik van publieke prestatie-indicatoren voor schoolverbetering. In A. B. Dijkstra, S. Karsten, R. Veenstra, & A. J. Visscher (Eds.), *Het oog der natie: scholen op rapport; standaarden voor de publicatie van schoolprestaties* (pp. 155–173). Assen: Koninklijke Van Gorcum BV.
- Dronkers, J., & Veenstra, R. (2001). Schoolprestatie-indicatoren in het voortgezet onderwijs: start, reacties en vervolg. In A. B. Dijkstra, S. Karsten, R. Veenstra, & A. J. Visscher (Eds.), *Het oog der natie: scholen op rapport; standaarden voor de publicatie van schoolprestaties* (pp. 21–36). Assen: Koninklijke Van Gorcum BV.
- Dubs, R. (2006). Bildungsstandards: Das Problem der schulpraktischen Umsetzung. *Netzwerk—Die Zeitschrift für Wirtschaftsbildung*, 1, 18–29.
- Ehren, M. C. M., & Visscher, A. J. (2006). Towards a theory on the impact of school inspections. *British Journal of Educational Studies*, 54(1), 51–72.
- Ehren, M. C. M., & Visscher, A. J. (2008). The relationship between school inspections, school characteristics and school improvement. *British Journal of Educational Studies*, 56(2), 205–227.
- Ehren, M. C. M., Leeuw, F. L., & Scheerens, J. (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.
- Elmore, R. F., & Fuhrman, S. H. (2001). Research finds the false assumption of accountability. *Phi Delta Kappan*, 67(4), 9–14.
- Eurydice (2004). *Evaluation of schools providing compulsory education in Europe*. <http://www.eurydice.org/portal/page/portal/Eurydice>. Accessed January 2011.
- Eurydice. (2007). *School autonomy in Europe: policies and measures*. Brussels: Eurydice.
- Faubert, V. (2009). School evaluation: current practices in OECD countries and a literature review. *OECD Education Working Papers*, No. 42, OECD Publishing. doi:10.1787/218816547156.

- Gärtner, H., Füsemann, D., & Pant, H. A. (2009). Wirkungen von Schulinspektion aus Sicht betroffener Schulleitungen. *Empirische Pädagogik*, 23, 1–18.
- Geijsel, F., Berg, van den, R., & Slegers, P. (1999). The innovative capacity of schools in primary education: a qualitative study. *Qualitative studies in education*, 12(2), 175–191.
- Geijsel, P., Slegers, P. J. C., Stoel, R. D., & Kruger, M. L. (2009). The effect of teacher psychological and school organizational and leadership factors on teachers' professional learning in Dutch schools. *The Elementary School Journal*, 109(4), 1–22.
- Grabensberger, E., Freudenthaler, H. H., & Specht, W. (2008). *Bildungsstandards: Testungen und Ergebnisrückmeldungen auf der achten Schulstufe aus der Sicht der Praxis*. Graz: Bifie.
- Groß Ophoff, J., Koch, U., Helmke, A., & Hosenfeld, I. (2006). Vergleichsarbeiten für die Grundschule– und was diese daraus machen (können). *Journal für Schulentwicklung*, 10(4), 7–12.
- Gustafsson and Myrberg (in prep). *School inspections of Swedish schools: a critical reflection on intended effects, causal mechanisms and methods*. http://schoolinspections.eu/wp-content/uploads/downloads/2012/05/Sweden_PT.pdf. Accessed November 2011.
- Hanushek, E. A., & Raymond, M. E. (2002). Lessons about the design of state accountability systems. Paper prepared for 'Taking Account of Accountability: Assessing Policy and Politics', Harvard University.
- Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analysis relating to achievement*. Milton Park: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High stakes: testing for tracking, promotion, and graduation*. Washington: National Academy Press.
- Hosenfeld, I., Groß Ophoff, J., & Koch, U. (2007). *Vergleichsarbeiten in Klassenstufe 3 (VERA 3)*. Präsentation bei der 7. Tagung "Empiriegestützte Schulentwicklung" in Mainz. Landau: Universität Koblenz-Landau.
- House, E., & Howe, K. (2000). Deliberative democratic evaluation. *New Directions in Evaluation*, 85, 3–26.
- Hoxby, C. (Ed.). (2003). *The economics of school choice*. Chicago: University of Chicago Press.
- Hughes, G., Mears, R., & Winch, C. (1997). An inspector calls? Regulation and accountability in three public services. *Policy and Politics*, 25(3), 299–314.
- Husfeldt, V. (2011). Wirkungen und Wirksamkeit der externen Schulevaluation. Überblick zum Stand der Forschung. *Zeitschrift für Erziehungswissenschaft*. doi:10.1007/s11618-011-0204-5.
- Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behaviour in organizations. *Journal of Applied Psychology*, 64(4), 349–371.
- Karsten, S., & Visscher, A. J. (2001). Ervaringen met het openbaar maken van schoolprestaties in Engeland en Frankrijk. In A. B. Dijkstra, S. Karsten, R. Veenstra, & A. J. Visscher (Eds.), *Het oog der natie: scholen op rapport* (pp. 36–53). Assen: Koninklijke Van Gorcum.
- Kerr, S. (1975). On the folly of rewarding A, while hopping for B. *The Academy of Management Journal*, 18(4), 769–783.
- Klerks, M. (submitted). The effect of school inspections: a systematic review. *Oxford Review*.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance. *Psychological Bulletin*, 119(2), 254–284.
- Kogan, M., & Maden, M. (1999). An evaluation of evaluators: the OFSTED system of school inspection. In C. Cullingford (Ed.), *An inspector calls: Ofsted and its effect on school standards* (pp. 9–32). London: Kogan Page Limited.
- Koretz, D. M. (2003). Using multiple measures to address perverse incentives and score inflation. *Educational Measurement*, 22(2), 18–26.
- Kotthoff, H. G., & Böttcher, W. (2010). *Neue Formen der „Schulinspektion“: Wirkungshoffnungen und Wirksamkeit im Spiegel empirischer Bildungsforschung*. In H. Altrichter, K. Maag Merki (Eds.), *Handbuch Neue Steuerung in Schulsystem* (pp. 295–325). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kuper, H. (2005). *Evaluation im Bildungssystem*. Stuttgart: Kohlhammer.
- Leeuw, F. L. (2003). Reconstructing program theories: methods available and problems to be solved. *American Journal of Evaluation*, 24(1), 5–20.
- Leutner, D., Fleischer, J., Spoden, C., & Wirth, J. (2007). *Schulrückmeldungen in landesweiten Lernstandserhebungen*. Präsentation bei der 7. Tagung "Empiriegestützte Schulentwicklung" in Mainz. Landau: Universität Koblenz-Landau.
- Luginbuhl, R., Webbink, D., & De Wolf, I. (2009). Do inspections improve primary school performance? *Educational Evaluation and Policy Analysis*, 31(3), 221–237.
- Maier, U. (2006). Können Vergleichsarbeiten einen Beitrag zur Schulentwicklung leisten? *Journal für Schulentwicklung*, 10(4), 20–28.
- Maier, U. (2007). *Lehrereinschätzungen zu zentralen Tests und Leistungsrückmeldungen*. Präsentation auf der 7. Tagung "Empiriegestützte Schulentwicklung" in Mainz. Landau: Universität Koblenz-Landau.

- Malen, B. (1999). On rewards, punishments, and possibilities: teacher compensation as an instrument for education reform. *Journal of Personnel Evaluation in Education*, 12(4), 387–394.
- Mason, R. O., & Mitroff, I. I. (1981). *Challenging strategic planning assumptions; theory, cases and techniques*. New York: Wiley.
- Matthews, P., & Sammons, P. (2004). *Improvement through Inspection*. London: Ofsted.
- McGarvey, N., & Stoker, G. (1999). *Intervention, inspection, regulation and accountability in local government. DETR—Interim literature review*. London: DETR.
- McNamara, G., & O'Hara, J. (2008). The importance of the concept of self-evaluation in the changing landscape of education policy. *Studies in Educational Evaluation*, 34, 173–179.
- McNamara, G., & O'Hara, J. (2009). *Where global meets local: contexts, constraints and consensus in school evaluation in Ireland. Sage international handbook of educational evaluation* (pp. 273–291). London: Sage Publications.
- Nevo, D. (2006). In S. J. Greene & M. Mark (Eds.), *Evaluation in education. Sage handbook of evaluation I* (pp. 441–460). London: Sage Publications.
- Nichols, S. L., Glass, G. V., & Berliner, D. C. (2006). High-stakes testing and student achievement: does accountability pressure increase student learning? *Education Policy Analysis Archives*, 14(1), Retrieved 14 November 2008 from <http://epaa.asu.edu/epaa/v14n1>.
- Ofsted (2009). *Parents' perceptions of Ofsted's work—a report of the NFER for Ofsted*. London: Ofsted.
- Peek, R. (2004). Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik (QuaSUM). *Empirische Pädagogik*, 18, 82–114.
- Peek, R. (2006). Dateninduzierte Schulentwicklung. In H. Buchen & H.-G. Rolff (Eds.), *Professionswissen Schulleitung* (pp. 1343–1366). Weinheim: Beltz.
- Reback, R. (2005). *Supply and demand in a public school choice program*. New York: National Center for the Study of Privatization in Education.
- Rosenthal, L. (2004). Do school inspections improve school quality? *Economics of Education Review*, 23 (2), 143–151.
- Scheerens, J. (2009). *Review and meta-analyses of school and teaching effectiveness. The Netherlands/department of Educational Organization and Management*. www.iqb.hu-berlin.de/lehre/dateien/rapportScheerens.pdf. Accessed November 2011.
- Scheerens, J., Seidel, T., Witziers, B., Hendriks, M., & Doomekamp, G. (2005). *Positioning and validating the supervision framework. Positioning the supervision frameworks for primary and secondary education of the Dutch Educational Inspectorate in current educational discourse and validating core indicators against the knowledge base of educational effectiveness research*. Enschede/Kiel: University of Twente/IPN
- Schildkamp, K., Visscher, A., & Luyten, H. (2009). The effects of the use of a school self-evaluation instrument. *School Effectiveness and School Improvement*, 20(1), 69–88.
- Schwippert, K. (2004). Leistungsrückmeldungen an Grundschulen im Rahmen der internationalen Grundschul-Lese-Untersuchung (IGLU). *Empirische Pädagogik*, 18, 62–81.
- Stecher, B. M. (2002). Consequences of large-scale, high-stakes testing on school and classroom practices. Tests and their use in test-based accountability systems. In L. S. Hamilton, B. M., Stecher, S. P. Klein (Eds.), *Making sense of Test-based accountability in education*. Santa Monica: Rand cooperation. http://www.rand.org/pubs/monograph_reports/MR1554/.
- Steffens, U. (2009). *Lernstandserhebungen in den deutschen Ländern—Probleme und Perspektiven*. Unv. Ms. Wiesbaden: Institut für Qualitätsentwicklung.
- Toulmin, S. (1958). *The uses of argument*. Cambridge: University Press.
- Toulmin, S. (1964). *The uses of argument*. Cambridge: University Press.
- van Ackeren, I. (2003). *Evaluation, Rückmeldung und Schulentwicklung*. Münster: Waxmann.
- van Bruggen, J. C. (2010). *Inspectorates of education in Europe; Some comparative remarks about their tasks and work*. SICI report: www.sici-inspectorates.org.
- van de Grift, W. (2007). Quality of teaching in four European countries: a review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152.
- Visscher, A. J., & Coe, R. (2002). *School improvement through performance feedback* (27–39). Lisse: Swets & Zeitlinger.
- Visscher, A. J., & Coe, R. (2003). School performance feedback systems: conceptualisation, analysis, and reflection. *School Effectiveness and School Improvement*, 14(3), 321–349.
- Whitby, K. (2010). *School inspection: recent experiences in high performing education systems; literature review*. Reading: CfBT Education Trust.
- Wood, E. (2003). The power of pupil perspectives in evidence-based practice: the case of gender and underachievement. *Research Papers in Education*, 18(4), 365–383.