Inspections of Swedish schools: A critical reflection on intended effects, causal mechanisms and methods

Jan-Eric Gustafsson*, Rolf Lander** & Eva Myrberg***

Abstract
The article outlines a programme theory for the Swedish school inspection. The theory has a format of ‘if... then... because’, whereby the last term states one or more generative mechanisms behind the reactions to inspection, and the former terms imply what the inspectorate does and which reactions it receives. The assumptions of the theory are tested regarding their precision of definition, consistency and empirical status. No research has as yet confirmed a general positive effect of the Swedish inspection on learning and school development. Programme theory, however, suggests that such effects are mostly context-dependent, and thus will vary between schools and school authorities. A complicating feature of the present inspection is its objectivist ethos, which is at odds with the Swedish tradition of a transactional ethos in inspections.

Keywords: school inspection, quality audit, regular supervision, programme theory, generative mechanisms, transactionalism

A programme theory is suggested here for two activities of the present Swedish school inspection (SSI): the regular supervisions of all schools, and the thematic quality audits of samples of schools. We regard these as holding the greatest relevance for the operations of inspected pre-schools, primary and secondary schools, and adult education.

We attempt to interpret what we believe are necessary or likely rationalist assumptions underpinning the inspection task. This is a social science approach to programme theory (Chen, 1990). We take as a starting point the main formulations of programme theory and reformulate these arguments as a step in what Pawson and Tilley (1997, 218) call “the wider cycle of ‘enlightenment’ between research and policy fields”. We also try to test the assumptions of the inferred programme theory of the Swedish school inspections.
Another intention is to discuss the ethos of the inspection, which is at odds with the historical tradition of transactionalism. By transactionalism we mean an approach to evaluation which favours an exchange between different perspectives on what is true or best (House, 1981). Inspection is commonly identified as an expert approach, but a transactional mode may also go along with it. This calls for a brief historical background.

**Background: Structural reform and inspection ethos**

Three big structural school changes involving more powerful inspectorates have been made in modern times.

For more than 300 years, the state and church had a “master-servant relation” to primary and secondary education (Isling 1988, 486), but in 1862 municipalities were separated from the state church parishes. A state school inspection was installed in 1858 to force local taxpayers to take their responsibilities for the obligatory primary education more seriously. After 1871 inspectors could veto state subsidies for municipalities that did not follow state regulations (du Rietz et al. 1987). However, inspectors also gave advice to municipalities and teachers (Thelin 1994). Inspection was welcomed by the primary teachers’ union. At the beginning of the 20th century, municipalities became more positive about schooling and many cities set up their own inspectorates (Nytell 2006). At the same time, central bureaucracies were implemented for secondary and primary education. Tight financial and administrative state rule-government was successively put into place, causing the state inspection to diminish in importance.

The second big structural change came with the introduction of the comprehensive school in 1962. Twenty-five regional state boards were added to the National Board of Education (NBE) in 1958. They supported schools during the experimentation phase before the reform, and inspected its later implementation. One task was to prevent informal ability grouping. The inspectorate’s sanctions for rule-breaking were, however, very rarely used. Boards primarily recruited their members among local politicians, who were thus co-opted, and further municipalities were often strong adherents to the new school. The regional state boards also allocated state subsidies to schools, and hired teachers and head teachers (du Rietz et al. 1987). Between 1962 and 1982 within the NBE there was a transactional and subject-oriented inspection for upper secondary schools, which included advice to teachers about instruction.

During the 1970s and 1980s decentralisation was an important issue for policymakers. The NBE regional boards reactivated the transactional role of school inspections, calling it ‘active supervision’, which focused on helping schools adapt to new political signals about local governing. A new system for economic government was launched in 1978 which promised great local control of the organising of instruction (du Rietz et al. 1987, 273). However, the subsequent outcome was that the
actual effects of the decentralisation policies were small (Lewin 2014, 513). During both decades comprehensiveness continued to be a leading policy in education.

The third big change came in 1991 when a Social-Democratic government made municipalities the primary authorities of schools. This also meant the abolition of state inspection, together with the NBE and its regional boards. Instead, the National Agency of Education (NAE) was established to uphold information governing within the decentralised system.

In 1992 a Conservative-Liberal government introduced a voucher system for both municipal and independent schools. No private fees are allowed in this system, but both small and big companies have found the school market profitable. About 13 percent of students at the compulsory level, and about 25 percent at the secondary level, are enrolled in independent schools. The Swedish school system is now more market-based than most (Rönnberg 2011).

The ‘New public management’ international discourse, favouring governing through contracting, a performance-based economy and competition, has especially been sponsored by the Swedish Association of Local Authorities and Regions (SALAR) which since 1991 has been coordinating the employer function for municipal schools. Market choice has become the new ‘cutting edge’ of the education system, thereby replacing comprehensiveness.

Decentralisation was also an international discourse that included the idea of school-based management (Levačić 1995). Municipalities tried to support schools with their own evaluation systems, and several chose to implement school inspections. Two municipal inspectorates in big cities have been analysed for their rhetorical ethos in trying to make themselves legitimate (Lander and Granström, 2000). One of them used a reasoning ethos (reflecting the views of several parties), and the other a didactical ethos (giving criticism and advice), both transactional in different ways. Inspection reports were also used in transactional ways by head teachers and municipal authorities to communicate with each other and with teachers.

The ambition of the NAE was to support reflection and deliberation through rich and nuanced evaluations (Pettersson and Wallin, 1998). According to Forsberg and Lundahl (2009), such an open-minded evaluation approach was too complicated for many to grasp, and the NAE had to report that most municipalities had failed in using evaluations as hoped for (Nytell 2006). The basic prerequisites for information governance thus met with difficulties.

School choice is the most likely cause of the increased between-school variation in results (NAE 2012). Teachers’ scoring of national tests has also been judged to be too lenient (Skolinspektionen 2012a; Gustafsson and Erickson, 2013). The SSI was therefore given the assignment of commenting on samples of teacher-marked national tests each year. The reliability and validity of public data about results is thus being questioned. This is a problem because the state has put more and more emphasis on management by objectives and results.
A gradual re-centralisation of governing has taken place (NAE 2012). Nytell (2006, Ch. 7) calls the result “a quality regime” including the present state inspection. Re-centralisation is also compatible with messages from international organisations and some research, claiming that decentralisation should be combined with strong central control (Sun, Creemers and de Jong 2007).

Demands from the political opposition and the national accounting authority (RRV) about stricter accountability were heard more intensely. The NAE was reluctant to move from its position that it should not cross municipal borders. However, in 1999 a Social-Democratic government introduced a thematic quality audit, and in 2001 the NAE launched a model called developmental dialogues. It was a deliberate attempt to avoid harsher accountability thinking (Ekholm 2012). A field trial was approved by the government.

NAE field teams engaged with individual municipalities for half a year or more in joint searches for strengths and weaknesses, and brought money to the dialogues in order to stimulate actions. The aim was to integrate organisational change with developmental work directed at classroom learning in a way that supported municipal authority (ASD 2008; Ekholm 2012). “System weak” municipalities, and those with great social problems gained the most from the first rounds of dialogue (Danmarks evalueringsinstitut 2003).

In the end, the model was not approved by the new Social-Democratic minister of education, probably because it was not seen as being sufficiently politically useful for responding to continuing criticism from the opposition. In 2003, the NAE was given the assignment to start ‘full inspections’. This led to a conflict between the minister and the general director of the NAE, which was solved by the director’s change of tasks: to build up a new central bureaucracy, the Agency for School Development (ASD), which for a time continued the developmental dialogues, until it was closed down at the same time as the present SSI was established by a Conservative-Liberal government in 2008.

In the governmental commission report (SOU 2007: 11, 129) which outlined the inspection model of 2008, we can catch a glimpse of transactionalism. The report admits difficulties in reaching a common understanding, but believes that a model of quality can be used if “teachers take part and cooperate in ... (defining) how it should work and what variation ... that is acceptable”. It can be seen as an invitation to engage in negotiations about what quality of instruction means.

This kind of reasoning has, however, not been taken up by the SSI. Instead, the ethos of SSI reports from supervision inspections is of an objectivist kind. The SSI takes on an expert role in relation to school staffs and school authorities. In combination with its choice to report (almost) only negative deviations from the rules and quality criteria, it certainly gives the impression of an inspection with teeth “and the ability to bite hard” (the minister of education cited in Rönberg, Lindblom and Segerholm 2013, 182).
As seen above, school inspection has been used previously to orchestrate power shifts in educational policy. First, to make primary education a secular and more productive part of the education system, and later to defend political ambitions about comprehensiveness. And now it has to handle quality problems together with the dysfunctions of the decentralisation and marketisation policy.

Transactionalism has been a major element in the historical practice of school inspection in Sweden. The roles of the state inspector for the period which ended in 1991 have been characterised as initiator, controller and advisor (Thelin 1994). The dialogue model also fits into this description. But the present state inspection does not.

**Method**

For our programme theory we have chosen to: analyse rational assumptions about school inspection as steps within a policy scientific approach; identify causal mechanisms linked to the goals of the programme, reformulate statements into conditional ‘if … then … because’ propositions; and evaluate the propositions logically and empirically (Ehren, Leeuw and Scherens 2005). This approach is used within a research project about the impact of inspection in six European countries (Ehren 2011), of which our study forms a part.

Causal mechanisms are often confused with programme activities, but Weiss (1997, 46) explains: “The mechanism of change is not the program service per se but the response that the activities generate”. Pawson and Tilley (1997) argue that mechanisms are about choices and reasoning and the capacities people derive from group membership and collective resources. Mechanisms constitute the regularities that are usually upheld in certain situations, so they are said to be generative. Through mechanisms, individuals and groups interpret and are attracted to or repelled by programme goals and activities. The programme has to compete with, suppress or change existing mechanisms in specific contexts. That is why an inspection programme may be predicted to have different effects at different sites.

**Reconstruction sources**

The Swedish Educational Act (SFS 2010, 800) legally regulates school inspection. The government issued an instruction (SFS 2011) and, like for all governmental bodies, there are annual “Letters of Regulation”. Official descriptions of the SSI (Skolinspektionen 2008, 2009a, 2009b, 2010a, 2010b, 2011, 2012b) explain the procedures and expected effects. We studied a sample of specific reports of regular supervision and thematic audits. The Swedish National Financial Management Authority (ESV 2006) outlined a programme theory for the NAE school inspection. We also met with some SSI officers to obtain their views on possible elements of the theory.
The design of regular supervisions

The government decides on objectives, guidelines and allocation of resources for the SSI. The SSI works with interpretations of the national curricular and other steering documents, as well as research-based characteristics of good education and successful schools. The use of research is most obvious in the thematic quality audits, but overall the inspectorate seeks to keep in touch with research. In the letter of regulation for 2014, the government asked the SSI: (a) to increase its analyses of the supervisions’ contribution to the quality of instruction; and (b) to evaluate how its supervision and auditing has so far contributed to school improvement and results. Neither task is possible without referring to scientific theory.

The inspectorate has the right to inspect municipal and independent schools whenever it decides, but it usually follows a five-year cycle. The Education Act from 2011 gave the inspectorate the right to impose sanctions. Some authorities have been threatened with financial penalties, but very few cases make it to the necessary court proceedings. Between 2009 and 2012, about 20 independent schools per year have had their permission (approved by the SSI) to operate schools withdrawn (Skolinspektionen 2012b). Since 2011, teachers have been subjected to a legitimation process by the NAE. If the SSI finds a teacher’s behaviour is not in line with the legitimation regulation, it can report him or her to the Teachers’ Committee of Accountability (Lärarnas ansvarsnämnd), which can issue a warning or withdraw the legitimation. Very few cases of this have been reported.

During 2010, a system of differentiated supervisions was introduced (Skolinspektionen 2010). Before a visit, results are studied, and a survey is completed by the individual schools and the responsible authority. A so-called basic supervision is done in schools which are judged to be well-functioning. If this is uncertain, a ‘widened supervision’ is made. Clear problems call for a ‘deepened supervision’. During 2011 and 2012 about 70 percent of the supervisions were of the widened type.

All schools are visited within a regular supervision when inspecting an authority. School leaders, teachers and school nurses are interviewed, as are pupils and the board members in charge. Some lesson observations are carried out. Two inspectors participate in the school visit, which lasts one or two days, but occasionally longer. Oral feedback is given to the head teacher, together with a preliminary report for the school and its authority to respond to. The formal report is published on the Internet, and also presented to the media via a press release.

The design of thematic quality audits

Every thematic quality audit focusses on a particular issue (e.g. leadership, a subject). A relatively small number of schools is randomly sampled for every audit. A programme theory for the issue is constructed together with a reference group with researchers and other experts. Two inspectors visit each school. Time limits are less fixed than for supervisions. The reference group offers suggestions for the final report.
Feedback to schools happens in the same way as for supervisions, but a summary report is also published and disseminated through the media, conferences and seminars. Between six and seventeen thematic audits have been made each year over the last four years.

Audits aim to establish a knowledge basis which supports improvement of the quality of all schools in Sweden. Since 2011, the SSI has intensified its efforts to obtain a close view on instruction and the learning process, and a model for subject audits has been developed.

**Programme theory reconstructed**

With some modifications, we follow the strategy used by Ehren, Leeuw and Scheerens (2005) who analyse a programme theory of the Dutch inspection. Like them, we match several analogous arguments from the sources to form single assumptions, and then put them in a logical format. They give explicit references in footnotes to the sources of each statement. We do not. However, the sources mentioned are easy to find by anyone who wants to scrutinise our analysis.

Ehren et al. (2005) define mechanisms as the factors that “are expected to solve the problem” and that “drive the policies or program and are believed to make them effective” (62). With this, mechanisms look like programme activities, which they should not do according to Astbury and Leeuw (2010), among others. However, Ehren et al. treat mechanisms as the combined effect of programme activities and reactions among stakeholders. For example, the inspectorate’s framework of quality “will be found both realistic and relevant by schools. Because schools will support a framework that is developed in close cooperation with umbrella organizations in the educational field which ensures that the framework is feasible for schools” (table 1, 67). The mentioning of realistic and relevant is an implied then-thesis, which is explained by the because-thesis. There are only four because-theses compared to fourteen if-theses about quality. Three of the because-theses are about the legitimacy just mentioned. We think that a programme theory becomes clearer when theses about if, then, and because are expressed in each statement. As a consequence, we have relatively many more mechanisms in our theory compared to that of Ehren et al.

In our 14 statements below, mechanisms (because-theses) in statements 1–3, 10 and 14 refer to the inspectorate while statements 4–9 and 11–13 refer to schools. The mechanisms for schools refer to motivation (e.g., are eager, welcome, feel rewarded, stimulates morale, feels threatened in statements 4–8, 12), to capacity, expectations and habits (in statements 9, 13) and to relations with clients (6, 11).

A statement can refer to supervision inspection or to thematic audit and supervision inspection. If these forms of inspection are not explicitly mentioned, the statement speaks about both. No one refers to audits only, but of course this is possible (as well as more detailed statements about supervision). We offer a theory as a general overview, well aware of the need for specifications in more in-depth
studies. Such studies are likely to reveal mechanisms that are counteracting the inspectorate’s efforts. They are not included in the theory here due to a lack of solid evidence, but are mentioned in the discussion.

1. If rules, regulations and quality criteria for schooling are precisely communicated from parliament and the government, then the inspectorate will be able to interpret which local behaviours will harm or stimulate pupils’ learning and school staffs’ professional work, because inspectors will trust that such rules, regulations and quality criteria are validly based on good and commonly accepted scientific evidence and professional experience.

2. If interviews and observations during visits, as well as analyses of documents and statistical information, are timely performed, then the strengths and weaknesses of the inspected school will be validly interpreted and reported by the inspectors, because inspectors have good methods and training for such tasks.

Reading yearly reports and other publications from the SSI, one obtains a picture of an agency that is eager to uphold internal discussions and training. It gives assignments to researchers to investigate relevant issues of quality and improvement. It is also a member of the Standing International Conference of Inspectorates. This forms the basis for the third statement:

3. If the inspection staff continuously engages in training and reflection over issues about school governing, quality, and methods for inspection, as well as interaction on these matters with researchers and other inspectorates, then they will be able to form an epistemic culture, because the recruitment of officers, the spirit of the organisation, and the economic frames make these urgent matters for the inspection’s leadership.

Let us assume that the authority and the school have the same capacity for receiving and using feedback and are supporting each other so that for the time being we can omit the authority from the theoretical considerations here.

4. If schools in advance and on request report on their own results and processes, and the supervision complements them with statistics and other public reports, then the inspectorate can validly decide which schools should have basic and widened regular supervisions, because schools are well informed about what is needed for the inspectorate’s decision, and are eager to belong to the right category.

5. If the inspection gives clear and trustworthy feedback about its results, then schools will feel obliged to work with them within given time frames, because schools welcome impartial feedback and advice.

6. If the inspectorate makes findings from the supervision of individual schools public, then designated schools will improve, because they want to keep up a responsible attitude towards the public, and their present and potential clients.

7. If schools receive basic supervision, then they will continue to improve their good work, because such schools will feel rewarded, and their professional collective efficacy will be improved.

8. If schools receive widened supervisions, then the extra time devoted to them by the inspectors are enough to convince them about the need to change, because the feedback and advice will be sufficiently instructive and supportive to keep staff morale on a high level.
9. If schools during supervision receive negative feedback on certain aspects of their work and results, then they will improve considerably within given time frames, because they will generally find the suggestions valid and also have the capacity to organise their quality and school development processes in line with them.

10. If schools do not have the needed capacity for school development in the face of external critique, then they will learn about the conditions to obtain such capacity from the supervision process, because the inspectorate pays considerable attention to developmental processes and has valid knowledge about this to disseminate to schools.

11. If schools are receiving heavy criticism during supervisions, then their clients (pupils and their parents) will not abandon them, because schools generally have established good relations with clients in order to have their loyalty for at least the time needed for improvement.

12. If schools receive renewed heavy criticism for not having improved after supervision, then schools will generally comply, because they fear the consequences of penalties or other sanctions the inspectorate may impose.

13. If the inspection recurrently exerts pressure on all schools to not deviate from rules and regulations and good quality, then school results will increase generally, because expectations among schools and their clientele will institutionalise a proactive habit of continuous improvement in line with the inspectorate’s criteria.

14. If the inspectorate publishes results from thematic audits as well as regular supervisions, then it will have an impact on public, professional and political opinion, because the inspectorate has the capacity to write good reports, and support them with seminars and other arenas for debate and information.

**Evaluation**

We now turn to an evaluation of whether the programme theory assumptions are valid.

**Precision of definitions**

Operations of the SSI are still being developed so it is natural that the documentation of procedures is lagging behind. The definition of terms is thus often implicit, and the interpretations made in the analysis of the SSI’s programme theory may be different to those intended.

The term “regular supervision” is quite clearly defined, probably because it has a relatively long history. However, the separation between the supervision of legal aspects and quality aspects needs further clarification. The newly introduced “differentiated school supervision” is briefly characterised by the SSI (2011), but for the distinctions between “basic”, “widened”, and “deepened” school supervisions clear descriptions have not yet been offered. Particularly in comparison with the description of the Dutch system for identifying potentially failing schools (see Ehren 2011), the lack of precision is striking. The definition of “strong” and “weak” schools also lacks in precision for the same reason.

“Good education” is defined in terms of all students reaching the educational goals. While this definition is clear, it might also be regarded as narrow. It could
mean that all students should reach the level of “Pass” in the criterion-referenced grading system. But that would imply a ceiling on the achievements of many students. Further, “good education” also implies good social development, which is not measured by grades, if it is measured at all.

Thematic quality audits involve definitions based on reviews of research. However, given the complexity of the issues under investigation and the enormous amount of research available, it is a difficult task to arrive at precise, research-based definitions of all pertinent concepts. The fact that there always exist several different research positions also implies great difficulties in arriving at unequivocal definitions of the key concepts.

Consistency
If a school does not agree with how the SSI interprets theory or professional norms, or the results of the supervision, a conflict with the freedom of decision making which is granted to schools may arise. The problem does not occur for thematic audits. This possible inconsistency between different parts of the governing is not discussed very much in policy documents.

Empirical value
Do the assumptions of the programme theory align with the results from the research?

Taken together, all statements assume that inspections will lead to improved performance. There is little Swedish research addressing this issue. However, Ekholm and Lindvall (2008) followed added value-adjusted marks in a time series five to six years before the inspection of all the inspected schools in 2003–2004, and three to four years after. Marks improved in 29 percent of the schools, were unaffected in 25 percent, and declined in 46 percent.

The United Kingdom and the Netherlands may be said to have the most experienced inspectorates in Europe. Their results are nevertheless ambiguous. In the Netherlands, Luginbuhl et al. (2009) found that test scores of pupils in primary education improved slightly in the two years following an inspection visit. In contrast, studies conducted in England by Rosenthal (2004) and Shaw et al. (2003) demonstrated weak negative effects of inspections in comprehensive schools, although Shaw et al. also found positive effects of school inspections in schools with a selective intake. Hussain (2012) found positive effects of inspection on the results of failing English schools, which he claims is not due to the statistical regression effect. These results indicate that the effects of inspections cannot easily be generalised across school systems and types of inspections, which is exactly what a programme theory approach would predict. It is still quite problematic that “very little hard empirical evidence exists on the effects of inspection systems” (Hussain 2012, 2).
The main assumption in statement 1 is that deviations from rules and regulations in a school’s activities cause negative effects on the learning, development or safety of the students. However, only a few of a school’s activities are governed by rules from which deviations are easy to identify. Further, rules and regulations are not necessarily based on scientific or professional knowledge but, instead, the results of ideological standpoints and compromises.

Statements 2 and 3 assume that inspectors have good training and valid methods for evaluating schools and, furthermore, that they engage in building an epistemic culture. Such cultures are typical of science, but they also exist within “expert processes and expert systems that are epitomized by science . . . [in] . . . all areas of social life” (Knorr Cetina 1999, 1). Epistemic praxis focuses on new knowledge and often tries to complicate matters. This is different from habitual praxis, which tries to make its objects of knowledge uniform and simple. An example shows how the SSI is drawn between such positions. In discussing teachers’ marking of national tests, the inspectorate praises essay writing for its validity, but prefers simpler tasks because of their higher reliability (Skolinspektionen 2012a, b).

Possibilities for epistemic curiosity must be highest within thematic audits. However, the research ambitions within thematic audits are largely realised through methods used in regular supervision. The analyses in summary reports generally show low ambitions, without inferences being drawn about the theoretical implications of the findings. In contrast to research, there is no process of peer review, and no feedback of the results to the research community (Gustafsson and Erickson, 2013). It is highly questionable whether these audits can manage the dual tasks of both inspecting schools for quality and creating more generally applicable knowledge.

Statement 4 presumes that a valid choice between the supervision modes can be made using public statistics and other information, including questionnaires given to the inspected school. One such piece of information is value-added measures of results (grade 9 marks for schools adjusted with regard to sex, parental education, and country of birth among pupils) which is published annually for all comprehensive schools by the NAE. Such measures are used internationally in order to make indicators of school quality more just, but in fact the measures’ validity is quite low (Hansen and Lander, 2009). Further, increasing segregation has been noticed in that pupils with similar study motivation, regardless of their socio-economic background, gather at the same schools (NAE, 2012). Such processes cannot be adjusted by the value-added measure used, but still influence teaching conditions. Considering this, it may be safe for the inspectorate to select 70 percent of schools for widened supervision.

Generally, stakeholders’ positive attitudes to inspection are reported from several countries (de Wolf and Janssens 2007), including Swedish head teachers (Skolinspektionen 2012b). Still, trust is a contested issue. Husfeldt (2011, 278) thinks that responses about inspection can be heavily affected by social desirability. De Wolf
and Janssens (2007) found many studies reporting on “window dressing”, and that teachers show sub-optimal “proceduralism” or “tunnel vision” during the inspectors’ visits. Hargreaves (1995) posits that inspection scares teachers not to demonstrate innovative ideas. Such data tell us about unease and anxiety, but of course feelings depend on what school staff learns about effects for themselves in the long run. Statements 5, 8, 9 and 10 assume such learning from feedback to be positive, but statement 12 acknowledges that compliance may only be achieved by threat. If there is a naming-and-shaming component of the feedback to local stakeholders, statements 11 and 12 are connected to this.

However, in a questionnaire study of head teachers within six European countries (including Sweden) the perceived quality of inspection feedback had no relationship to improvement actions. Instead, a most important factor was how sensitive to the inspection report parent and pupil representatives and the school authority were felt to be (an issue taken up in statement 6). It thus seems like the need for legitimacy among such groups was more important for improvement actions taken than the quality of the inspection feedback. This is not according to the rationalistic logic mainly assumed for inspection programme theory, but is more likely from a neo-institutionalist theoretical view (Gustafsson et al., submitted). Here we may see how the inspectorates themselves become part of the structural tensions they are intended to supervise.

Statement 8 assumes instructive and supportive feedback from inspectors. The SSI has developed a model for feedback seminars which should be used if the problems pointed to are difficult to change. In this respect, there is a transactional element in the SSI model. Above, we characterised the SSI reports as objectivist in ethos, but we have also heard anecdotal evidence from schools that oral feedback is more lenient and approving. The feedback assumed in statement 10 may be similar to a style found among Dutch inspectors by Ehren and Visscher (2008). With a reserved inspection style, the school receives information about its weak and strong processes, without any elaboration. With a directive style, the inspector also discusses probable causes and potential improvement measures. The oral Swedish supervision style may look like the latter one, but we do not know for sure. In the Netherlands the directive style was neither expected nor wished for.

Ehren and Visscher also state that more complex changes cannot be achieved as results of the inspection in the absence of more frequent visits and feedback. Here it may be constructive to make a comparison with the former Swedish developmental dialogues. An analysis (Sandström, Arvidson and Landahl 2005) showed that the most influential dialogues combined perseverance in time and feedback with a focus on objectives and results and also deep participation from both municipal and school staff.

Statement 9 assumes that schools (and authorities) have the capacity to change following criticism from external sources. In some cases this is an easy task. The SSI
acknowledges that most criticised issues are immediately taken care of, indicating that they are not of a very difficult kind. The SSI (Skolinspektionen 2012b, 34) also reports less optimistic findings. In a questionnaire to ‘target groups’, only 60 percent of the head teachers, 50 percent of leading local politicians, 33 percent of superintendents, and 25 percent of the staff thought that supervisions will promote change “to a high extent”. When we include “to some extent”, about 90 percent think so, suggesting that most think that something, but not enough, will follow due to inspection in serious cases, or that easier cases can be dealt with due to inspection.

Teachers consider national tests as valid against curricula. They are believed to have some impact on teachers’ choice of content and working methods (Erickson and Lander, 2007). The inspectorate’s new task of controlling teachers’ marking of tests may, therefore, also have some effect on teaching, but it is not easy to say of what kind. If it is market mechanisms that induce teachers to overstate test results and marks to improve competitiveness, the problem is probably reoccurring. Anyway, such changes, as well as others that the inspections may actualise, require that schools have a certain habit of continuous improvement, which is taken up in statement 13.

Support structures for ‘failing schools’ in England have long been criticised for being inadequate (Levacic 1995, 195). The Dutch inspectorate was criticised by Ehren, Leeuw and Scheerens (2005) for having an unrealistic programme theory because it neglected the conditions concerning how schools shall make use of external feedback. In Sweden, the ESV (2006) has pointed to the dependence of school improvement after inspection on local support and contextual factors. This problem is strongly related to statements 10/11.

How market mechanisms actually work in relation to quality statements from an inspection is unknown. The logic says that a failing school should close down if reasonable support within a certain time does not help. We do not know what the media and, ultimately, parents think is reasonable support and time limits. It is possible that the inspectorate’s judgement of a school triggers reactions among parents, which leads to withdrawal of voucher resources. This might be a wanted effect, but the risk is that the reaction is premature and ignores the possibilities for improvement in the school.

Statement 14 assumes that the inspectorate will influence public opinion. The English inspection, OFSTED, has been called a “campaigning organization” with an offensive media policy combined with an aggressive inspection style dividing schools into losers and winners (Smith 2000, 342). Rönberg et al. (2013, 194) report from interviews with SSI officers that the inspectorate is highly aware of the possibilities of governing the school system, as well as individual schools, through interchange with the media. The SSI wants to promote democratic accountability, transparency and public debate, but “media usage is also a means to convince and coerce municipalities and schools into doing what is asked of them after the inspections” (184). This is characterised as ‘inspection by spin’, which we normally associate with
political parties. An analysis of the SSI report on the placement of pupils in special education schools identified a romantic rhetorical style which gives the municipalities the roles of villains, the pupils the roles of victims, and the SSI, of course, the role as the hero (Lundgren 2013).

The noted convincing and coercing inspection power related to media usage can be compared with Astbury and Leeuw’s (2010) theoretical discussion of naming and shaming policies in governing. They use quite another example, but mention its relevance for league tables. We find it easy to apply this discussion to school inspection as another theoretical example. It is possible for the inspectorate to attract the media into an established agenda covering inspections’ activities and results, which is likely to be noted by national and local networks. By interpreting and aggregating opinions, they put pressure on schools to comply with inspection messages. This may be an example of the interchange of mechanisms on different structural levels. In Norway, Elstad (2009) studied accountability mechanisms within individual schools exposed to the combination of low results on national tests and media’s naming and shaming publishing. Some schools reacted with panic, without any improvement, and in some schools collegial pressure was triggered leading to better results when internal conditions and external support were good.

Rönnberg et al. (2013) find that the habit of focussing on negative deviations in Swedish supervision reports links very well with a favoured media format. A combined effect on public opinion through the reporting, and an effect on schools’ institutionalised improvement habits “in line with the inspection’s criteria” assumed in statement 13, may be impressive over time.

Discussion
We have raised questions about the precision of the concepts used by the SSI, the consistency of its programme theory, and the empirical support for the effectiveness of inspection. We do not claim that our interpretation of the programme theory – outlined in 14 statements – is the only one possible. It is instead meant as a basis for further empirical inquiries. It is important because inspection is an expensive tool and the meaning of a ‘good education’ is contested. At the same time, governing is more complicated in the present market system than before. Market mechanisms give actors much stronger incentives to tailor regulations to their own advantage, and to protect themselves against external observation. De Wolf and Jansens (2007) describe possible negative consequences of school inspection. In a highly competitive system, strategic behaviour of schools, including window dressing, fraud and misinterpretation, ought to be expected. Schools may implement procedures in order to be assessed more positively, regardless of their effect on children’s learning or social development.

The institutionalisation of the inspection is not given, even if it at present has powerful protection. Pagil (2008) found two pro and contra discourses in responses
to the commission report supplying the basis for the SSI (SOU 2007, 11). On the concurring side were the national school leader association, the pupils’ national organisations, the SALAR and the teacher union organising subject teachers. On the opposite side were two universities, and the teacher union organising pre-school and primary teachers. Two arguments were especially disputed: that strict control is good for development, and that inspections can give valid, comparable and objective information.

It may be noted that three of the actors with positive views of large-scale inspections may have their own reasons for that. The SALAR has aligned with the decentralisation reform agenda and, if inspection can be seen as a defence for this, it is a price municipalities may be willing to pay. The national school leader association also has its own interest in a policy change. According to Jarl, Fredriksson and Persson (2011), the ‘New public management’-concept became a catalyst for the professionalisation of head teachers because it supported their transformation from a narrower teacher-like profession to organisationally defined professionalism. The subject teachers’ union has demanded that schools should once again mostly be governed by the state and may see state inspection as a step in that direction.

The present governing policy is at odds with both the tradition of transactionalism in Sweden and with how decentralisation was first defined. Under a ‘quality regime’ governing is technified and delegated to managers and specialists of administration competing with professional groups and local political rule (Jarl 2012). The dialogue model applied in 2003–2008 first by the NAE, then by the ASD, is the most elaborated attempt to realise transactional thinking in Swedish inspection. However, as discussed by Forsberg and Nordzell (2013), once freely accepted by municipalities it worked with both sticks, carrots and sermons. Municipalities adhered to “absolution through confession” in that they, in order to gain support, had to publicly accept that they had problems (to a great extent defined by the state), and then promise to implement a certain logic of change. Forsberg and Nordzell (2013, 197) discuss possible identity building within the dialogue process where the municipalities are defined as the receivers, and the state agency as the donor. However, as Forsberg and Nordzell say, the dialogue model is not easily characterised, its core is about agreement, negotiating and networking, but a mandatory structure is also embedded. We think the dialogue model should not be idealised because it may show quite strong paternalism but, principally, and as it also seems in practice, dialogue implies greater respect for professionalism and the inherent ideological tensions in education.

Inspection not only promotes certain structural changes but, if institutionalised, it will be a powerful structure in itself. Leeuw and Furubo (2008) suggest that there are five current evaluation systems, among them “the System of (Performance) Audit, Inspection and Oversight”. It has clear connections with ‘New public management’ and a unique feature is the focus on accountability. An evaluation system is characterised
by a distinctive epistemological perspective, organisational responsibility, permanence, and a focus on the intended use of evaluations.

Leew and Furubo (2008) list some warnings about possible malfunctions of evaluation systems: They may produce largely routinised information relevant to day-to-day practices and single-loop learning, but of little relevance for fundamental reassessment and double-loop learning; they give an assurance of procedural safety; they confirm rather than question policies; and they tend to breed evaluation in itself.

The functioning of an evaluation system of the kind pointed to here can be illustrated by the Standing International Conference of Inspectorates. It promotes the idea of self-evaluation as part of inspection, which Lawn and Grek (2012, Chap 9) characterise as a “travelling idea” originating from the Scottish inspectorate, and now supported by the European Commission. Lawn and Grek interpret the concept as intended to be a response to the growing criticism of the regulatory regimes of ‘New public management’ in terms of their costs and “their installation of performative cultures of distrust within organizations” (136). Self-evaluation makes inspections look fresh and somewhat bottom up, though it has been constructed to follow the inspections’ performance criteria and benchmarks, and thus represents no challenge, but a support, to the inspectorates. This may be an example of the tendency of institutionalised organisations to work just as much with their own legitimacy as with the tasks given to them. The interchange between the Swedish inspection and the media may be seen as another such example. But, as Rönberg et al. (2013, 194) observe, there is a lack of “critical mass-media scrutiny of the Inspectorate” in Sweden.
References

ASD. (2008). Utvecklingsdialog för skolutveckling – metod och förhållningsätt [Developmental dialogue – method and attitude]. Temaskrift. Stockholm: Myndigheten för skolutveckling [ASD].

Astbury, B. & Leeuw, F. L. (2010). Unpacking black boxes. Mechanisms and theory building in evaluation. American Journal of Evaluation, 31(3), 363–381.

Danmarks Evaluatoringsinstitut. (2003). Udviklingsdialoger. Strategi, forløb og effektvurdering [Developmental dialogues. Strategy, process and evaluation of effects]. København.

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.

Du Rietz, L., Lundgren, U. P. & Wennås, O. (1987). Ansvarsfördelning och styre på skolområdet. Ett beredningsunderlag utarbetat inom utbildningsdepartementet [Distribution of responsibility and governing in schooling]. DsU 1987:1. Stockholm: Utbildningsdepartementet.

Chen, H. (1990). Issues in constructing program theory. In Advances in Program Theory, L. Bickman (ed.). San Francisco: Jossey-Bass.

Ehren, M. C. M. (2011). Risk-based school inspections of Dutch schools and school boards. A critical reflection on intended effects and causal mechanisms. Manuscript.

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.

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.

Ekholm, M. (2012). Samverkan i styrenings [Cooperation in governing]. http://www.eveduc.com/images/doc/vagval_i_skolans_historia_samverkan_styrningen.pdf

Ekholm, M. & Lindvall, K. (2008). Skolinspektioner - i tid och otid [School inspections – in due time and not so due time]. Pedagogisk Forskning i Sverige, 13(1), 41–58.

Elstad, E. (2009). Schools which are named, shamed and blamed by the media: school accountability in Norway. Educational Assessment, Evaluation and Accountability, 21(2), 173–189.

Erickson, G. & Lander, R. (2007). Der Kitt, der ein wachsendes System zusammenhält? Nationale Test als Kern der Qualitätssicherung in Schweden. Pädagogik, 59 Jahrgang, Heft 3, 32–35.

ESV. (2006). Skolverkets utbildningsinspektion – ger den några effekter? [The school inspections conducted by the National Agency for Education – do they have any effects?]. Stockholm: Ekonomistyrningsverket.

Forsberg, E. & Nordzell, A. (2013). ‘Dialogue’ for school improvement? Nordic Studies in Education, 33, 187–203.

Gustafsson, J.-E., Ehren, M.C.M., Conyngham, G., McNamara, G., Altrichter, H. & O’Hara, J. (Submitted). School inspections and school improvement: relations according to principals’ reports.

Gustafsson, J.-E. & Erickson, G. (2013). To trust or not to trust? Teacher markings versus external marking of national tests. Educational Assessment, Evaluation and Accountability, 25(1), 69–87.

Hansen, M. & Lander, R. (2009). Om statens verktyg för skoljämförelser. Vem vill dansa SALSA? [On the state’s instrument for school comparisons . . .]. Pedagogisk Forskning i Sverige, 14(1), 1–22.

Hargreaves, D. H. (1995). Inspection and school improvement. Cambridge Journal of Education, 25(1), 53–61.
House, E. R. (1981). *Evaluation with validity*. Beverly Hills: Sage.

Husfeldt, V. (2011). Wirkungen und Wirksamkeit der externen Schulevaluation. Überblick zum Stand der Forschung. *Zeitung für Erziehungswissenschaft*, 14, 259–28.

Hussain, I. (2012). Subjective performance evaluation in the public sector. Evidence from school inspections. CEE DP 135. London School of Economics. Centre for the Economics of Education.

Isling, Å. (1988). *Det pedagogiska arvet. Kampen för och emot en demokratisk skola 2* [The pedagogical heritage. The fight for and against a democratic school, part 2]. Stockholm: Sober Förlag.

Jarl, M. (2012). *Skolan och det kommunala huvudmannaskapet* [The school and the municipal authorityship]. Malmö: Gleerups.

Jarl, M., Fredrikson, A. & Persson, S. (2011). New public management in public education. A catalyst for the professionalization of Swedish school principals. *Public Administration*. doi: 10.1111/j.1467-9299.2011.01995.x.

Knorr Cetina, K. (1999). *Epistemic cultures. How the sciences make knowledge*. Cambridge: Harvard University Press.

Lander, R. & Granström, K. (2000). Skolinspektion i England och Sverige. Hjälp till självhjälp eller självstyrning med betslet i munnen? [School inspection in England and Sweden ...]. *Pedagogisk forskning i Sverige*, 5(3), 215–234.

Lawn, M. & Grek, S. (2012). *Europeanizing education. Governing a new policy space*. Oxford: Symposium Books.

Leeuw, F. L. & Furubo, J.-E. (2008). Evaluation systems. What are they and why study them? *Evaluation*, 14(2), 157–169.

Levačić, R. (1995). *Local management of schools. Analysis and practice*. Philadelphia: Open University Press.

Lewin, L. (2014). *Staten får inte abdikera – om kommunaliseringen av den svenska skolan* [The state must not abdicate – about municipal ruling of schools]. Stockholm: Utbildningsdepartementet.

Luginbuhl, R., Webbink, D. & De Wolf, I. (2009). Do inspections improve primary school performance? *Educational Evaluation and Policy Analysis*, 31(3), 221–237.

Lundgren, E. (2013). Subjektspositioner spelar roll. Hur diskurser kan ... [Subject positioning plays a part. How discourses ...]. Examensarbete (PDA 252) vid institutionen för pedagogik och specialpedagogik, Göteborgs universitet.

NAE. (2012). *Likvärdig utbildning i svensk grundskola. En kvantitativ analys av likvärdighet över tid* [Equal education in the Swedish comprehensive school ...]. Stockholm: Skolverket [The National Authority of Education].

Nytell, H. (2006). *Från kvalitetsidé till kvalitetsregim. Om statlig styrning av skolan* [From ideas of quality to a regime of quality. About state governing of schools]. Uppsala Studies in Education No, 114. Uppsala. Acta Universitatis Upsaliensis.

Pagil, M. (2008). När staten axlar sitt ansvar - en studie av den lägmälda kampen kring skolinspektionen [When the state takes responsibility – a study of the low-voiced struggle about the school inspection]. Uppsats inom lärarutbildningen, Rapportnummer HTo8-2611-213, Göteborgs universitet.

Pawson, R. & Tilley, N. (1997). *Realistic evaluation*. London: SAGE.

Pettersson, S. & Wallin, E. (1998). Målstyrning och resultatansvar [Management-of-objectives and responsibility for results]. I *Analyser och utvärderingar 1996–97*, Skolverket. Stockholm: Liber.

Rosenthal, L. (2004). Do school inspections improve school quality? Ofsted inspections and school examinations in the UK. *Economics of Education Review*, 23, 143–151.
Ronnberg, L. (2011). Exploring the intersection between marketisation and central state control through school inspection. *Education Inquiry, 2*(4), 689–707.

Ronnberg, L., Lindgren, J. & Segerholm, C. (2013). In the public eye. Swedish school inspection and local newspapers. Exploring the audit–media relationship. *Journal of Education Policy, 28*(2), 178–197.

Sandström, B., Arvidson, C. & Landahl, J. (2005). Engagemang i slott och koja. Om effekter i utvecklingsdialogerna [. . . about effects in the developmental dialogues]. Stockholm: Myndigheten för skolutveckling.

SFS. (2011:556). Förordning med instruktion för Statens skolinspektion [Instruction for the Swedish school inspection]. Regeringskansliet.

SFS. (2010.800). Skollag. [Educational Act 2010.800]. http://www.riksdagen.se/Webbnav/index.aspx?nid=3911&bet=2010.800

Shaw, I., Newton, D. P., Aitkin, M. & Darnell, R. (2003). Do Ofsted inspections of secondary schools make a difference to GCSE results? *British Educational Research Journal, 29*(1), 63–75.

Skolinspektionen. (2008). Årsredovisning 2008. [Yearly report 2008]. Dnr 10-2008-411. Stockholm.

Skolinspektionen. (2009a). Vi vassar pennan. *Skolinspektion för kunskap och kvalitet [Sharpening the pen. School inspections focusing on knowledge and quality]. Stockholm: Skolinspektionen and Norsteds JuridikAB/Fritzes.*

Skolinspektionen. (2009b). Årsredovisning 2009 [Yearly report 2009]. Dnr 10-2009-4224. Stockholm.

Skolinspektionen. (2010a). *The inspectorate of educational inspection of Sweden.* http://www.skolinspektionen.se/PageFiles/1854/SwedishSchoolsInspectorate2009.pdf?epslanguage=en

Skolinspektionen. (2010b). Årsredovisning 2010 [Yearly report 2010]. Dnr 10-2010-6658. Stockholm.

Skolinspektionen. (2011). Årsredovisning 2011 [Yearly report 2011]. Dnr 10-2011.6465. Stockholm.

Skolinspektionen. (2012a). Lika för alla? Omrättning av nationella prov i grundskolan och gymnasieskolan under tre år [Equal for all? Re-correcting national tests . . .]. Stockholm.

Skolinspektionen. (2012b). Årsredovisning 2012 [Yearly report 2012]. Dnr 10-2012-5700. Stockholm.

Smith, G. (2000). Research and inspection. HMI and OFSTED, 1981–1996 – a commentary. *Oxford Review of Education, 26*(3), 333–352.

SOU. (2007.11). *Tydlig och öppen. Förslag till stärkt skolinspektion [Transparent and open. Reinforcing school inspection, commission report]. Stockholm: Utbildningsdepartementet.*

Sun, H., Creemers, B. P. M. & de Jong, R. (2007). Contextual factors and effective school improvement. *School Effectiveness and School Improvement, 18*(4), 93–122.

Thelin, B. (1994). Skolinspektörer minns (School inspectors remember). Årsböcker I svensk undervisningshistoria. Vol. 74. Uppsala. Föreningen för svensk undervisningshistoria.

Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation Review, 21*, 501–524.