

The effect of school inspections: a systematic review

Paper presented at the ORD, Wageningen, The Netherlands

June 20, 21 and 22, 2012

Marielle C.J.L. Klerks

Inspectorate for Youth Care, the Netherlands

Correspondence

Dr. Marielle C.J.L. Klerks

Inspectorate for Youth Care

PO Box 483

3500 AL Utrecht

The Netherlands

+31627743627

MC.Klerks@InspectieJZ.nl

The effect of school inspections: a systematic review

Abstract

This article is based on a systematic review. It aims at obtaining evidence based insight into the effect of school inspections on the educational quality of schools. Fourteen articles which discuss various types of research, were identified as relevant. A meaningful distinction concerns the so-called outcome measures against which the effect of school inspections are measured: 1. school improvement; 2. behavioural change of teachers; or 3. student achievement results. Overlooking the overall results of the systematic review, no evidence has been found that school inspections automatically lead to the improvement of the educational quality. Also it cannot be concluded that characteristics of school inspections themselves lead to the improvement of educational quality. Instead, research shows that in practice there is a complex interaction between different characteristics of school inspections and the inspector on the one hand, and the school with its pupils, teachers and management on the other hand.

Keywords

Educational Quality, Evidence Based, Inspection, School Improvement, Review

Introduction

In many countries, especially in Europe, school inspections are used as an instrument to monitor and improve the educational quality of schools. Most Inspectorates of Education expect from visited schools improvement actions which are related to the feedback given during the inspection visit and in the inspection report. They also expect that the improvement actions are related to the inspection frameworks and ultimately lead to more effective teaching and learning in schools (De Wolf & Janssens, 2007; Ehren, Altrichter, McNamara & O'Hara, sub.). There is nevertheless little insight into what in practice school inspections mean in terms of improvement of the quality of education. There are several reasons why a better understanding of the effects of school inspections is important. For an Inspectorate itself, a better insight in the effects it brings about is necessary in order to improve the quality of regulation and to make responsible choices which priorities must be set. Especially in times of cutbacks, the available capacity must be used as effectively as possible. In relation to schools, an Inspectorate can use these insights to prove its added value and justify itself to the public and to policy makers (cf. De Winter, 2010; Jansen & Smits, 2011; Smeets & Verkroost, 2011).

The study presented in this article is a literature review conducted of relevant scientific literature, according to the principles of a systematic review (cf. Hemingway & Brereton, 2009). The aim of this review is to obtain evidence based insight into the effect of school inspections on the educational quality of

schools.¹ In the next section the research questions will be presented. This will be followed by the methodology and the results of the systematic review of the scientific literature. The article ends with a conclusion, discussion and interpretation of the results.

Research questions

The systematic review has two main research questions:

1. What effect do school inspections have on the improvement of the educational quality of schools?
2. Which characteristics of school inspections contribute to the effect on the improvement of the educational quality of schools?

It should be noted that both research questions concern the effect of school inspections on the improvement of the quality of education. In research with a quasi-experimental design on level 3 or higher on the Maryland Scientific Methods Scale (MSMS) for internal validity, an effect refers to a causal relation. In studies with a lower MSMS-level or qualitative research, the effect refers to a supposed connection which cannot be stated with certainty (cf. Farrington, Gottfredson, Sherman & Welsh, 2002).

In 2010 Winter and De Ridder (2010) conducted a literature review of international studies on the effectiveness and impact of regulation. The results of this review showed that research -certainly also within educational

regulation- is still in its infancy. Of the in total seventy identified studies at MSMS-level 3 or higher, there were four in the field of education. Three of the four studies are relevant in relation to the aforementioned aim². First, the study of Luginbuhl, Webbink and De Wolf (2009) examined whether school inspections of the Dutch Inspectorate of Education had effect on the test scores of pupils in primary education. The two other studies by Shaw, Newton, Aitkan and Darnell (2003) and Rosenthal (2004) investigated whether an inspection visit carried out by the Office of Standards in Education (Ofsted), had an effect on the examination results of secondary schools in the United Kingdom.

Method

Inclusion criteria

Three criteria for the inclusion of articles in the systematic review were adopted. The articles had 1. to be peer reviewed; 2. to be published recently, i.e. after 1 January 2000; and 3. to describe empirical research on the effect of school inspections, regardless of the MSMS-level.

In the aforementioned literature review of Winter and De Ridder (2010) it became clear that research with a quasi-experimental design is fairly unique within educational regulation. Given this result, the third inclusion criterion in

the present systematic review which was added states that the articles have to describe empirical research, regardless of the MSMS-level.

Search strategy

The data was collected between June 2011 and January 2012. Given the fact that education and inspection are studied from different scientific disciplines, different databases were simultaneously searched: ERIC, a database for educational research; PsycINFO and PsycArticles, two psychological databases; SocINDEX, a sociological database; and EconLit, an economic database. The databases were searched using a combination of terms, as presented in Table 1.

**** Table 1 near here.**

The overall result of the searches was 73 unique articles. All titles, authors and keywords of the 73 articles were scanned. When this information suggested a relevant article, the summary of the article was read. Fourteen articles were identified which met the inclusion criteria and which were relevant for the research questions (see Table 2, 3 and 4). Also one relevant literature review, albeit not a systematic review, was found (De Wolf & Janssens, 2007).

The reference sections of the articles were checked to locate additional articles, which however did not result in additional relevant articles. Instead, background information was found in the form of three reports from Ofsted (Matthews & Sammons, 2004; McCrown et al., 2007; Ofsted, 2007) and one dissertation (Ehren, 2006).

Results

In the above mentioned fourteen articles various types of research were discussed. A meaningful distinction concerns the so-called outcome measure by which the effect of school inspections is measured: 1. school improvement; 2. behavioural change of teachers; or 3. student achievement results (cf. De Wolf & Janssens, 2007).³ Arranged by outcome measure, for each article a number of specifications is presented in Tables 2, 3 and 4. Eight of the fourteen articles were concerned with the impact on school improvement (Table 2). Three articles were concerned with the impact on behavioural change of teachers (Table 3). One of these three articles also paid attention to the effect on student achievement results (Matthews & Sammons, 2004). The other three articles were only concerned with the effect on student achievement results (Table 4).

Five of the fourteen articles are based on a case study. These articles, all by the same author, described the same case, each time from a different

perspective (Perryman, 2005; 2006; 2007; 2009; 2010). In four other articles a multiple case study is presented. In two of these case studies only qualitative data was collected (Case, Case & Catling, 2000; Willis, 2010); in the other two both qualitative and quantitative data was collected (Chapman, 2001; Ehren & Visscher, 2008). The results of each of these case studies do not refer to a causal relation, but a supposed connection between school inspections and the quality of education.

Of the five fully quantitative studies three are concerned with a quasi-experimental design in which a causal effect is measured (Shaw et al., 2003; Rosenthal, 2004; Luginbuhl et al., 2009). The other two, the survey of Dederling and Muller (2011) and the study of Matthews and Sammons (2004) which consisted of the secondary analysis of three sets of data, are identified as studies at MSMS-1 level. Once more, the results of these two studies do not refer to a causal relation, but to a supposed connection between school inspections and the quality of education.

The case studies of Perryman (2005; 2006; 2007; 2009; 2010) and Willis (2010), and also in the research of Matthews and Sammons (2005), focus on unsatisfactory or weak schools under special measures of Ofsted⁴. The other five articles pay no special attention to this group of schools. Notably, eight articles date from at least five years back, and the three studies with a quasi-experimental design are based on data at least six years old (Shaw et al., 2003; Rosenthal, 2004; Luginbuhl et al., 2009).

Finally, in two articles the Dutch situation is studied (Luginbuhl et al., 2009; Ehren & Visscher, 2008). The most recent article is a study from Germany (Dedering & Muller, 2011). The other eleven articles are concerned with research that was conducted in the United Kingdom, and thus with the regulation of Ofsted. Four of the fourteen studies were conducted in primary education (Case et al., 2000; Ehren & Visscher, 2008; Luginbuhl et al., 2009; Willis, 2010). One article covered research in various educational sectors (Dedering & Muller, 2011) and the other nine articles describe research in secondary education.

**** Table 2 near here.**

School improvement

The specifications of the articles concerning the effect on school improvement are presented in Table 2. Central in the study of Ehren and Visscher (2008) is the relation between school inspections, school characteristics and school improvement. In ten case studies of Dutch primary schools the data was collected in seven different methods in a two year period: before, during and after an inspection visit of the Dutch Inspectorate of Education. The results showed that the inspection of schools only will not automatically improve schools. The interactions of the characteristics of inspections and schools may

or may not lead to school improvement. What appears to make a difference in promoting school improvement are the provision of feedback about weaknesses, the assessment of these weak points as unsatisfactory, and the agreements between an inspector and school regarding improvement activities.

The study of Dederling and Muller (2011) also showed that feedback by the inspector plays an important role. In the survey of 468 principals of primary and secondary schools in Germany, it became clear that school inspections provided the school with knowledge they so far lacked. Around 85 percent of the principals found the verbal feedback at the end of the inspection visit, and also the feedback in the inspection report, useful and meaningful for the development process of the school. A few months after the inspection visit development activities took place at the majority of the schools.

School improvement is also the outcome measure of the qualitative case study in which Perryman followed an English secondary school for four years (2005; 2006; 2007; 2009; 2010). From different perspectives she described the improvement of the school in the last one and a half year the school was under special measures of Ofsted and the three subsequent years. Perryman argued that the special measures regime is an example of 'panoptic performativity':

... an inspection regime in that teachers and pupils feel as if they are constantly being observed, and perform accordingly in order to escape the regime. (2006: 155)

According to Perryman, the question is if there is true long-term school improvement which continues in the years following the period under special measures. A question that one year later was answered positively, when it appears that the school improvement process at the school is continued.

In all five articles by Perryman (2005; 2006; 2007; 2009; 2010) a negative attitude towards Ofsted resonates: intensified monitoring as in the special measures regime leads to schools which learn how 'to perform the good school'. An attitude that can also be seen in the study of Willis (2010) which emphasizes the lack of confidence in Ofsted. In her research Willis described the perspective of three principals of primary schools in the United Kingdom when being placed in special measures by Ofsted. Even Willis finally has to admit that two out of three principals, she herself was the third principal, found that being in special measures was a 'good thing'. It was experienced as a wake-up call to the lack of focus on student achievement results.

**** Table 3 near here.**

Behavioural change of teachers

Table 3 shows the specifications of the articles concerning the effect on behavioural change of teachers. Chapman's study (2001) investigated teachers' perceptions, responses and intentions to change their practice as a

result of an Ofsted inspection in the United Kingdom.⁵ At five secondary schools qualitative and quantitative data was collected after the inspection visit. The results of this study indicated that an appropriate level of high quality feedback from inspectors may be the key to teachers' intention to change classroom practice as a result of the inspection. About twenty percent of the teachers felt that the feedback received from the inspectors prompted them to change aspects of their teaching practice.

The study of Case et al. (2000) showed that an Ofsted inspection achieved just the opposite. The research was concerned with the effect of an inspection visit on the overall wellbeing of teachers. At ten primary schools in the United Kingdom qualitative data was collected in the period running up to, including and the year following the Ofsted inspection. The results showed that the observations of the lessons and the related assessment, including the brief verbal feedback given to teacher and the report to the principal, resulted in a three-year disruption of teaching practice. They concluded that:

...far from improving performance, Ofsted is actually having a detrimental impact on the wellbeing of teachers, the education process and hence the 'qualitative standard' of schooling. (p.618)

It should be noted that throughout the article a negative attitude echoes towards Ofsted (cf. Perryman 2005; 2006; 2007; 2009; 2010; Willis; 2010).

Matthews and Sammons (2005) concluded in their research in the United Kingdom on the development of unsatisfactory secondary schools which

required special measures, a substantial improvement in the quality of teaching as assessed by inspectors of Ofsted. On average there was a decrease of eighty percent in the percentage of unsatisfactory or poor lessons at the time of going into special measures, and the inspection two years after coming out of this regime. The study by Matthews and Sammons (2005) is a small part of the results of a large-scale evaluation of the work of Ofsted about which they reported the year before (Matthews & Sammons, 2004). That evaluation showed that if the researchers looked at the overall group of schools for primary and secondary education something had actually changed in the classroom practice. When Ofsted in 1994/1995 started with the observations, one in five lessons was assessed as unsatisfactory or poor. In 2003/2004 this was only one in twenty lessons. It is debatable, however, whether this difference can be fully attributed to the regulation of Ofsted.

**** Table 4 near here.**

Student achievement results

The specifications of the articles concerning the effect on student achievement results are presented in Table 4. The study of Luginbuhl et al. (2009) considered whether school inspections of the Dutch Inspectorate of Education had effect on the test scores of pupils in primary education. Both approaches

that the researchers used to estimate the effect, indicated positive effects. The effects were quite small and not always statistically significant. Therefore, they concluded that school inspections do no harm, but seem to have little or no effect on student performance.

Similar results are found in two other studies concerning the effect on student achievement results. Shaw et al. (2003) and Rosenthal (2004) conducted research which examined whether an inspection visit carried out by Ofsted, had effect on the examination results of secondary schools in the United Kingdom. Shaw et al. (2003) found that at schools where the achievement of students was already much higher or lower than the average, a slight improvement on examination achievement was noticeable. At those schools with average results the inspection visit had no positive effect: 'If anything, they made it worse' (p.70). Also Rosenthal (2004) found a negative effect on the exam performance in the year of the inspection. In absolute terms the magnitude is small, but nevertheless significantly negative. The explanation is found in the fact that around the inspection period:

...teachers and school administrators might well concentrate on the requirements for success on Ofsted standards and benchmarks rather than their student exam performance. (p.150)

Diametrically opposed to the above are the results of Matthews and Sammons (2005) on the development of unsatisfactory secondary schools in the United Kingdom which required special measures. At these schools the examination

results improved an average of ten percent in the period between the start of the special measures and two years after the special measures had been concluded. As mentioned earlier, the study by Matthews and Sammons (2005) is a small part of the results of large-scale evaluation of the work of Ofsted which they reported in the previous year (Matthews & Sammons, 2004). This evaluation demonstrated that if the researchers looked at the overall group of secondary schools, the examination results in about two-thirds of the schools were higher in the two years after an inspection visit than the results in the year before the visit. Compared with schools that were not visited, the inspected schools were on the average slightly better. The conclusion was that an inspection is neither a catalyst for improving results, nor a significant inhibitor. A conclusion that is similar to the conclusion of Luginbuhl et al. (2009).

Conclusion

The aim of the systematic review was to obtain evidence based insight into the effect of school inspections on the educational quality of schools. In the systematic review fourteen articles were identified which met the inclusion criteria and which were relevant for the research questions. Also one relevant literature review was found. This result shows again that research on the effect within educational regulation is scarce.

The first research question was concerned with the effect school inspections have on the improvement of the educational quality of schools. Overlooking the overall results of the systematic review, no evidence has been found that school inspections automatically lead to the improvement of the educational quality. In the review three outcome measures were distinguished by which the effect of school inspections is measured. It can be concluded that the research designs are the most advanced in those studies that examined the student achievement results, but that the found causal effects of school inspections are small (Shaw et al., 2003; Rosenthal, 2004; Luginbuhl et al., 2009). Less advanced designs are used in research that more directly study the changes on the level of the school or the classroom. Precisely in these studies evidence is found of the effect of school inspections on school improvement and behavioural change of teachers (Chapman, 2001; Matthews & Sammons, 2004; 2005; Ehren & Visscher, 2008; Dederling & Muller, 2011). With this evidence a causal relation cannot be established. However, a supposed connection can be made plausible and explained.

The second research question was as follows: which characteristics of school inspections contribute to the effect on the improvement of the educational quality of schools? Based on the results of the systematic review it can be concluded that none of the characteristics of school inspections in themselves lead to the improvement of educational quality. Research shows that in practice there is a complex interaction between different characteristics of school inspections and inspector on one side, and the school with its pupils,

teachers and management on the other side (cf. Ehren & Visscher, 2008). There is evidence that a positive role is reserved for one aspect of regulative measures, namely feedback: the verbal feedback at the end of the inspection visit and also the written feedback in the inspection report. The feedback provided the school with knowledge they had so far lacked. The provision of feedback about weakness, the assessment of these weak points as unsatisfactory, and the agreements between an inspector and school regarding improvement activities, appeared to make the difference in enhancing school improvement (cf. Ehren en Visscher, 2008; Dederling & Muller, 2011). Evidence for feedback to the teacher after class observation is contradictory. In one study a positive effect is reported while other researchers emphasize the negative aspects (cf. Chapman, 2001; Case et al., 2000).

Discussion and interpretation

Research on the effect of educational regulation is scarce. On the other hand, in the Netherlands as well as in the United Kingdom, there is a decline of schools which were found inadequate (Ofsted, 2011; Inspectie van het Onderwijs, 2012a). The Netherlands⁶ has approximately 7200 primary schools and on 1 September 2011 exactly 3.8 percent of these schools was assessed as an unsatisfactory school with shortcomings in the quality of education, while 0.6 percent was assessed as a weak school. Four years earlier, on 1 January

2008 9.2 percent of primary schools was unsatisfactory and 1.4 percent was considered weak (Inspectie van het Onderwijs, 2012a).

In addition, research shows that in general most schools are satisfied with school inspections (e.g. McCrone et al., 2007; Dederling & Muller, 2011; Inspectie van het Onderwijs, 2012b). For example, Ofsted found in research on the impact of the new inspection regime, which was introduced in 2005 and was called Section 5 inspections⁷, that the vast majority of schools were satisfied with the inspection process. The schools reported that they were implementing all or most of the recommendations for improvement. Most schools took the view that the inspection had contributed to school improvement (McCrone et al., 2007).

Aforementioned examples indicate that the inspection process may have an effect on educational quality, even if this effect has not been established in an evidence based study. However, the result of the systematic review is an initial insight into the effect of school inspections on the educational quality of schools. This insight can only grow and become more evidence based as further research into the effect of school inspections is carried out. An example of such research at the Dutch Inspectorate of Education are two pilots which were recently conducted in order to examine the effect of verbal feedback, given by the inspector to the teacher after class observation (Roijmans, 2010; Dobbelaer, 2011). In the first study it was concluded that the feedback does not lead to behavioural change of teachers (Roijmans, 2010). In the second study it became clear that feedback that is given to stimulate the development

of teachers is appropriate, provided that the inspectors get feedback training (Dobbelaer, 2011).

Other examples of research were presented at the seminar of the European Standing International Conference of Inspectorates (SICI) in November 2011, which was titled 'Impact of Inspection'.⁸ During this seminar ongoing research was presented from Germany, Sweden and Scotland. The first results of follow-up inspections at schools in Lower Saxony in Germany showed for example that the quality of teaching improved nearly to the level of schools that did not fail the first inspection. Whether these improvements have an impact on student achievement results, has not been studied yet. (Homeier & Kluth, 2011). Also during this seminar there was a presentation of the ongoing project on the impact of school inspections on teaching and learning, funded by the European Commission's Lifelong Learning Programme (Education, Audiovisual and Culture Executive Agency, 2010).⁹ This project examines in six European countries what the effects and negative consequences of school inspections are in terms of changes in educational quality of schools and changes in student outcomes; and also what aspects of school inspections contribute to these effects and negative consequences. The project started in January 2011 with the analysis of the program theories of the participating Inspectorates. In December 2013 the project will be finished (Gustafsson, 2011; Ehren et al., sub.).

The above initiatives do not automatically lead to evidence based insight into what school inspections mean and can mean in practice in terms of

improvement in the quality of education. It is important to continue research on all three outcome measures by which the effect of school inspections can be measured. For example, the Dutch Inspectorate of Education has had access to the student achievement results of all primary schools since 2003, the implementation of the Dutch Education Supervision Act.¹⁰ Quasi-experimental research can be established in which can be examined whether there is a comparable improvement in the student achievement results in different types of primary schools: weak, unsatisfactory or basic. Several designs are possible. In research on the other two outcome measures, school improvement and behavioural change of teachers, it appears to be more complicated to establish a causal relation. For example, since 2003 the Dutch Inspectorate of Education has been working with a framework. The framework includes five domains subdivided in indicators which are assessed during school inspections (Inspectie van het Onderwijs, 2011b). The possibility of a research design, with the inspectors' judgments on the indicators of the domains teaching-learning process, special needs provision and guidance, and quality assurance, has to be carefully examined. For each primary school the inspection scores on the indicators are available, for at least two moments in the past nine years. Also the possibility for new research on the effect of school inspections has to be examined. It is important to embed the research within the existing regulation practice where it should be developed to a more evidence based approach. Cooperation between the European Inspectorates is important, as

well as cooperation between Inspectorates and external researchers of scientific research institutions.

Finally, it is clear that the relation between school inspections and the improvement in the quality of education is not a simple one. Nevertheless there is ample evidence that an approach to school inspection that takes into account the findings of some of the discussed studies may have a positive effect on school improvement, behavioural change of teachers or student achievement results. Of course in the end it should not be forgotten that regulation can only act as a catalyst and that schools themselves must bring about the actual improvements.

Notes

¹ It should be noted that it is emphatically not our intention to compare the educational regulation or the Inspectorates of Education in different countries.

² The fourth study is a study into the effects of social skill interventions on aggressive and disturbed behavior in school children (study 37 of Winter & De Ridder, 2010).

³ De Wolf and Janssens (2007) distinguish a fourth outcome measure, i.e. the effect of school inspections on the satisfaction of schools and school leaders. Given the aim of the present systematic review which deals with the effect on the quality of education, research on this outcome measure is excluded from the review.

⁴ In the United Kingdom, according to the Education Act 2005, schools requiring special measures are 'Schools which are failing to give pupils an acceptable standard of education and where the persons responsible for leading, managing or governing the school are not demonstrating the capacity to secure the necessary improvement.' (Ofsted, 2007: 15).

⁵ Chapman (2001) frequently refers in the article to the article underlying MA dissertation from 2000. However, since it is an unpublished dissertation and not available for inspection.

⁶ In the Netherlands a primary school is an unsatisfactory school if the school has shortcomings in the quality of education, only because of insufficient student achievement results. A primary school is weak if the school has shortcomings in the quality of education, because of insufficient student achievement results and an inadequate teaching and learning process (cf. Inspectie van het Onderwijs, 2011b; 2011c).

⁷ The main elements of Section 5 inspections include: 'shorter notice of inspection, smaller inspection teams, more frequent inspections, an increased emphasis on the school's own self-evaluation evidence, and shorter reports with fewer, clearer recommendations for improvements.' (McCrone et al., 2007: iii).

⁸ <http://www.sici-inspectorates.eu/>

⁹ http://ec.europa.eu/education/lifelong-learning-programme/doc78_en.htm

¹⁰ http://wetten.overheid.nl/BWBR0013800/geldigheidsdatum_07-04-2011

References

- Case, P., Case, S. & Catling, S. (2000). Please Show You're Working: a critical assessment of the impact of OFSTED inspection on primary teachers. *British Journal of Sociology of Education*, 21(4), 605-621.
- Chapman, C. (2001). Changing Classrooms Through Inspections. *School Leadership and Management*, 21(1), 59-73.
- Dedering, K. & Muller, S. (2011). School Improvement through Inspections? First Empirical Insights from Germany. *Journal of Educational Change*, 12(3), 301-322.
- De Winter, H.B. (2010). *Zicht op toezicht? Over de meerwaarde van toezicht in de risicosamenleving (oratie) [View of regulation? On the added value of regulation in the risk society (inaugural speech)]*. Groningen, The Netherlands: University of Groningen.
- De Wolf, I.F. & Janssens, F.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.
- De Wolf, I.F. & Verkroost, J.H. (2011). Evaluatie van de theorie en praktijk van het nieuwe onderwijstoezicht [Evaluation of the theory and practice of new approach of regulation in education]. *Tijdschrift voor Toezicht*, 2(2), 7-24.
- Dobbelaer, M.J. (2011). Mondelinge Feedback door Inspecteurs aan Leraren Primair Onderwijs: Kwaliteit, Perceptie, Effect, en Feedbacktraining (Masterthesis) [Oral Feedback by inspectors to teachers in primary

- education: Quality, Perception, Effect, and Feedback training (Master thesis)]. Online: <http://igitur-archive.library.uu.nl/student-theses/2011-0729-200530/Masterthesis%20Dobbelaer,%20MJ-3458210.pdf> [assessed January 2012].
- Education, Audiovisual and Culture Executive Agency (2010). *Compendium 2010, Lifelong Learning Programme, Key Activity 1, Policy cooperation and innovation, Studies and Comparative Research (SCR)*. Online: http://eacea.ec.europa.eu/llp/results_projects/documents/compendium_ka1_scr_2010.pdf [assessed June 2011].
- Ehren, M.C.M. (2006). *Toezicht en schoolverbetering [School inspections and school improvement]*. Delft: Eburon.
- Ehren, M.C.M. (2010). Effecten van toezicht op het basisonderwijs [Effects of school inspections on primary schools]. *Pedagogische studiën*, 87(3), 165-182.
- Ehren, M.C.M., Altrichter, H., McNamara G. & O'Hara, J. (submitted). Impact of school inspections and learning: Describing assumptions on causal mechanism in six European countries. *Educational Assessment, Evaluation and Accountability*. Retrieved from authors in January 2012.
- Ehren, M.C.M. & Visscher, A.J. (2008). The Relationships between School Inspections, School Characteristics and School Improvement. *British Journal of Educational Studies*, 56(2), 205-227.
- Hemingway, P. & Brereton, N. (2009). What is a systematic review? Online: <http://www.whatisseries.co.uk/whatis/> [assessed June 2011].

Homeier, W. & Kluth, J. (2011). *The impact of Inspection. Lower Saxony.*

Presentation on the SICI workshop from 23-25 November 2012 in Stockholm, Sweden. Online: <http://www.sici-inspectorates.eu/en/87600-impact-of-inspection-stockholm-sweden> [assessed January 2012].

Farrington, D.P., Gottfredson, D.D., Sherman, L.W. & Welsh, B.C. (2002). The Maryland Scientific Methods Scale. In L. W. Sherman, D. P. Farrington, B. C. Welsh, & D. L. MacKenzie (Eds.), *Evidence-based crime prevention* (pp. 13-21). London, England: Routledge.

Gustafsson, J. (2011). Impact of school inspections on teaching and learning. An ongoing project. Presentation on the SICI workshop from 23-25 November 2012 in Stockholm, Sweden. Online: <http://www.sici-inspectorates.eu/en/87600-impact-of-inspection-stockholm-sweden> [assessed January 2012].

Inspectie van het Onderwijs (2005). *Toezichtkader primair onderwijs 2005* [Inspection framework primary education 2005]. Utrecht, The Netherlands: Inspectorate of Education.

Inspectie van het Onderwijs (2010). *Zeer zwakke scholen in het basisonderwijs* [Weak primary schools]. Utrecht, The Netherlands: Inspectorate of Education.

Inspectie van het Onderwijs (2011a). *De staat van het onderwijs. Onderwijsverslag 2009/2010* [The state of education. Education Report 2009/2010]. Utrecht, The Netherlands: Inspectorate of Education.

- Inspectie van het Onderwijs (2011b). Toezichtkader 2011. Primair onderwijs, voortgezet onderwijs [Inspection framework 2011. Primary education, secondary education]. Utrecht, The Netherlands: Inspectorate of Education.
- Inspectie van het Onderwijs (2011c). *Toezicht op zwakke en zeer zwakke scholen. Zo werkt het [Regulation on unsatisfactory or weak schools. How it works]*. Utrecht, The Netherlands: Inspectorate of Education.
- Inspectie van het Onderwijs (2012a). De staat van het onderwijs. Onderwijsverslag 2010/2011 [The state of education. Education Report 2010/2011]. Utrecht, The Netherlands: Inspectorate of Education.
- Inspectie van het Onderwijs (2012b). Tevredenheidsonderzoek sectoroverstijgend. Periode oktober 2010 - september 2011 (versie 0.1) [Satisfaction survey multisectoral. Period October 2010 - September 2011 (version 0.1)]. Internal report of the Dutch Inspectorate of Education.
- Jansen, M. & Smits, A. (2011). *Effecten van toezicht en Handhaving meten. Een handreiking [Measuring the effects of regulation and enforcement. A guideline]*. Utrecht, The Netherlands: CCV.
- Luginbuhl, R., Webbink, D. & De Wolf, I.F. (2009). Do School Inspections Improve Primary School Performance? *Educational Evaluation and Policy Analysis*, 31(3), 221-237.
- Matthews, P. & Sammons, P. (2004) *Improvement through inspection. An evaluation of the impact of Ofsted's work*. Londen, Engeland: Ofsted.

- Matthews, P., & Sammons, P. (2005). Survival of the Weakest: The Differential Improvement of Schools Causing Concern in England. *London Review of Education, 3*(2), 159-176.
- McCrone, T., Rudd, P., Blenkinsop, S., Wade, P., Rutt, S. & Yeshanew, T. (2007) *Impact of Section 5 inspections: maintained schools in England: Final report*. Online: <http://www.ofsted.gov.uk/resources/impact-reports-2007> [assessed December 2011].
- Ofsted (2007). Review of the impact of inspection (Reference no: 070042). Online: <http://www.ofsted.gov.uk/resources/impact-reports-2007> [assessed December 2011].
- Ofsted (2011). The Annual Report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2010/11. Online: <http://www.ofsted.gov.uk/resources/annualreport1011> [assessed December 2011].
- Perryman, J. (2005). School Leadership and Management after Special Measures: Discipline without the Gaze? *School Leadership and Management, 25*(3), 281-297.
- Perryman, J. (2006). Panoptic Performativity and School Inspection Regimes: Disciplinary Mechanisms and Life under Special Measures. *Journal of Education Policy, 21*(2), 147-161.
- Perryman, J. (2007). Inspection and Emotion. *Cambridge Journal of Education, 37*(2), 173-190.

- Perryman, J. (2009). Inspection and the Fabrication of Professional and Performative Processes. *Journal of Education Policy*, 24(5), 611-631.
- Perryman, J. (2010). Improvement after Inspection. *Improving Schools*, 13(2), 182-196.
- Rosenthal, L. (2004). Do School Inspections Improve School Quality? Ofsted Inspections and School Examination Results in the UK. *Economics of Education Review*, 23(2), 143-151.
- Roijmans, E.A.J. (2010). *Effecten van mondelinge feedback door inspecteurs aan leerkrachten in het primair onderwijs (Masterthesis) [Effects of oral feedback by inspectors to teachers in primary education (Master thesis)]*.
Online: http://inspectienet.owinsp.nl/kennis/pdf/thesis_Eveline_Roijmans
[assessed January 2012 from the intranet of the Dutch Inspectorate of Education].
- Shaw, I.I., Newton, D.P., Aitkan, M.M. & Darnell, R.R. (2003). Do OFSTED Inspections of Secondary Schools Make a Difference to GCSE Results? *British Educational Research Journal*, 29(1), 64-75.
- Smeets, G. & Verkroost, J. (2011). *Selectief en slagvaardig. Werken met de WOT (2000-2010) [Selective and decisive. Working with the Dutch Educational Supervision Act (2000-2010)]*. Utrecht, The Netherlands: Inspectorate of Education.
- Willis, L. (2010). Is the Process of Special Measures an Effective Tool for Bringing about Authentic School Improvement? *Management in Education*, 24(4), 142-148.

Winter, H.B. & De Ridder, J. (2010). *'Toezicht en effectmeting. Het kán!'*
Researchsynthese van onderzoek naar de effectiviteit van
toezichtsinterventies ['Regulation and effect. It can!'] Research synthesis of
research on the effectiveness of regulation]. Groningen, The Netherlands:
Faculty of Law, Department of Administrative Law and Public Administration
and Pro Facto.

Table 1

Terms used to search the databases

| No. | Terms | Results | | |
|-----|--|-------------|--------------------|---------------------------|
| | | Found | New ^A . | Selected |
| 1. | School Inspection and Improvement | 27 articles | 27 articles | 8 articles ^B . |
| 2. | School Inspection and [Effect or Effects or Effective or Impact] | 58 articles | 39 articles | 6 articles 1 review |
| 3. | School Inspection ^C and Monitoring Improvements | 0 articles | 0 articles | 0 articles |
| 4. | School Inspection and Intensified | 1 article | 0 articles | 0 articles |
| 5. | School Inspection and [External Control or Quality Control or Compliance Assistance] | 10 articles | 3 articles | 0 articles |
| 6. | School Inspection and [Special Measures or Poor Performances or Failing Schools or Underperforming Schools] | 5 articles | 0 articles | 0 articles |
| 7. | School Inspection and [Quality of Education or Educational Quality or School Performances or Student Performances] | 23 articles | 4 articles | 0 articles |

^A. New stands for articles not found in previous searches.

^B. One article was also published in Dutch. In the review is chosen to use the English article (Ehren, 2010; Ehren & Visscher, 2008).

^C. Also the term Inspection did not yield new articles in combination with the other terms.

Table 2

Specifications of the articles concerning the effect on school improvement

| Specification | Article 1 | Article 2 | Article 3 | Article 4 |
|------------------|---|------------------------------|---|---|
| Title | School Improvement through Inspections? First Empirical Insights from Germany | Improvement after Inspection | Is the Process of Special Measures an Effective Tool for Bringing about Authentic School Improvement? | Inspection and the Fabrication of Professional and Performative Processes |
| Authors | Dederling, K. and Muller, S. | Perryman, J. | Willis, L. | Perryman, J. |
| Journal | Journal of Educational Change | Improving Schools | Management in Education | Journal of Education Policy |
| Publication year | 2011 | 2010 | 2010 | 2009 |
| Country | Germany, North Rhine-Westphalia | United Kingdom | United Kingdom | United Kingdom |
| Sector | Primary and secondary education ^A . | Secondary education | Primary education | Secondary education |
| MSMS-level | MSMS-1 | - | - | - |
| Design | Survey | Case study | Multiple case study | Case study |
| Data | Quantitative | Qualitative | Qualitative | Qualitative |
| Research period | 2008 | 1999-2003 | Not explicitly mentioned | 1999-2003 |

^A. Including special and vocational education.

| Specification | Article 5 | Article 6 | Article 7 | Article 8 |
|------------------|---|--------------------------------|--|---|
| Title | The Relationships between School Inspections, School Characteristics and School Improvement | Inspection and Emotion | Panoptic Performativity and School Inspection Regimes: Disciplinary Mechanisms and Life under Special Measures | School Leadership and Management after Special Measures: Discipline without the Gaze? |
| Authors | Ehren, M.M. and Visscher, A.J. | Perryman, J. | Perryman, J. | Perryman, J. |
| Journal | British Journal of Educational Studies | Cambridge Journal of Education | Journal of Education Policy | School Leadership and Management |
| Publication year | 2008 | 2007 | 2006 | 2005 |
| Country | The Netherlands | United Kingdom | United Kingdom | United Kingdom |
| Sector | Primary education | Secondary education | Secondary education | Secondary education |
| MSMS-level | MSMS-1 | - | - | - |
| Design | Multiple case study | Case study | Case study | Case study |
| Data | Quantitative and qualitative | Qualitative | Qualitative | Qualitative |
| Research period | 2002-2006 | 1999-2003 | 1999-2003 | 1999-2003 |

Table 3

Specifications of the articles concerning the effect on behavioural change of teachers

| Specification | Article 9 | Article 10 | Article 11 |
|------------------|---|--|---|
| Title | Changing Classrooms Through Inspections | Please Show You're Working: a critical assessment of the impact of OFSTED inspection on primary teachers | Survival of the Weakest: The Differential Improvement of Schools Causing Concern in England |
| Authors | Chapman, C. | Case, P., Case, S. and Catling, S. | Matthews, P. and Sammons, P. |
| Journal | School Leadership and Management | British Journal of Sociology of Education | London Review of Education |
| Publication year | 2001 | 2000 | 2005 |
| Country | United Kingdom | United Kingdom | United Kingdom |
| Sector | Secondary education | Primary education | Secondary education |
| MSMS-level | MSMS-1 | - | MSMS-1 |
| Design | Multiple case study | Case study | Secondary analysis |
| Data | Quantitative and qualitative | Qualitative | Quantitative |
| Research period | 1999-2000 | Not explicitly mentioned | 1993-2004 |

Table 4Specifications of the articles concerning the effect on student achievement results^A.

| Specification | Article 12 | Article 13 | Article 14 |
|------------------|---|---|---|
| Title | Do School Inspections Improve Primary School Performance? | Do School Inspections Improve School Quality? Ofsted Inspections and School Examination Results in the UK | Do OFSTED Inspections of Secondary Schools Make a Difference to GCSE Results? |
| Authors | Luginbuhl, R., Webbink, D. and De Wolf, I.F. | Rosenthal, L. | Shaw, I.I., Newton, D.P., Aitkan, M.M. and Darnell, R.R. |
| Journal | Educational Evaluation and Policy Analysis | Economics of Education Review | British Educational Research Journal |
| Publication year | 2009 | 2004 | 2003 |
| Country | The Netherlands | United Kingdom | United Kingdom |
| Sector | Primary education | Secondary education | Secondary education |
| MSMS-level | MSMS-3 | MSMS-3 | MSMS-3 |
| Design | Quasi-experimental design | Quasi-experimental design | Quasi-experimental design |
| Data | Quantitative | Quantitative | Quantitative |
| Research period | 1999-2003 | 1993-1997 | 1992-1997 |

^A. In the article by Matthews and Sammons (2004) which is shown in Table 3, attention is also paid to the effect on student achievement results