Literature review: Impact of inspections on the improvement of schools and of networks of schools

Contents

Review questions and key terms

The leading question in this review is:
- What is the impact of school inspections on improvement of schools and of networks of schools?
- What are the mechanisms of impact: how does inspection lead to improvement?

Key terms:

- **School inspections**: School inspections are a specific case of external evaluations of schools, undertaken by officials outside the school with a mandate from a national/local authority. Regular visits to schools are an essential part of school inspections to collect information about the quality of the school, check compliance to legislation and/or evaluate the quality of students' work (e.g. through observations, interviews and document analysis). School inspections are a special case of external evaluation in having formal or informal consequences in place for failing and/or high performing schools (sanctions/rewards) and use central (national/district/state) standards to assess schools. These can be performed by the school district, the state department of education, or a ministry of education using professional evaluators or regional inspectors, or a district/state/national evaluation department. In the literature these are variously referred to as (external) reviews, (external) evaluations or supervisions.

- **Internal evaluations:**
  "At the school level, internal evaluation can be performed by a teacher or a group of teachers, by other members of the school's professional personnel, by the principal or other school administrators, or by a special staff member designated by the school to serve as a school evaluator" (Nevo 2001 p.95). The impetus for this type of evaluation should come from the school itself and involve a process in which relevant stakeholders Plan (meet up, discuss, agree on aims of evaluation); Do (carry out evaluative process, such as observing lessons, interviewing staff, students and examining documentation); Check (assess the extent to which the planned aims are being met and other peripheral findings); Act (follow up the findings and recommendations from the review process). Internal evaluations are formative when focused on assessing strengths and weaknesses to inform continuous improvement and school development processes, or have a summative aspect when self-evaluation reports (resulting from the 'Check' phase) feed into external accountability or external evaluations. Internal evaluations are variously described as (school) self-evaluation; inquiries; internal accountability; or (internal) reviews. *NB: These were not included in this review.*

- **Interaction:** This is where there are parallel, sequential or cooperative models used to evaluate schools, where self-evaluations are, in some way, shape or form, part of an external inspection. In *parallel evaluation*; internal and external evaluators do not participate in each other's evaluation. According to Kyriakides and Campbell (2004)and Christie et al (2004), both the school and the external body conduct their own evaluations and maybe compare and share findings afterwards. In sequential evaluation, schools conduct their own evaluation and the external body (for example an Inspectorate) then uses the results of the self-evaluation as a basis for its external evaluation. The external evaluator can for example analyse the internal evaluation data or use it in a meta-evaluation (Christie, Ross et al. 2004, Van Petegem and Vanhoof 2007)). External evaluators may also evaluate internal evaluation criteria related to the functioning of the school, on the request of the school. Evaluation of these criteria can be described in a supplement to the external evaluation report and has a formative function according to Vanhoof and Van Petegem (2007). According to Kyriakides and Campbell (2004), sequential evaluation may also work in the opposite direction when the external body provides feedback to the school which is expected to be used in self-evaluations and improvement of the

school. In this case, the action plan for improvement of the school will reflect the criteria and judgments of the external inspectors instead of the school-defined criteria. Vanhoof and Van Petegem (2007) describe how internal evaluations may also be focused on the extent to which the school complies with external legislative criteria. The aim of internal evaluation in this case is to respond better to these external criteria. The third, and final, type of combination integrates internal and external evaluations in all three phases of evaluation. This type of combination is described as the *cooperative model* (Kyriakides and Campbell, 2004), collaborative evaluation (Christie et al, 2004) or joint evaluation (Eurydice, 2004). In this model, internal and external evaluators collaborate to plan, design, conduct and report on the evaluation together. As a result, the interests and viewpoints of external evaluators as well as the school-defined criteria are taken into account simultaneously. According to Kyriakides and Campbell (2004), measurement criteria come up from both types of evaluation and an attempt is made to combine the results in order to conduct a holistic evaluation and satisfy the needs of both parties.

By interaction we do not mean studies that look at changes in self-evaluations as a result of inspections (in this case self-evaluations would be the outcome variable of inspection). *NB: While many papers informed this interaction category, these were counted as either inspection or internal evaluation in the data extraction spreadsheets according to their main focus. A separate review was conducted of internal evaluation.*

- **Other:** These may include externally-mandated school self-evaluations, for example where a school is required to undertake a review by its umbrella academy trust, federation or local district and models of peer review that may mix elements of internal/external evaluation and even inspection. The unit of focus of activities may not be the school, rather a district or network of schools or a department within one school or departments cooperating across schools (in the case of some peer reviews). *NB: As with the above category, papers were classified as either inspection or internal evaluation for the purposes of data extraction tables.*

Inclusion criteria:

We focused on evaluation of <u>schools</u>, and excluded evaluation and inspection of individuals (e.g. head teachers, teachers). Additional inclusion criteria were year of publication, articles published from 2000 until April 2015, high income countries, setting (primary and secondary education), and language (English). Literature reviews on inspections for improvement were also used as a reference point for searches of relevant articles and books. Articles and books prior to 2000 were included if they were mentioned in these literature reviews and were relevant to the above four categories.

We looked at non empirical studies (e.g. think pieces, exploratory studies) to increase our understanding of potential mechanisms of impact. In reporting findings we clearly distinguish which results are from empirical studies and which are from exploratory studies or think pieces. Publications for review were drawn from educational research, official government and international body publications, and a variety of other sources of evidence, including internet resources. Unpublished work drawn from an EU project, the Impact of School Inspections on Teaching and Learning was also included. Overall, the literature reviewers erred towards including studies where there was a degree of ambiguity regarding their relevance in order to aim for a comprehensive coverage of the field.

See below for databases searched and search terms.

The literature review process:

This was carried out in steps and involved two research officers with general educational knowledge but non- expert in the field. The process was overseen by a Senior Lecturer who is an expert in the field of Inspection and evaluation. Refinements to the process of data extraction, understanding of key concepts and search terms took place through regular meetings and discussions.

Step 1

Searches: A research officer carried out an initial search to identify possible books, articles and reports. Lists of titles of possible references were identified through searches of general databases, journal searches and searches of library catalogues. All keywords listed below were used in the searches of general databases and lists of titles were scanned for relevance. This resulted in a large number of titles that were further filtered for relevance according to the above categories of inspection, internal evaluation, interaction and other.

Journal searches and library catalogue search used more restricted sets of search terms to search 'keywords' and 'all text': 'school inspection', 'school evaluation' and 'school accountability', 'data and school improvement'. An additional sweep of articles on evaluation was carried out by a researcher by referring to bibliographies and searches for authors frequently published in the field when the first set of searches were found to have relatively few articles compared to those on inspection. Where titles were ambiguous, abstracts were referred to, if available, and more recent titles were prioritised. Manual scanning of database records was then used to refine title lists and to identify those authors who had published frequently in the field. As the steps of the search proceeded, additional titles were added, based on bibliographies of relevant titles and through use hand searches of journals and other relevant sources.

Step 2

Data extraction and summaries: Literature was filed and classified for inclusion in separate databases for empirical studies, non-empirical studies and literature reviews. Previous literature reviews were not included in counts but used to inform searches relevant to this review. These are summarised in separate excel and word files. Where empirical papers contained elements of separate data sets for each country, these were counted as discrete studies. Also, if individual papers referred more than once to one data set/study, these were aggregated in the data extraction and counted only once. Empirical studies were classified methodologically according to the Maryland Scientific Methods Scale (MSMS).

The spreadsheets include brief summaries of findings and conclusions with regard to the how inspections or self-evaluation related to school improvement or important mediating and contextual factors. Separate, extended summaries were compiled in corresponding word documents to aid the writing of the literature review.

Step 3
The summaries of the literature were used to inform evidence tables which provide an overview of the different types of effects and mechanisms and conditions explaining such effects. These evidence tables (see below) were then used to write four chapters on effects and side effects of school inspections, mechanisms of change from school inspections, and effectiveness of school self-evaluation.

	Type of effect					
	No effect	*Reflection on school quality and intentions to improve*	*School improvement*	*Improvement of self-evaluation and innovation capacity*	*Improved student achievement*	*Side effect*
Number of studies						
Country						
Year in which studies were reported						
Authors						

	Type of mechanism			
	Performance feedback	*Setting expectations*	*Stakeholder involvement*	*Capacity-building*
Number of studies				
Country				
Year in which studies were reported				

Databases searched and search terms:

Sources were identified through:

- General databases (e.g. Digital Resource Archive (DERA), British Education Index (BEI); Australian Education Index (AEI); ERIC; Web of Science); internet search engines and gateways (e.g. Google Scholar); websites of inspectorates, education charities, policy 'thinktanks', conferences, unions.
- Library catalogues, such as:
  - American Educational Research Association
  - Australian Centre for Economic Performance
  - British Educational Research Association
  - Bristol Institute of Public Affairs
  - Business in the Community
  - CfBT Education Trust
  - Consortium on Chicago School Research
  - Department for Education
  - Education Scotland
  - Education Sector
  - Estyn
  - FORUM: for promoting 3-19 comprehensive education
  - Her Majesty's Inspectorate of Education
  - London School of Economics Centre for Economic Performance
  - National Education Policy Center

- Ofsted
- Research for Action.
- Royal Society of Arts and Manufactures
- Social Policy Association Conference
- US Department of Education
- Professional fora (e.g. Times Education Supplement, Guardian Professional,…)
- Manual keyword searching of journals:
- American Journal of Sociology
- British Journal of Educational Studies
- British Journal of the Sociology of Education
- British Educational Research Journal
- Cambridge Journal of Education
- Canadian Journal of Educational Administration and Policy
- Comparative Education
- Current Issues in Comparative Education
- Early Education and Development
- Education 3-13
- Education Inquiry
- Educational Action Research
- Educational Administration Quarterly
- Educational Assessment, Evaluation and Accountability
- Educational Management Administration & Leadership
- Educational Policy
- Education Policy Analysis Archives
- Educational Research Review
- Ethnography and Education
- European Education Research Journal
- European Journal of Education
- European Journal of Training and Development
- Improving Schools
- International Journal of Educational Management
- International Journal of Education Policy and Leadership
- International Journal of Qualitative Studies in Education
- Journal of Education Policy
- Journal of Educational Change
- Journal of School Choice
- London Review of Education
- Management in Education
- Quarterly Review of Comparative Education
- Research in Education
- Scandinavian Journal of Educational Research
- School Effectiveness and School Improvement
- School Leadership & Management
- Studies in Educational Evaluation
- Studies in the Cultural Politics of Education
- Teacher Development
- Scanning lists of references
- Contacts with those in the professional networks of the research team and those suggested by key informants.

Keywords for internal/external evaluations, monitoring and inspections

School inspection, internal evaluation, external evaluation, Internal/external audit, Critical friend, Compliance, Teaching school networks/alliances, underperforming schools, special measures schools, failing schools, schools (requiring or needing) improvement, good schools, outstanding schools, satisfactory schools, unsatisfactory schools, (coasting or stagnating or stuck) schools, administrative organization, educational monitoring, administrator evaluation, bureaucracy, database management systems, decision support systems, educational indicators, information management, information systems, information utilization, management information systems, management systems, performance information, performance factors, performance management, performance indicators, program monitoring, progress monitoring, school performance, progress reporting, recordkeeping, records, school-level data, data and school improvement, school self-evaluation, SSE, self-assessment, student evaluation of teacher performance, teacher evaluation, total quality management, database management systems, school monitoring, EMIS, school performance data, monitoring systems, school governance, education governance, school boards, Governing education, school autonomy, school efficiency, national information systems, school marketisation, school league tables or school rankings, feedback and school, external review, inspection & review, quality control, quality review, quality management, dynamic school improvement, school inspectors, school supervision, school visitation, supervision, supervisor qualifications, supervisor- supervisee relationship, supervisors, teacher supervision, institutional evaluation, inspectorate, school evaluation, school regulation, school peer inspection/review, bottom-up/top-down inspection, school self- inspection, school self- review, school self- regulation, Ofsted, Estyn, HMIE plus other names of inspectorates, accountability and gaming, educational accountability, standards based accountability.

References:

Christie, C. A., et al. (2004). "Moving toward collaboration by creating a participatory internal-external evaluation team: A case study." Studies in Educational Evaluation 30(2): 125-134.

Kyriakides, L. and R. J. Campbell (2004). "School self-evaluation and school improvement: A critique of values and procedures." Studies in Educational Evaluation 30(1): 23-36.

Nevo, D. (2001). "School evaluation: internal or external?" Studies in Educational Evaluation 27: 95-106.

Van Petegem, P. and J. Vanhoof (2007). "Towards a Model of Effective School Feedback: School Heads' Points of View." Educational Research and Evaluation 13(4): 311-325.

Annotated bibliography of empirical research.

This annotated bibliography summarises those texts consulted for empirical studies on the impact of inspections and includes some not counted in the spreadsheet that were referring to the same data set. It only notes aspects of the studies which relate directly to the focus of the review.

Allen, R. & Burgess, S. (2012). How should we treat under-performing schools? A regression discontinuity analysis of school inspections in England. Bristol, University of Bristol, Centre for Market and Public Organisation, Bristol Institute of Public Affairs.

Allen and Burgess compared the performance statistics for two groups of secondary schools in England over the period 2002- 2009. The two groups were identified as 'just failing' or 'just passing' an Ofsted inspection through analysis of a large number of sub- criteria within inspection judgements. Neither sanctions nor specific, targeted support were applied to either category of schools, both of which would be likely to be described as 'coasting' or fall into the current (January, 2014) Ofsted descriptor of 'requires improvement'.   Samples were otherwise matched for other factors that are likely to have influenced student achievement. They found evidence of greater improvement in performance in compulsory core subjects for schools identified as 'just failing' as 'moderate to large at around 10% of a pupil-level standard deviation in test scores' (author abstract).   However, there is little improvement for lower ability pupils, with gains noted for average and above average ability pupils.

Altrichter, H. and D. Kemethofer (2014). Impact of School Inspection on Teaching and Learning in Primary and Secondary Education in Austria: Final-technical-report-Austria, schoolinspection.eu.

The theoretical framework summarizes program theories (Leeuw 2003) of six European Countries1. On the basis of these six program theories a conceptual model was developed (Ehren et al., 2013) to describe the mechanisms by which inspectorates aim to monitor school quality and stimulate school improvement. A longitudinal design was used, administering an online questionnaire over three years (2011, 2012 and 2013) in the federal state of Styearia. This investigated intermediate mechanisms (setting expectations, accepting feedback, promoting/improving self-evaluations, taking improvement actions, actions of stakeholders) and the outcomes (improvement capacity, effective school and teaching conditions) of school inspection. A five point Likert scale (strongly agree-strongly disagree) and another five point scale related to time spent (much less time=1, much more time=5). The target sample was all primary and non-academic secondary schools in Styearia. The sample included 693 cases with 190 secondary schools, 451 primary schools and a number of 37 cases which refer to the school type "other". The response rates were higher in year one (69% and 77% respectively), dropping to 35% and 48% in year 3). The report was interested also in perceptions about the new model of 'team inspection' introduced in Styearia, in 2007/8. This inspection regime, with the intention to drive school improvement, includes a 2 or 3 day visit by a district inspector and one outside inspector and covers: Classroom observation, Group interviews with parents' representatives, teachers, and students, Interview with the mayor, Meeting with the headperson including the analysis of documents and a conversation about potential development fields and Site inspection. Some days after the school visit the inspectors present their preliminary inspection report in a feedback conference to the school staff and discuss it with teachers. The headteachers, usually in collaboration with school staff, then draw up a school development plan and timetable. 4-5 weeks later the plan is discussed and agreed (or amended) with the district inspector. There are no positive or negative sanctions attached to the results of the inspection.

In comparison with schools that had not been inspected in the previous year, inspected schools scored higher on "accepting feedback" and a trend in "setting expectations" and "stakeholders'

sensitiveness". In both years we observe significant effects for stakeholders' sensitiveness and setting expectations on accepting feedback and promoting/improving self-evaluation. Accepting feedback itself has no significant effect on development activities. Self-evaluation has a significant effect on capacity building and improvement of school effectiveness (only year 1). Schools that report improvement in capacity building also report of more activities in school effectiveness. The results showed no constant relationship between causal mechanisms of school inspections and activities in school improvement. The results indicate the need for more research to further investigate the interaction of the mechanisms of school inspections and schools" development activities.

Altrichter, H. & Kemethofer, D. (2015) Does accountability pressure through school inspections promote school improvement? School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice, 26:1, 32-56.

 "New" school inspections are essential parts of "evidence-based governance" concepts and have been implemented by many European countries as a major strategy to assure and improve the effectiveness and quality of their education systems. However, national inspection systems vary in their composition and in their contextual features. Using online survey data from approximately 2300 principals in 7 European countries, the paper explores the role of "accountability pressure" as an element for understanding the operation of inspection systems. The results indicate that principals who feel more "accountability pressure" are more attentive to the quality expectations communicated by inspections, more sensitive to stakeholders' reactions to inspection results, and more active with respect to improvement activities. However, also the number of unintended consequences is increasing with pressure. Inspection systems in different countries are seen by school leaders as applying differential degrees of "accountability pressure", which is reflected in system-specific amounts of improvement activities.

The authors consider inspections systems in 7 countries. Based on factors such as the availability of sanctions following inspection, the publication of inspection reports, differentiated inspections depending on outcomes they classify these on a scale from 5 to 0 of high to low accountability as follows: The Netherlands 5, England 5, the Czech Republic 3, Sweden 2, Ireland 2, Austria 0, Switzerland 0-1. Survey responses from targeted samples in each country of both primary and secondary schools were used as the basis for analysis. Nearly two thirds of respondents felt pressure to do well on measures of inspection, and these respondents were also more likely to have reported putting in place measures such as self- evaluation, build development capacity among staff, improve teacher participation in decision making, enhance teacher co-operation, and improve transformational leadership. Those principals who reported feeling pressure to do well also reported more effects related to the intermediate mechanisms proposed by Ehren et al. (2013) in that they agreed that the inspection process set expectations for performance. However there was no difference in responses from this group in relation to stakeholders' reception of reports or acceptance of feedback. School leaders who reported experiencing more pressure reported significantly more unintended consequences: discouraging new teaching methods; narrowing the curriculum and instructional strategies.

On the whole, their findings support their analysis of the pressure exerted by the inspection systems in the countries covered, with principals in England and the Netherlands most likely to note the pressure of inspection and those in Austria and Switzerland reporting the least pressure. The accountability system in each country is also linked to responses about actions following inspections and to the prevalence of unintended consequences. However, the responses in relation to the intermediate mechanisms identified in Ehren et al. (2013) are inconclusive.

Table 6. (p45) Pressure to do well on the inspection standards by countries.

I feel pressure to do well on the inspection standards

Country	1	2	3	4	5	mean	SD
ENG	1.70%	0.00%	3.00%	32.80%	62.80%	4.54	0.717
NL	2.20%	0.00%	8.90%	66.70%	22.20%	4.07	0.72
SE	2.50%	4.20%	15.50%	54.60%	23.10%	3.92	0.882
IE	4.10%	20.70%	18.20%	38.80%	18.20%	3.46	1.133
CZ	8.90%	6.70%	35.60%	44.40%	4.40%	3.29	0.991
CH	14.50%	16.10%	25.80%	38.70%	4.80%	3.03	1.159
AT	14.10%	22.10%	32.20%	27.40%	4.00%	2.85	1.097
Total	6.40%	10.20%	18.60%	40.80%	23.90%	3.66	1.137

Notes: N = 1169; n(ENG) = 235; n(NL) = 45; n(SE) = 355; n(IE) = 121; n(CH) = 62; n(CZ) = 45; n(AT) = 298;

chi2 = 479.468, df = 24, p = .000.

1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree.

Table 7. (p46) Means of accountability pressure by country.					
Pressure					
Country	High	←→		Low	
ENG	4.54				
NL		4.07			
SE		3.92			
IE			3.46		
CZ			3.29		
CH				3.03	
AT				2.85	

Notes: The table represents means of accountability pressure reported by school leaders

in the respective countries. Figures within a column do no differ significantly; between

columns there is a significant difference (p ≤ .05; Duncan test).

For further analyses, we included countries which conformed to the following criteria: Countries were (a) to represent different "accountability traditions" (b) to be characterized by an empirically distinctive amount of "accountability pressure" (indicated by significant differences in Table 7), and (c) to include an adequate number of cases for further analyses. England, Sweden, and Austria fit these criteria.

England represents a system with high "accountability pressure", in Sweden principals feel a little less, but still a considerable amount of "accountability pressure", while Austrian school leaders seem to live in a "low accountability pressure" system. Data from The Netherlands, Ireland, the Czech Republic, and Switzerland were excluded.

Improvement process	Accountability system	N	Mean	SD	p (≤ .05)
Promoting/improving self-evaluation	Austria	482	3.19	0.657	ENG > SE > AT
	Sweden	998	3.51	0.57	
	England	217	4.03	0.791	
Improvement in capacity building	Austria	496	3.52	0.533	ENG ≈ SE > AT
	Sweden	1007	3.82	0.438	
	England	221	3.82	0.535	
Improvement in teacher participation in decision making	Austria	487	3.52	0.711	ENG > SE > AT
	Sweden	1005	3.79	0.636	
	England	220	3.95	0.705	
Improvement in teacher cooperation	Austria	490	3.64	0.702	ENG > SE > AT
	Sweden	998	4.07	0.59	
	England	220	4.18	0.638	
Improvement in transformational leadership	Austria	482	3.36	0.612	SE ≈ ENG > AT
	Sweden	998	3.6	0.633	
	England	220	3.56	0.713	

Table 8. (p47) Improvement process by accountability system.

Note: The last column presents the results of the comparison of group means (Duncan test) and is to be read in the following way: ">" … the group mean is significantly bigger (p ≤ .05) than …; "≈" … the group mean does not significantly differ (p ≤ .05) from …

Table 9 (p48) Intermediate inspection mechanisms by accountability system.					
Intermediate mechanisms	Accountability system	N	Mean	SD	p (≤ .05)
Setting expectations	Austria	310	3.34	0.536	ENG > SE > AT
	Sweden	357	3.55	0.689	
	England	238	3.96	0.634	
Stakeholders sensitive to reports	Austria	288	3.32	0.614	ENG > SE > AT
	Sweden	351	3.56	0.626	
	England	238	4.03	0.626	
Accepting feedback	Austria	305	3.88	0.661	AT ≈ SE > ENG
	Sweden	359	3.86	0.606	
	England	238	3.71	0.747	

Note: The last column presents the results of the comparison of group means (Duncan test) and is to be read in the following way: ">" … the group mean is significantly bigger (p ≤ .05) than …; "≈" … the group mean does not significantly differ (p ≤ .05) from …

Table A1. (p55) Summary of sample characteristics in each country.						
	Target sample		Actual sample		Response rate	
Country	Primary	Secondary	Primary	Secondary	Primary	Secondary
Netherlands (NL)	408	359	73	15	18%	4%
England (ENG)	1422	637	189	101	13%	16%
Sweden (SE)	1167	897	567	464	49%	47%
Ireland (IE)	3200	729	123	42	4%	6%
Czech Republic (CZ)	150	170	56	69	37%	41%
Austria (AT)	503	194	345	149	69%	77%
Switzerland (CH)	465		132		28%	

Note: In Switzerland, no information about school type in the target sample is available. Three principals responded that they work in a special school, and four did not respond to this question.

Araújo, M. 2009. A Fresh Start for a 'failing school'? A qualitative study. British Educational Research Journal, 35, 599-617.

This qualitative study looked at the effects of an initiative called 'Fresh Start' in England in the late 1990s. In this programme, schools that were considered 'failing' due to very poor pass rates in GCSEs were closed, new building, staff and leadership were introduced in their place. No additional funds were available to the staff, the principle being that increased accountability of staff and leadership would lead to improved outcomes. The study focused mainly on a form in Year 7 (aged 11) and their teachers. Semi-structured interviews, direct observation and the collection of documents were used to elicit data on the teachers' and pupils' experiences of schooling. A total of 23 pupils out of the 26 in the form studied, were interviewed in friendship groups of two in Year 7 and individually in Year 8 (aged 12). Over a period of 18 months. Results data (% A-C GCSE are also shown for the period between 1994 and 2006. (this started at 19% and ended at 48%) Their teachers and other school staff (such as the headteacher and learning mentors) were also interviewed, some twice. Observation focused on Science and Personal and Social Education lessons, due to the diverse statuses and classroom atmospheres associated with these two subjects, and also included Year assemblies and parent days. In addition, I collected and analysed school documents, including the prospectus, discipline policy, pupil records, incident reports and tables of attainment. This was a case study of Millhaven High was a co-educational comprehensive created in the early 1980s and located in a large English city. Before Millhaven closed, it served a largely disadvantaged community: nearly three-quarters of the school population were eligible for free school meals,5 over four times the national average (DfEE, 1998) and almost 30% of the pupils were refugees in England. These teachers believed that at Greenfield (tne newly created school) they would be able to 'start from scratch' and develop innovative pedagogical practices, an expectation created by the Fresh Start. However, the pressure on the initiative to raise 'standards' did not encourage this. The authors reveal (and compare to other research) that the newly opened school attracted a greater proportion of children with advantaged backgrounds. In 2000 this was 50% compared to approx 75% before. By year 7 this dropped to 30%. This case study warns of the dangers of such high stakes accountability: the school was never given credit for the work it did with refugee students or EAL students; Greenfield was creating an identity based on 'traditional' approaches, through the policing of teachers' work, strict discipline and increased selection within the school, favouring the 'more able'.

Bates, A. (2013). Transcending systems thinking in education reform: implications for policy-makers and school leaders. Journal of Education Policy, 28(1), 38-54.

Bates' case studies of two primary schools in England showed that a national emphasis on pupil standards as demonstrated in test scores had become a frequent feature of discourse within the schools studied, with comments such as 'The overriding improvement has to be on the scores on the doors, SATs …'(quote from member of school staff, Bates, 2013, p 46). She suggests that not only have school staff 'embraced' standards (p 46) but that this was reported in relation to parental expectations, who expect high test performance as one of the measures by which they judge the schools, although parents are also believed to want a more rounded education and curriculum. Staff in the schools studied expressed concern about the extent to which the significance of test results may cause stress and anxiety among pupils. She comments on the schools' reaction to a number of policy interventions, noting that they had welcomed the introduction of National Literacy and Numeracy Strategies in 1998/9 as introducing clear structure and progression as well as expectations about consistency within individual schools and throughout the system, but that they resented the frequent policy changes since as a distraction. Overall, Bates' study suggests that expectations for consistency and quality and measurement through test outcomes have become institutionalised in

the case study schools. Although this may result in superficial 'performativity' in teaching and learning, combined with frustration, cynicism and loss of commitment, her study also suggests that teachers are able to make professional judgements about the usefulness, or otherwise, of policy interventions and that the introduction of aspects of policy is moderated by underpinning values and beliefs about children's learning. She illustrates this as follows:

I never just accept … you've got to be brave and you've got to have the SATs results to back it up. But you can do it, you don't have to say 'we've got to follow this' … When something comes along we don't just say 'Oh dear', we look at it and we say 'okay, what's the best way we can use this for us?' And it's the way we can get round it, the little loopholes, can we cut corners … (headteacher, quoted in Bates, 2013, p 49).

Bates criticises the approach of Ofsted inspectors in judging schools solely on nationally- defined performance criteria without attention to the context of the school. Her study also illustrates the significance of the style of members of the inspection team in providing feedback, with an illustrative quote from the headteacher of a school that was judged to the 'good with outstanding features':

Jenny, the Headteacher, recalled her experience of conversations with Ofsted inspectors thus:

… some of them just come in and criticise … a lot depends on the team, it really does. If you've got somebody that's just chanting from a bit of paper … whatever evidence we gave them, they kept chanting the rules … (Bates, 2013, p 45).

Bates suggests that feedback which focuses on the few negative issues in a school which is otherwise performing well and improving may undermine confidence and commitment or, as she suggests happened in this example, increase cynicism and resentment about the inspection process.

Baxter, J. & Hult, A. 2013. Different systems, different identities: school inspectors in England and Sweden a comparative study. European Conference for Educational Research: Educational Improvement and Quality Assurance, Network 11. Istanbul.

This conference paper examines changes in Swedish and English inspectorates: in Sweden, this involves the use of people with legal and investigative backgrounds and stricter penalties for non-compliance. In England there has been an increase in the use of part time inspectors of current serving school leaders and a change to a more rigorous inspection combined with a developmental approach. The paper addresses two key questions:

"Which competencies are required for school inspectors within both systems, and why?

How do changes in the both systems affect school leader perceptions of inspection as a governing tool?" (p.1)

Their paper uses Jacobsson's theory of governance to examine differences in both systems as a result of the changes. This theory proposes the existence of:

Regulative activity, i.e. reliance on formal laws, directives and penalties for their violation

Inquisitive activity, i.e. compelling the inspected party to 'show and tell' and open up for scrutiny

Meditative activity, i.e. discussion, dialogue and negotiation about what constitutes best practice in that context (see p.2)

60 Semi structured interviews were carried out on inspectors, inspector trainers, school leaders, and heads of inspection services in both Sweden and England. The researchers (in the wider research from which the paper was derived) also did documentary analysis of official literature, inspection handbooks and inspection reports within each system sample; interviews with key 'system actors' at the international, national and local levels [90 in total] and investigation of the background, training, experience of each national Inspectorate. There are also detailed case studies of a sample of inspection 'events' (5 in each system). This paper is based upon the case studies carried out in Sweden and England and the interview data emanating from the studies. Both systems raise concerns about the relative weighting of statistical information gathered (mostly prior to the inspection) and the small amount possible to gain over a short inspection visit. Therefore, concerns were raised that inspectors had already made up their minds about judgements before arriving. Swedish inspectors were viewed negatively by some school leaders who described the process as more of an 'inquisition' and the interpersonal relations were not helped by the use of legalistic language that clashed with the realities of school decision making. By contract, English inspectors had to balance the inquisitive and meditative functions and sometimes found it hard to remain impartial. As serving school leaders, the inspectors valued the experience as professional development and many described the inspection process as being a shared process rather than being 'done to'.

While Swedish inspectors do not observe lessons or give individual feedback they do hold seminars based on a report they write in which school district and head teachers are invited to discuss the improvement plans. Overall conclusions: The English system uses more meditative activities to govern and this has potential for school improvement. The major issue appears to be the accusations of bias based on the nature of the data relied upon to make the judgement. This led to the 2013 NAHT rebellion and creation of INSTEAD and also 1 in 12 schools in 2012 submitting formal complaints. In the Swedish system, the judicial background of many inspectors means that governance is quite inquisitive but less open to accusations of partiality.

Baxter, J. A. 2013. An independent inspectorate? Addressing the paradoxes of educational inspection in 2013. School Leadership & Management, 34, 21-38.

his paper is based upon documentary analysis of English inspection frameworks, thematic reports and Ofsted inspection reports. The inspection reports were taken from five case study areas throughout England, geographically dispersed throughout the country. 40 interviews with key actors, including school leaders, heads of inspection services, inspectors. The new inspection standards (2012) were compared to the 2009 framework. This emphasise the professional expertise of inspectors and also the developmental nature of inspections. 'Professional dialogue' was emphasised in the 2012 framework, reflecting the nature of inspectors as current serving school leaders, one aspect of which could be considered a discussion between 'equals'. The inspector training "attempts to address the technical paradox of independence in attempting to moderate challenges to the reliability and robustness of Ofsted's methodology via collusion between inspectors and inspected. In addition it also addresses the social paradox of agency independence by aligning inspector identities with those of school leaders."(p.29). Comments from schools indicate that they suspect that the use of data about teaching and learning (i.e. attainment measures) dominated the judgements made by inspectors since the time given to the inspection visit was so little to make adequate judgments about teaching and learning. Inspectors in the new framework were not supposed to judge lessons according to any one preferred pedagogical framework, therefore some school leaders suspected that any approach would be fine as long as pupils' results were excellent. Policy documents now show a subtle shift in the role of Ofsted, not simply to inform,

but now to be the ultimate arbiter of what is considered 'acceptable standards of education in England'. Inspectors are now considered to be both expert and professional.

Baxter, J. A. 2013. Professional inspector or inspecting professional? Teachers as inspectors in a new regulatory regime for education in England. Cambridge Journal of Education, 43, 467-485

Drawing from a sub set of a larger ongoing study, comparing inspection systems in England, Scotland and Sweden, this paper looks at the increased role of serving teachers and school leaders in the revised inspection framework for England and the skills and professional judgements required in order to successfully implement the new system. The reports sampled included older 'section 10' inspections as well as more recent ones. There were analysed qualitatively and using discourse analysis to identify trends in use of key words such as 'teaching, standards and evidence'. The analysis looked for tensions between professional judgement, interpersonal relations and the regulatory role. 40 interviews across England in the South, North and Midlands were conducted of key 'system actors', who included inspection contractors from Serco, Tribal and CfBT, HMIs, headteachers and local authority personnel as well as a small number of people responsible for recruitment and training of inspectors. The new framework requires schools to have 'good' for teaching and learning in order to receive this grade overall. It also places more emphasis on observations of classes and describing the 'story of the school' therefore drawing more on inspectors' professional judgements. At the same time, the judgement needs to reflect on pupils' progress over time, not just an aggregate of lesson ob grades during the inspection visit. The papers examines whether employing serving teachers is based on the theory that they are best placed to make judgements about teaching or whether it is a way of having 'colluding' teachers to 'justify the existences' of the Ofsted inspectorate. Using Boyne's criteria (detector, director, effector) the paper looks at the extent to which inspectors operate at the third level, 'effector' i.e. in securing behavioural change in line with that required by a higher agency or government. Comparing teacher professionalism to that needed for an inspector, there were some discrepancies, for example, lead inspectors sometimes mentioned the difficulty of 'parking their baggage', i.e. the need for teachers to be aware that their own practice may not apply well to the situation they are inspecting. The teacher may be less able to apply 'relational distance' and therefore be a less tough inspector. Another tension is that in the new framework, no one way of teaching is suggested and therefore teachers are likely to apply their existing beliefs about what makes good pedagogy and learning. These are not explicit in the framework itself. Some teachers complained about the lack of discussion about schooling, education and pedagogy in the training. There is also a dictotomy, since inspectors may observe teaching that they would normally see as not good practice but the existing data (progress over time) upon which they rely to make their judgements for the report may contradict this. The article raises issues about the tensions in identity between teachers and inspectors. One example is that those teaches with resilient identities or good practitioners often come from very positive environments that may not apply to schools were this is not the case: "teaching inspectors from outstanding or good schools may have the most fixed ideas about what constitutes good teaching and learning. Concomitantly, they may also be less likely to accept that practices that work in their schools and contexts may well not be appropriate in other contexts and in relation to other schools' stories" (p479). Having an inspector from an outstanding school in a very different socio-economic area also opened the inspectorate to more complaints. The professional judgement also opens itself to criticisms of unreliability. There was a further tension: the new framework emphasised professional dialogue much more throughout the process, therefore the skills of the inspector to build effective relationships and collegiality were more important. This tended to ignore the high stakes nature of the inspection process which meant that the power relationship was highly unequal. Some other disadvantages: inspectors were unlikely to be released

to do more than one inspection per term, and therefore not enough time (only 5 days of training too) to develop the professional identity of an inspector.

Baxter, J., & Clarke, J. (2013). Farewell to the tick box inspector? Ofsted and the changing regime of school inspection in England. Oxford Review of Education, 39(5), 702-718.

The shift towards professional judgement of inspectors is useful in principle but in practice there are tensions with data based accountability. More consideration needed to how professional judgement can be usefully harnessed towards school improvement.

Bitan, Kristin; Haep, Anna & Steins, Gisela (2014): School inspections still in dispute – an exploratory study of school principals' perceptions of school inspections, International Journal of Leadership in Education: Theory and Practice. Prior to publication DOI: 10.1080/13603124.2014.958199

This paper assesses attitudes of principals from one German federal state (North Rhine- Westphalia) toward school inspections. In the study, 50 school principals from a mix of primary, secondary and special schools that had not yet been inspected were interviewed about their attitudes toward school inspections. The contents of their statements were analysed in a qualitative study. The central result was that 42% of the respondents were generally cooperative and had positive attitudes toward school inspections. The analysis of the school principals with rather negative statements (32% of the respondents) reflects a criticism of the anticipated workload in relation to the expected efficacy of school inspections and emotional as well as practical problems. Prior to the visit by the interview, who shared information about the expectations and process of an inspection, Knowledge by hearsay was the most frequently named source of knowledge about school inspections. Although the system in Germany is intended to be supportive of school improvement, rather than for accountability purposes, not all respondents believed this 'Because QA is aimed at supporting school development, this implies that it is a cooperative procedure rather than one of control. Rumours about side effects of the inspection, such as window dressing or game playing nevertheless show that it is perceived as a procedure of control. Statements related to this topic were often voiced in our study.' (p 15)

On the basis of their findings they suggest that there should be greater dialogue between inspection teams and school principals, to remove misconceptions based on rumour and for sharing the decision making process on judgements made.

Braun, A., Maguire, M., & Ball, S. J. (2010). Policy enactments in the UK secondary school: examining policy, practice and school positioning. Journal of Education Policy, 25(4), 547-560.

Braun et al. used case studies to examine policy enactments in four secondary schools, thus providing examples of the extent to which and the way in which such policies are institutionalised. They argue that school context is critical and includes 'aspects such as school intake, history, staffing, school ethos and culture, as well as "material" elements like buildings, resources and budgets, and external environments such as local authority relations and pressures from league tables and national bodies such as Ofsted' (p585). They report a 'preoccupation' (p586) in all four schools with Ofsted inspection and test results, as well as recognition that these may be positive influences. One of the schools had been considered a 'coasting' school in 2002 with complacent staff, 'you know, it was a really good thing that we had that Ofsted report because it did challenge those perceptions, it did enable us to really start moving things forward'. (teacher comment, p 594)

A further specific example is shown of how an Ofsted inspection prompts action at teacher level and also illustrates the acceptance of the inspection findings:

I'm looking at group work within schools and how to . . . improve group work . . . because in our school, one of the things that Ofsted raised was the fact that we've got _ some of our female students they don't involve themselves in lessons as much. (teacher comment, p 585).

This comment also illustrates a point made elsewhere in the article that external pressures may serve to reinforce stereotypes about students and learning, with, in this case perhaps, reinforcing beliefs about female students needing to involve themselves actively in lessons in order to learn and achieve.

Brimblecombe, N., Ormston, M., & Shaw, M. (1995). Teachers' Perceptions of School Inspection: a stressful experience. Cambridge Journal of Education, 25(1), 53-61. doi: 10.1080/0305764950250106

The way that inspector behaves mediates the stress that teachers feel for inspection as well as the preparation stage of inspection. Questionnaire data from 821 teachers, at all levels, in 35 different secondary schools nationally in the first year of full Ofsted inspections and interview data from 30 staff, at all status levels, in five schools.

Brimblecombe, N., Shaw, M., & Ormston, M. (1996). Teachers' Intention to Change Practice as a Result of Ofsted School Inspections. Educational Management Administration & Leadership, 24(4), 339-354.

38% of teachers intended to make changes shortly after an inspection visit, particularly in relation to the way they teach and organise classes. The tendency to change increased the higher up the teacher is within the organisation.

Brookes, C (Gen Sec of NAHT) Ch. 8 We need an inspection process. But not this one. In *DE WAAL, A. (Ed) & INSTITUTE FOR THE STUDY OF CIVIL SOCIETY 2008. Inspecting the inspectorate: Ofsted under scrutiny, London, Civitas.

"Inspection outcomes give authority to leadership teams in cases where the findings of the inspection propel the direction of travel defined by school self-evaluation

• The quality of the inspection teams is variable

• The reliance on external data is too heavy

• Ofsted inspections contribute to poor retention and recruitment of school leaders" (p. 85).

Survey of 2,789 Head Teachers in 2008 (interim findings). In Survey most said that inspection made no difference to pupil behaviour or learning and a negative impact on teacher morale and wellbeing. Also a strong feeling that inspection reports failed to capture aspects not covered in measurable outcomes (at the time it was CVA particularly).

Carlbaum, S. 2014. Equivalence and performance gaps in Swedish school inspection: context and the politics of blame. Discourse: Studies in the Cultural Politics of Education, 16.

This article looks at national school inspection reports in four municipalities in Sweden, exploring "performance gaps, equality and justice" in the educational system. The four municipalities of Gothenburg, Gävle, Växjö and Linköping were inspected in 2011 or 2012, after the implementation of the new school act (SFS – Swedish Education Act – 2010); geographical area and SSI division; and

municipals size (50,000 or more). The decision reports of all independent schools (43), public schools (78) and the four municipalities, which in addition included 10 city district reports for Gothenburg, were analysed. Analysis of the reports followed Braun's, professional, material and external contexts, when looking at educational achievement and performance gaps. "The situational context refers to a schools setting, its intake and its history. By focusing on the situational context I examine how the school intake, the location and the competition/local school market is described. The professional dimension refers to contexts such as values, teacher commitment and experience. The professional dimension focuses on questions of who or what is considered responsible, such as teachers' reasons and explanations for specific outcomes, and what credit such explanations are given in the decision reports by the SSI. The material dimension of context refers to resources in terms of budgets, staffing and buildings. How is the material context described in the decision reports, and what explanatory value is given?" (page 4). In terms of the material dimension, the assumption was made that by schools and districts providing attention to equivalence in resourcing this would reduce performance gaps. However, the Swedish inspection reports also stressed the need for schools to adapt pedagogy to take account of the different school intake. Reports (professional dimension) also stressed the need to reduce performance gaps by individualising and personalising teaching and learning practices. Stressing high expectations was seen as a vehicle for reducing gaps but environments where expectations were low were explained by teachers as problems with resourcing aspects which led to low expectations. The NAE has a statistical model (rather like contextual value-added CVA in England, called SALSA) which takes into account features such as parental educational level, proportion of students with one or both parents not born in Sweden, proportion of students not born in Sweden and proportion of boys and girls.

SSI decision reports did not take the SALSA statistic into account and these have been criticised as providing 'excuses' for underperformance. However, the reports themselves did make several references to other contextual factors such as the prevalence of students newly arrived in Sweden, type of housing and single-parent families, thus given some legitimacy to the idea that performance gaps were to do with external context. Nevertheless, the SSI report put the onus on the school to adapt practices in order to rectify these issues and sometimes made assumptions that teachers and school leaders set low expectations for certain pupils. Context reasons given by the school were seen as 'excuses' by the SSI. The authors suggest a contradiction in the prescription towards individualisation of instructions (implying differences between students) and the requirement to have 'high expectations' of all students. They also suggest that by focusing on individualisation of instruction, the focus is taken away from structural inequalities and reinforces a neo-liberal politics of 'blame'. The authors also suggest that the inspection regime reinforces the idea that central government is blameless, sets high expectations and wants equality of performance between groups but the responsibility lies with local government, schools and teachers.

Case, P., et al. (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.

Ethnographic method was used to collect data over a 3-year period, with one of the authors acting as a participant observer in three different schools. Field notes were complemented by in-depth key informant interviewing (Taylor & Bogdan, 1984) of six teachers and two head teachers. Two focus group meetings were conducted with two groups of teachers (six in each) from two schools, approximately two terms after their experience of inspection. These meetings were recorded and transcribed. Follow-up telephone interviews were pursued with the focus group members individually 1 year after inspection. Negative effects on well-being of teachers and disruptions to teaching before and after Ofsted inspections

Chapman, C. (2001). Changing Classrooms through Inspection. School Leadership & Management, 21(1), 59-73.

5 secondary schools. Questionnaires of staff in 5 schools and semi structured interviews in 2 of them. Case study of five English schools just after an Ofsted inspection found that high quality feedback may be the key to teachers' intentions to change practice. Approximately 20% of teachers studied felt that inspectors' feedback had prompted changes in teaching practice.

Chapman, C. 2002. Ofsted and School Improvement: Teachers' Perceptions of the Inspection Process in Schools Facing Challenging Circumstances. School Leadership & Management, 22, 257-272.

Preliminary findings from a study investigating teachers' views of Ofsted as a mechanism for improving secondary schools 'facing challenging circumstances' . There were schools in England that were the lowest attaining; and/or in Special Measures. Much higher (36% vs 19% national average) FSM students. Case studies of 10 secondary schools; from the Midlands and had Headteacher permission for access. They were sampled for variation in socio-economic context, cultural typologies (Stoll and Fink 1996) and leadership characteristics. A range of data were used to indicate the trajectory of improvement, including performance data, development plans and past Ofsted and HMI reports and the school management conditions survey.

3 Research q's:

How do teachers perceive the inspection process?

To what extent does Ofsted inspection generate changes in teaching and non-teaching practice in schools in challenging circumstances?

Does Ofsted inspection identify similar priorities for change to those identified internally by schools in challenging circumstances?

Semi-structured interviews: 45mins group interviews with smt, middle managers and teachers; 2 with headteachers. These were balanced in terms of gender, age, and experience, and varied by subject area. Documentary analysis contextualised the interview data. Perceptions and attitudes to the inspection were looked at. Contribution to school improvement:

Senior and middle managers had the most positive perception, particularly as a lever to school change, however they criticised it as encouraging short term thinking. Others saw the inspection as a useful 'audit', allowing this data to be used later to drive improvements. Teachers were less positive, citing examples of the poor judgement of the inspector, particularly with regard to lesson observations. The lack of discussion with teachers meant that it was not seen to help improve teaching (judgements hidden in the report). The preparation to the inspection was seen as more beneficial but the process draining for staff and pupils.

Internal leadership: Important issues highlighted were: the preparation for the inspection to mitigate the effects of timing of inspections and potentially variable quality of the inspectors. The good relationship of leadership with inspectors also helped them to challenge judgements, it was felt. Some SLT mentioned the need to adopt a more autocratic style of leadership due to time constraints leading up to the inspection (get on with it).

Changes to the inspection: the most important development was that inspectors were able to feed back to teachers and this helped them to improve their teaching. However, concerns existed over the quality of the inspectors.

The use of labels: Schools in 'special measures' or having 'serious weaknesses' gave legitimacy and impetus to the need for change. The need for external inspection was accepted by SLT in this process. It also meant that additional funding could be brought it. Removal of the special measure label was seen as a significant milestone for the school that signalled improvement. Improvement was seen in improved behaviour, attendance and increase in first choice applications. However, the 'naming and shaming' policy overall was seen as morale sapping and divisive and did not take into account variation among teachers and departments (some of which were very good). Others felt that the labels were applied inconsistently and incorrectly in some cases. The promotion of a 'low-risk culture' was seen as a consequence of inspection outcomes. HMI vs OFSTED inspectors: the long relationship built up with HMIs was seen as more helpful than the snapshot gained by OFSTED inspectors. Teachers' professional and personal health: Senior managers were seen to be under a lot of pressure and this was passed down to other staff. Generally the preparation period was very stressful. Changes to non-teaching practice: the inspection process was seen to encourage reflection on practices, particularly ones that led to better achievement. However, there was an absence of concrete examples of change. Changes to teaching practice: Middle managers and teachers reported little or no change to their teaching practice as a result of the inspection process. The former mentioned how the inspection spurred the writing of schemes of work but that these were not amended after the inspection.

Identifying priorities for improvement: Most reported that they already knew their priorities for improvement and did not need Ofsted to tell them. However, some did admit that these issues became re-emphasised and some may have worked more strongly towards them having been 'endorsed' by the inspection team.

Coldron, J., Crawford, M., Jones, S. & Simkins, T. 2014. The restructuring of schooling in England: The responses of well-positioned headteachers. Educational Management Administration & Leadership, 42, 387-403.

The article analyses interviews in three LAs with 15 headteachers whose school were well positioned locally. This article examines headteachers' responses in the light of the government's position of making school independent of local authorities. The LAs were in a variety of geographical, historical and policy contexts: a large metropolitan authority (A), a large rural authority with a dispersed population (B) and a smaller authority centred on a town (C). Within each area, interviews with: headteachers actively engaged in building collaborations, federations or the creation of academies or teaching schools. In 2011/12 academic year, 15 semi-structured interviews with headteachers across the three LAs. None was a member of a large national chain of schools. All but one had been graded good or outstanding at the last inspection. The article highlights the social and cultural capital gained by a school that has an outstanding designation and this puts them in a hierarchical relationship with other schools (which current reforms highlight too). The inspection report's effect is complex: "The Ofsted category, measured performance, league tables, intake and parental preference interact in a complex way to locate a particular school in the local, regional and national school fields"(p 390). For example, the Ofsted grade affected that ability of the school to go its own way through the academisation process. The extent to which headteachers felt that poor performing schools were due to the characteristics of the local area or the qualities of the leader were mixed – there was a contradiction here. The prestige of the school was seen as precarious "Despite their enjoyment of relative independence and autonomy, there was a strong awareness that one poor

inspection report or set of results could change their position drastically" (p.393). In general the article borrows on Bourdieau's concepts of capital and habitus to argue that the ability of a school to improve or extend its reach (in social capital) in order to improve other schools around it, depends very much on a variety of factors that affect its reputation.

Courtney, S. J. (2012). Ofsted's revised school inspection framework: experiences and implications. Paper presented at BERA conference, Manchester.

Courtney approached schools that had been recently inspected in the first three months following the introduction of a revised Ofsted inspection framework in January 2012 (this framework has been amended since then, with the latest amendments made in July 2013). In 26 responses to a survey sent to the 175 schools inspected in this period, with six follow- up interviews, principals reported that they would focus more on framework priority areas with an increased focus on, specifically: quality of teaching; leadership and management; pupil progress and literacy. Although self-evaluation is no longer mandatory under the 2012 framework, 61% of survey respondents said that they would continue to retain the same level of emphasis on this. Other themes identified in responses were the variability in the quality of inspectors, the implications of inspection outcomes for headteachers' careers and that the contexts and challenges facing schools serving areas of disadvantage are insufficiently taken into account. It is suggested that this may make it more difficult to recruit headteachers for such schools. The paper argues that the framework is more successful as a tool for controlling behaviour than for improving schools. It calls for improved inspector training, a focus on a broad, values-driven agenda by headteachers and increased recognition of contextual diversity in the inspection framework.

Courtney, S. J. & Steven, J. C. 2013. Head teachers' experiences of school inspection under Ofsted's January 2012 framework. United Kingdom.

This study looks at the impact of a new Ofsted framework introduced in 2012 that reduced the number of inspection judgements from 27 to just four: The quality of teaching, pupil achievement, leadership and management, and the behaviour and safety of pupils. Authors used a mixed methods design combining surveys of secondary school head teachers self-selected from 175 eligible participants (those inspected under the new framework and published on the Ofsted website between feb and march 2012. The sample was representative in terms of proportion awarded each of the four possible grades overall, i.e. iandequate, satisfactory, good and outstanding (11%, 39%, 42%, 8%, compared to 17%, 42%, 36% and 5% from the 175 in the population). 12 headteachers volunteered for follow up semi structured interviews and 6 selected to reflect a range of grades and geographical spread. Results suggested that new framework focuses strongly on VA scores in English and Maths and that it is increasingly difficult for school in deprived areas to rate highly. However, largely Headteachers agreed with new framework. In terms of the impact of the new framework headteachers rated whether there would be more, less or no change I focus on a number of issues (from table 3, p.168):

Top answer was more focus:
The quality of teaching
Pupils' literacy skills
Pupils' progress
Leadership and management

Top answer was less focus:
My school's community engagement

Pupils' health

Top answer was no change:
Care, guidance and support
 Behaviour
Pupils' safety
Other impact suggested from the qualitative data was that outstanding headteachers would be less likely to take a job in a deprived area due to worse job security and greater difficulty in achieving a high grade. Overall, it was felt that there was a 'climate of fear' created by the new framework.

Cross, A. 2006. Looking for patterns in Ofsted judgements about primary pupil achievement in design and technology. Education 3-13, 34, 163-172.

"This analysis utilised a previous one made in 1996 (Cross, 1996) when a set of 34 school inspection reports were examined by the author for comments on teaching methods and pupil achievement in design and technology. The results of this earlier analysis were then compared with another conducted in 2004 and with the annual subject summaries of design and technology by HMI based on Ofsted data (Her Majesty's Inspectors, 2002, 2004, 2005). Within each individual school report the section on design and technology was read with comments and judgements noted within categories. These were often comments about what are referred to as 'standards' of pupils' work observed. The outcomes of these analyses were compared to HMI annual subject reports for design and technology. Questions were raised about the basis of these judgements, about possible patterns in pupil achievement and attainment and links between the two." (p.165). Inconsistencies in the language used across reports to make judgements makes it hard to say, for example what percentage of cases fell into certain categories (e.g. wording 'good' or 'satisfactory' and 'attainment' vs 'achievement'). The report suggests that consistency in criteria for judging quality and analysing data (e.g. across time/judged according to a cohort) would be needed to gauge improvements over time.

Day, C., & Gu, Q. (2010). The new lives of teachers. London: Routledge.

Day and Gu report on data collected as part of large- scale, longitudinal, mixed methods research in England involving 100 schools and 300 teachers (Day et al., 2006). They present the potentially negative impact of reforms, including increased oversight through inspection, as reducing teachers' professional autonomy and motivation, illustrating their commentary with case study examples and quotations taken from teachers at different stages of their careers, for example ' I don't feel able to relax now that my teaching is being rigorously monitored by local authority and national school inspections' (teacher quoted in Day and Gu, 2010, p 128). However, elsewhere, in a report based on the same data set, Day and Gu cite Day and Smethem (2009) as noting that 74% of the teachers studied had maintained their commitment to broad educational ideals. Day and Gu (2010) comment on the importance of school leadership in mediating the impact of external reform, of professional development and of the importance of support from colleagues, providing both negative and positive examples of the effects of leadership and the need for leaders to show trust and to involve staff in decision making. Day and Gu further draw on findings in a different large- scale project led by Day, into the impact of leadership on student learning outcomes (Day et al., 2009, 2010) to note the importance of the quality of leadership in mediating 'the boundary between high challenge and high stress' (p 144) and in successfully promoting staff ownership and well- being so that there was 'low staff turnover, high expectations and success for all.' (p144).

de Waal, A. 2006 Inspection, inspection, inspection: how Ofsted crushes independent schools and independent teachers. London, The Institute for the Study of Civic Society

The author makes the overall point that Ofsted inspections have led to a climate of fear that overwhelms schools and teachers. An emotive piece that describes Ofsted as a 'lapdog rather than a watchdog; having a Govt stranglehold; enforcing a top down dictat about how teachers should teach; sabotaging independent school freedoms, thwarting learning; failing the nation; arresting parental choice; intimidating schools into compliance; fostering fabrication of evidence to meet tick boxes and degrading teachers. The writing is strongly opposed to New Labour's approach to centralisation and top down control, which includes private schools.

Two case studies are researched over a period of two years, before and after an inspection. One was an independent 'prep' school in Cambridge with 200 students aged 3-11, with excellent SAT results and another a primary school in a very deprived area of London. There involved interviews with Heads and teachers (unspecified number).

The prep school: the Head decided to make minimal preparations for inspection (to avoid disruption to the school and apparently following official Ofsted guidance), however was criticised for this when inspectors arrived. Despite a history of excellent academic results, the school was criticised for its teaching and assessment methods and records of evidence. Also, resources to help with access to ICT (a new computer suite) and new teaching resources were bought. Demands for paperwork (re-written policies) were met and the curriculum was changed radically in order to become more in line with expectations. Regarding the latter, the Headteacher was surprised to discover that their students were 2 years ahead in some aspects, such as learning about verbs. The poor report led to problems with parent dissatisfaction and falling applications for the following year and significant financial costs (up to£198,000). The head accepted the need for external inspection but thought it too prescriptive; led to more paperwork and too much preparation of inspection. They did however, get a very positive report following the re-inspection and were told they would not be inspected again for six years.

The state school: the deprivation of the intake had several repercussions: many students had poor attendance due to their parents taking them out to visit their country; many parents spoke little English and therefore could not help their children with school work; previous inspections added to pressure on teachers and the leadership and turnover was high. The LEA came in to 'support' but this only added to the relentless pressure and zero tolerance of low achievement in relation to children's backgrounds. Improvement efforts relentlessly targeted aspects they knew Ofsted would be measuring, e.g. borderline children on SATS assessments were intensively tutored. The Ofsted inspection was demoralising and destructive and the Head felt that it did not help the school to improve, rather hindered its rate of improvement, setting it back by months.

The author argues that the heavy level of accountability was undue given that these schools were already accountable to their fee paying customers and that poor provision would be unlikely to survive for very long.

Further study: This involved analysis of inspection reports of eight primary schools in the independent sector. Four schools were interviewed in addition to the scrutiny of the 8 Ofsted reports (published in 2004). An HMI was also interviewed. The independent school report was six pages shorter and covered fewer areas than the state sector version, nevertheless was still highly comprehensive. The Ofsted criteria were considered much narrower compared to the ISI (Independent Schools Inspectorate) version that most schools preferred to undertake. There wa a

perception that Ofsted was trying to force schools to operate in practices that they suggested in the state sector. It was felt that some of these were unsuitable and that many were arbitrary and not evidence based and just a 'fad' (e.g. told how to do 'assessment'). More dictation of what is taught and how, and more paperwork to document everything for Ofsted, was required. The inspection template was perceived as too rigid and independent schools' autonomy under threat. The author also suggests that Ofsted is negating parental choice but forcing schools to make changes that many parents would not agree with.

Chapter 4; A senior member of ISI was interviewed: ISI has a wider scope, as it assesses performance of the school against its own aims and mission. It can allow for recognition of aspects outside its own criteria. In order to become ISI inspectorate-eligible, schools had to join the independent schools council and this was expensive. However, many consider doing this in order to retain more autonomy (although partly surrendering some of this to the ISC). While 'professionally independent' the ISI is under the supervision of Ofsted and the DfE. 10% are co-inspected and 15% of inspections are evaluated by Ofsted and this has a tendency to standardise procedures and hence provision by the schools. Many senior ISI inspectors and also Ofsted inspectors and are thus less likely to be critical of the latter.

Ch 5; Discusses the centralised top down control over teaching that turns teachers nto robots and that this contradicts their intention of creating a world class system, as it deprofeesionalieses teachers. The pressure of the audit society means that even private schools are finding the pressure to standardise in order for their outputs to be more easily measured

Dean, J. 1995. What Teachers and Headteachers Think about Inspection. Cambridge Journal of Education, 25, 45-52.

Study that is concerned with staff perceptions of the process rather than improvement. 5 Las, 1992-1993 , 14 schools, 5 secondary, 3 primary , 3 junior and 1 middle and 2 'first' schools. Semi-structured interviews with 11 headteachers and 37 teachers in nine groups. During a change to the inspection framework, a sample of schools across phases that included some where it was known that the inspection had been a 'disaster' and others average or good. Schools were pre-visit contact was made and where inspectors were reassuring said that this helped allay some anxiety. Some inspectors were seen to lack credibility, e.g. secondary experience inspecting primary schools. Some teachers felt that inspectors influenced the behaviour and contributions of pupils in lessons (more quiet than normal). Teachers did not like it when an inspector left without giving feedback. Teachers approved of inspectors looking at pupil's work.

Dedering, K., & Muller, S. (2011). School Improvement through Inspections? First Empirical Insights from Germany. Journal of Educational Change, 12(3), 301-322.

Dedering and Muller present positive findings on the impact of inspection systems in Germany.   In their survey of 468 recently inspected schools in North Rhine Westphalia, Dedering and Muller found that most schools, whether or not they had a positive or negative report, found it relevant and accurate, with appropriate awareness of the school context and its work. Most principals reported that reports were discussed in leadership teams and school conferences (consisting of representatives from staff, students and parents) and about half in subject conferences. 'Pedagogical' days had taken place, or were planned, in two- thirds of schools to consider aspects of the report. Dedering and Muller suggest that 'developmental activities are taking place at the majority of the schools with 86% of the principals stating that they had already started to put concrete measures into practice.' (p315) and 70% reporting having put in place an improvement

plan with assigned responsibilities although fewer (about half) had associated timelines and only 40% had put introduced their own internal evaluation measures. Overall the report and the process of preparing for this by compilation of school data was seen as helpful and supportive, with both oral and written feedback seen as helpful.

Dedering and Muller also report on two other studies in Germany (not available in English), from Brandenburg (Gaertner et al, 2009) and Bavaria (Huber, 2008). Dedering and Muller say that schools surveyed in both Federal states were similarly positive about inspection. Inspection systems in Federal states were introduced relatively recently and this research reports on schools' first experience of external inspection under newly introduced frameworks. A difference in German systems which is highlighted by Dedering and Muller is the emphasis on the use of inspection to guide the internal processes of the school, with no mandatory publication of inspection reports or test results.

Dobbelaer, M. J., Prins, F. J., & van Dongen, D. (2013). The Impact of Feedback Training for Inspectors. European Journal of Training and Development, 37(1), 86-104.

Dobbelaer et al. used an independent sample experimental design in the Netherlands to investigate the impact of inspection feedback to teachers. 15 inspectors provided feedback to 40 teachers in primary education. Nine inspectors had received short feedback training (the experimental group), while six others did not receive this training (the control group). Their results indicate that feedback provided by trained inspectors can foster professional development of teachers in primary education and that short feedback training has added value. "More than two out of every three teachers acted on the feedback. One-third of the teachers indicated they reflected on their teaching after receiving the feedback. About

20 per cent of the teachers adjusted their teaching methods and another 20 percent did put the suggestions into practice. A total of eight teachers (32 per cent) did not do anything with the feedback. Two teachers indicated they were planning to use the feedback in the near future, but did not have time to do so now. Another six teachers indicated they did not know how to improve their teaching as a result of the feedback" (p.97).

Dougill, P. R., Mike; Blatchford, Roy; Fryer, Lyn; Robinson, Carol; Richmond, John. (2011). To the next level: good schools becoming outstanding. Reading: CfBT Education Trust.

Dougill and his team conducted research into how schools that had been judged to be 'good' continued to improve to 'outstanding'. Based on a review of literature and case studies of eight schools, the researchers identified the following factors, which may be summarised as: a) Senior teams had clear and systematic evidence, rooted in regular lesson observations and teachers' own self-evaluations, on how all staff were performing in classrooms. b) Senior leaders asked all staff to read Ofsted reports on schools which had already been judged outstanding and asked them to identify and apply those practices that could lead to improvement in their own school. c) Professional development for all teaching and support staff ensured that all understood the difference between Ofsted's good and outstanding judgement, in relation to teaching and learning and personal development. d) Schools used self-evaluation as a key strategy. Different groups of staff were asked to focus on particular aspects, reviewing which aspects of practice needed to improve in order to cross the boundary from a grade two (good) to a grade one (outstanding). e) Leaders sought to share with all staff how the inspection process works, making it clear that inspectors focus on students' learning and progress as well as on quality of observed teaching. f) Headteachers believed that action should be immediate following inspection, to ensure benefits to

students were felt quickly. g) Headteachers and governors made use of data to alert staff to students at risk of under-achieving. h) Headteachers focused on eliminating 'in-school variation', for example through middle- leader support and training. I) Schools used partnerships with other schools and education providers to bring good ideas back into their schools. J) Good communication was high priority. k) Governors had high expectations and were prepared to challenge where necessary. I) 'External critical friends' are invited to the school to help celebrate achievement and point out where there is room for improvement.

Education Scotland (2012). Quality and Improvement in Scottish Education: Trends in Inspection Findings 2008-2011 (pp. 48). Livingston, Scotland.

Education Scotland presents findings from Inspection reports from 901 primary, 166 secondary, 34 independent and 112 special schools which were inspected between 2008 and 2011. They note improvement in self- evaluation since the previous inspection report in 2008, with examples of good practice in supporting school improvement in some schools. However the use of self- evaluation to drive improvement is noted overall as an 'aspect for improvement'. They conclude: 'approaches to self-evaluation need to impact on young people's learning and achievements, including their attainment' (p18).

Ehren, M. C. M. And A. J. Visscher (2008). "The Relationships Between School Inspections, School Characteristics And School Improvement." British Journal Of Educational Studies 56(2): 205-227.

Interviews, a questionnaire, observations during the inspection visits, and by analysis of documents such as school inspection reports and school documents. Also, some information gathered in advance, e.g. SIP and innovation capacity of the school. Cases were selected by taking high, low and average capacity for innovation schools and looking at inspectors that were directive, average or reserved.

Data collected before and after an inspection. Whether school improvement resulted or not depended on: the quality of feedback about weaknesses; the assessment of weak points as unsatisfactory; the agreement between an inspector and the school regarding improvement activities. Innovation capacity of the school was not significant.

M. Ehren., and N. Shackleton (2015) Risk-based school inspections: impact of targeted inspection approaches on Dutch secondary schools Unpublished manuscript

Compared changes in student attainment and additional performance indicators in secondary schools who have been assigned to different inspection treatment categories (basic, weak, very weak) following an early warning analysis in 2011. Schools in the basic inspection category are excluded from inspections and receive no visits or any kind of other inspection activity. Schools in the weak and very weak inspection category have received inspection visits in which they were assessed to be failing. These schools have to submit improvement plans which address the inspection standards that are below par and additional monitoring visits are scheduled in subsequent years to assess the implementation of this plan and improvement of their quality. Schools that fail to improve within two years after the first assessment can be put forward for financial and/or administrative sanctions by the Ministry of Education. The target population of secondary schools was therefore set to 454 schools (including both a HAVO and VWO department), of the total of 548 Dutch HAVO/VWO schools. The target sample included almost all HAVO and VWO departments in three different inspection treatments to reach sufficient response rates. Due to the limited number of schools in the 'very weak' inspection category, all schools in this category were included in the sample.

The national non-profit organization 'Schoolinfo' provided us with secondary data on the majority of these schools (266, 88%) at four time points between 2009 and 2013 (the year prior to the early warning analysis, year 1, year 2 and year 3). These data allowed us to measure change in the number of students in the school, parent and student satisfaction, scheduled and taught hours, the number of external evaluations, student achievement on school and central exams, and throughput indicators of lower and upper grades (i.e. number of students progressing without repeating a grade). Data from 2009-2010 acted as a baseline before the early warning analysis of the Inspectorate and the assignment of schools to different inspection categories.

Results:

Changes in student attainment of schools in different inspection categories

The results indicate differences in changes in student achievement results in secondary schools, but only in Dutch literacy. These differences become more prominent over time and are particularly significant in the second and third year after the early warning analysis, suggesting that it takes two to three years for inspection visits to have an impact on student attainment. 'Weak/very weak' schools in the VWO track showed a decline in school exam results in Dutch language compared to (stable) scores in the VWO schools in the 'basic' category, whereas 'weak/very weak' schools in the HAVO track show increasing scores on the central exam results in Dutch language compared to schools in the 'basic' category.

Changes in other performance indicators in schools in different inspection categories

Student satisfaction declined, as well as student numbers and student-staff ratios in weak and very weak schools over time. This would suggest that students are less likely to choose schools that are evaluated as weak or very weak by the Inspectorate, and students in weak and very weak schools become less satisfied when the school is assessed to be failing.

Conclusions: Authors question whether risk-based inspection models are the best way forward. Limited impact on improvement of schools and improvement only seemed to occur in weak and very weak schools. Risk-based models may work better in sectors where 'risk' is more easily measured and predictable (see Ehren, M. C. M., & Honingh, M. E. (2011).

Ehren, M. C. M., & Honingh, M. E. (2011). Risk-based school inspections in the Netherlands: A critical reflection on intended effects and causal mechanisms. Studies in Educational Evaluation, 37(4), 239-248. doi: http://dx.doi.org/10.1016/j.stueduc.2012.02.001

Ex ante evaluation of programme theories of two inspection frameworks

Paper questions the new framework, based on polycentric steering towards improvement. Is the Dutch system mature enough to have schools driving own improvements? Risk assessment based on outcome measures is flawed - i.e. it identifies school already failing, not 'in danger' therefore already too late. Questions whether risk can be adequately assessed in school systems.

Ehren, M. C., Honingh, M., Hooge, E., & O'Hara, J. (2015). Changing school board governance in primary education through school inspections. Educational Management Administration & Leadership, 1741143214549969.

Ehren et al. investigated the impact of 'inspection meetings' in which school boards of schools that have been identified at risk of failing are required to meet with members of the Dutch inspectorate (school boards may be responsible for one or for several schools in the Netherlands). Changes to the inspection method were made to a 'risk based' approach using school data to identify such schools

in 2007. School boards responsible for such schools are required to prepare an improvement plan, which is monitored by the Inspectorate of Education. Using a survey method, they were able to compare responses from schools that had had inspection visits after 2007 with those that had not. They were also above to compare schools that had received a visit in the year prior to the survey and those that had not. According to responses from all school boards, the changes in the inspection method had particularly affected their governance of data use and achievement orientation in schools, the development of quality assurance and self-evaluations in schools and the information they collect. School inspections did not seem to affect their governance of teaching time, curriculum and instruction in schools. Comparative results showed that a focus on data, quality assurance and self- evaluation was more likely in those school boards that had had an inspection visit than in those that had not. Moreover, the data showed a decline in the governance of curriculum and instruction in schools that had had a meeting. Boards responsible for larger numbers of schools reported more changes than those responsible for a small number.

Ehren, M. C., Shackleton, N., Tymms, P., Jones, K., Anane, E, Gustafson, J, Myearber, E. McNamara, G., O'Hara, J., Conyngham, G., Altricher, D., Kemethofer, E., Schmidinger, E., Greger, D. (2015) Technical report EU-project. Year 1-3 data (September-November 2011, 2012, 2013) 'Impact of School Inspections on Teaching and Learning'

 Overall methodology and sampling:

Our study includes a survey to principals in primary and secondary schools in six European countries in three subsequent years (2011, 2012 and 2013) to identify the mechanisms linking school inspections to the improvement of schools.

In each country, an online questionnaire was administered to principals in primary and secondary education. The questionnaire included questions on the intermediate mechanisms of inspection (setting expectations, accepting feedback, promoting/improving self-evaluations, taking improvement actions, actions of stakeholders) and the intermediate outcome variables (improvement capacity, effective school and teaching conditions) in the theoretical framework. Principals scored items on the effective school and teaching conditions in their school and the intermediate processes on a 5-point scale ranging from 'strongly disagree' to 'strongly agree'. Questions about improvement actions refer to actions the school has taken to develop its capacity to improve and specifically to enhance effective school and teaching conditions; questions are framed in terms of the amount of time principals have spent during the previous academic year to improve the school's functioning in these areas (using a 5-point scale ranging from 'much less time' to 'much more time').

Target samples in the Netherlands and England included schools in different inspection arrangements. These schools are expected to respond differently to inspection measures as they face different types of inspection visits and consequences. Schools in the most intensive inspection arrangement (high risk schools in the Netherlands and schools in England receiving monitoring visits) are for example inspected more frequently and have to submit improvement plans to the Inspectorate. Including these schools in the sample provides a broad range of potential responses. The other four countries used a different sampling design as the Inspectorate in these countries schedule regular visits to all schools instead of targeted visits. The designs include either a random sample of schools (Sweden) or use of the TIMSS sampling design (the Czech Republic). Austria (province of Styearia) and Ireland have selected all schools due to the small numbers of schools in the population. See Tables 2 and 3, page 7/8 report for summaries of the target sample and actual response rates.

Austria:

All schools in the province of Styearia were included in the target. Due to restructuring the target sample varied within the three years. Year 1: 503 primary, 194 secondary schools; year 2: 504 primary, 174 secondary schools; year 3: 480 primary schools, 171 secondary schools. This was representative of the target population for school type. Comparing primary and secondary schools: Secondary schools report significantly more activities in the improvement of capacity building including two out of three nested factors (improvement in teacher participation making and improvement in teacher co-operation. The data shows that in two out of three years (year 1 and year 3) primary schools report a significant higher status in capacity building. Primary schools also report more often about a higher status of school effectiveness whereas secondary schools score higher on the improvement of school effectiveness scale including the nested factor (improvement in clear and structured teaching). Comparing inspected versus non inspected schools: Schools with an inspection treatment score significantly higher on two out of three identified causal mechanisms (setting expectations and accepting feedback) with a trend of a higher mean in stakeholders' sensitiveness in year 1. The result of year 1 indicates that schools are more aware of the purposes of school inspection after they have been inspected. As no school inspection took place in the last year of our study we compared schools that were inspected either in year 1 or year 2 with schools without inspection treatment during our study. The results indicate that in the last year there are no differences in school development activities. However, schools that have undergone an inspection treatment significantly report about more stakeholders' sensitiveness and the acceptance of feedback.

The introduction of a new quality framework and the termination of school inspections led to small effect sizes for not inspected schools; they reported an increase in development activities. A small effect size can also be observed for schools which have undergone an inspection two years ago: these schools report less development activities.

Path model

The conceptual model (Ehren et al., 2013) was tested for the first two years. In both years we observe significant effects for stakeholders' sensitiveness and setting expectations on accepting feedback and promoting/improving self-evaluation. Accepting feedback itself has no significant effect on development activities. Self-evaluation has a significant effect on capacity building and improvement of school effectiveness (only year 1). Schools that report of improvement in capacity building also report of more activities in school effectiveness.

Changes in variables over time

The results only indicate a significant time effect of the last school inspection on changes in the means of change of school effectiveness. Follow up comparisons indicated that a pairwise difference between year 1 and year 2 (p = .040) and year 3 (p = .047) was significant. The interaction effect of time*school inspection however was not significant (p = .310). No other significant time effects of school inspections on any of the scales were found.

The Netherlands:

Sample: Three threshold groups (basic, weak and very weak inspection treatment) for both primary and secondary schools. Percentages are percentage of schools in target sample compared to number of schools in the total population:

Primary basic: 208 (3.10%) Primary weak: 152 (41.53%) Primary very weak: 51 (83.61%) Secondary HAVO basic: 321 (77.16%) Secondary VWO basic: 262 (73.39%)

Secondary HAVO weak: 33 (100%) Secondary VWO weak: 91 (100%) Secondary HAVO very weak: 5 (100%) Secondary VWO very weak: 6 (100%). Total target sample: 411 primary schools, 359 secondary school.

Impact of school inspections: cross-sectional analysis

Responses of both teachers and principals to the scales have not been the same in each year, and that changes over time have been different for principals and teachers. Both teachers and principals report few unintended consequences. Teachers versus principals

Teachers report significantly higher scores for accepting feedback and stakeholders' sensitivity to reports. Principals on the other hand report significantly higher scores for the school's capacity and some unintended consequences of school inspections.

Overall we can conclude that there are differences between principals and teachers in how they report about relations between the variables in our model, and about changes in the scales across the three years. The following table summarizes the results of the three analyses.

Differences teachers/principals (t-test)	Differences teachers/principals (longitudinal dataset: only when differences are found for three years)	Differences in reporting about relations between variables: teachers versus principals	Differences in reporting about changes in variables: teachers versus principals
Principals + : school's capacity and some unintended consequences

Teachers +: accepting feedback and stakeholders' sensitivity to reports | Principals + : setting expectations, improvements in capacity-building, improvements in school effectiveness (borderline)

Teachers +: school effectiveness | Principals +: improvements in capacity building year 2, improvements in school effectiveness year 2 and capacity building year 3, stakeholder sensitivity year 1 suggests that

Teachers +: school effectiveness year 3. | Principals: decline in stakeholder sensitivity,

Teachers: decline in setting expectations, increase in stakeholder sensitivity, decline in improvements in capacity building, decline in improvements in school effectiveness |

Primary versus secondary schools

Principals and teachers in primary schools report significantly higher scores for setting expectations, improvement of self-evaluations, improvement of capacity-building, and improvement of school effectiveness compared to secondary schools. Primary schools also report higher scores for the

innovation capacity of the school and the effectiveness of the school, as well as unintended consequences of school inspections.

Inspected versus non inspected schools: Inspected schools report significantly higher scores for setting expectations across all three years. In year 2, inspected schools also report higher scores for accepting feedback, stakeholder sensitivity, improvement of self-evaluations, improvement of building capacity and improvement of the school's effectiveness. Inspected schools in year 1 and 2 also report more unintended consequences, particularly in discouraging teachers to experiment with new teaching methods and the narrowing and refocusing of the curriculum.

Impact of school inspections: changes in schools, and after inspection visits

There is no evidence that principals or teachers in schools of different inspection categories responded consistently differently to the scales in all three years.

There are differences in how schools in the basic and weak/very weak category report about changes in the scales across the three years. Principals in schools in the basic category for example report larger decreases in the setting expectations scale between year 1 and 2 compared to principals in weak/very weak schools. Also the average score of principals to accepting feedback in weak/very weak schools increased from year 1 to year 2.

The results of the longitudinal path models for principals and teachers indicate differences between principals and teachers, where teachers in weak/very weak schools seem to have higher responses to setting expectations as an intermediate mechanisms of improvement, compared to principals.

Surprisingly 'accepting feedback' and 'stakeholder sensitivity' are not indicated as intermediate mechanisms in either of the longitudinal models, whereas 'improvement of self-evaluations' is central in both of the longitudinal models. The results of both models also indicate a link between improvements of self-evaluations, improvement of capacity-building and improvement of school-effectiveness. The association between these different types of improvements however partly disappears when the relations in the path model are controlled for the inspection category (and implicitly also controlling for whether schools had an inspection visit).

Changes over time in student attainment according to inspection category

No evidence of differential changes in Maths A and B, geography, chemistry, Biology, Economics grades on school or on central exams over time by inspection category.

However, grades in Dutch literature differed over time by inspection category.

The average school exam score for pupils in the VWO track changed differentially for schools in different inspection categories. Pupils in schools in the 'weak/very weak' category initially scored higher on the school exams compared to pupils in schools in the basis category of inspection. However the scores declined on the school exams in Dutch language and literature over the three years so that in years 3 they scored less than pupils in schools in the basis category of inspection. Pupils in schools in the basis category of inspection did not show the same decline in scores over time.

There is also some evidence that scores on the average central exam grade in Dutch differed by inspection category over time. There is a significant interaction between inspection category and year of survey for the average central exam grades in the HAVO track. Pupils in schools in inspection category 'weak/very weak' in the havo track initially have lower scores on the central exam than pupils in schools in the basis inspection category. However between year 1 and year 2 this changes

and pupils in schools in inspection category 'weak/very weak' have higher scores. By year 3 there is little discernible difference between schools in the different inspection categories.

Changes over time in other secondary data according to inspection category

Student satisfaction scores changed differently by inspection category over time. Initially student satisfaction scores are very similar in the inspection categories. By year 2 pupils in schools that are in inspection category 'weak/very weak' report significantly lower satisfaction than pupils in schools in the basis category of inspection. By year 3 there is no significant difference between students satisfaction by inspection category of the school.

There is some evidence that in year 3 the number of pupils per full-time employee was significantly lower in schools in the 'weak/very weak' inspection category compared to schools in the basis inspection category.

There is no evidence of differential changes in parental satisfaction, the ratio of pupils to management full time employees, the ratio of pupils to teacher full time employees, the proportion of pupils living in poverty areas, the proportion of sick leave days over time by inspection category.

Autoregressive models

Stakeholder sensitivity in year 1 has a significant positive association with changes in accepting feedback between year 1 and year 2.

There are no associations for changes in improvements in self-evaluations or changes in improvements in capacity building between year 1 and year 2.

However there is an association between improvements in capacity building in year 1, and promoting self-evaluations for changes in school effectiveness.

Whereas higher scores on promoting self-evaluations are associated with decreases in the improvements in school effectiveness between year 1 and year 2, controlling for the influence of the other scales in the model.

Again this may be because schools that report high levels of promoting self-evaluation don't need to take as many improvement actions as schools which report lower scores on promoting self-evaluations in year 1. For school effectiveness promoting self-evaluations is also associated with a decrease in changes in school effectiveness over time. Again this likely reflects a kind of ceiling effect where schools who are already scoring highly on school effectiveness, who are most likely school which do promote self-evaluations can't score much higher on school effectiveness so there is less change in the measure over time.

Stakeholder sensitivity was also associated with larger changes in school effectiveness between year 1 and year 2.

The inspection category of the school in year 2 and year 3 influence the change in accepting feedback between year 2 and year 3. Principals of schools who are in the "Zwak/zeer zwak" inspection category increase their accepting feedback scores by 0.59 points on average. However principals of schools which change from inspection category "Basis" in year 2 to inspection category "zwak/zeer Zwak" in year 3 have a significant reduction in their accepting feedback scores between year 2 and year 3 (-0.66).

Principals in schools with higher levels of stakeholder sensitivity tend to have reductions in the improvements in capacity building scale.

England:

Targeted sample: 211 primary and 211 secondary schools that were closest to the threshold for monitoring inspections (the "treatment" in the England study). Logistic regression models were developed to give good estimates as to how close each school was to the threshold for monitoring inspections.

Data was collected from two main samples. The first sample was selected from primary and secondary schools judged to be "satisfactory" in their main Section 5 Ofsted inspection in 2009/10. Schools were selected that lay close to either side of the threshold for monitoring inspections in order to use a regression discontinuity design. The second sample was also taken from primary and secondary schools that received a main Section 5 Ofsted inspection in 2009/10. A random sample of schools was selected across all inspection judgements.

On all scales in the first two years of data collection, and on almost all scales in year 3, the schools that received their main inspection plus an extra monitoring inspection scored higher on average than the schools that received their main inspection but no further monitoring inspections.

A number of these differences (particularly the scales where schools were commenting on their improvement activities compared to last year) were large and statistically significant in the first year of data collection. It is possible that this is because the non-monitored schools did not receive an extra inspection, whereas the monitored schools were in the middle of another year with an inspection. By the second year the differences are even larger, with more statistically significant differences. It is possible this is due to schools that were monitored continuing to make improvements following the monitoring inspection. Interestingly by the third year of data collection (the second year after the monitoring inspections) there were no statistically significant differences between the two groups.

As expected the outstanding schools judged their schools' effectiveness and capacity as being higher compared to good, satisfactory and iandequate schools. Outstanding schools were also more sensitive to reports compared to other schools and more likely to accept feedback. There was also more evidence of the unintended/side-effects of school inspection in schools with lower Ofsted grades, compared to schools with high Ofsted grades. There were no statistically significant differences between the inspection grades in the year 3 data.

Sweden:

The target sample included a random selection of 1167 primary schools, and 987 secondary schools from the population of 3468 primary schools and 1529 secondary schools.

The results of the different approaches to analysing the data from the longitudinal study of effects of school inspections in Sweden indicate that inspections influence the principals' reports of activities that support learning and instruction. The estimated effect sizes are small, however, and they tend to vary somewhat over the different methods of analyses.

It should also be noted that the growth modelling estimates were computed under the assumption that effects were lasting (i.e., for three years for the schools inspected in 2010/2011, and for two years for the schools inspected in 2011/2012). This thus implies an assumption that effects lasted for two years and there was empirical support for this assumption. It is reasonable to expect that the inspection effect vanishes successively.

There is considerable stability of the outcome variables over time, and the mediating variables Stakeholder Sensitivity to Reports and Accepting Feedback also show strong autoregressive relations over time. These relations create indirect effects of inspection over time, which are equally strong as the direct effects.

The modelling results showed that inspection affected Stakeholder Sensitivity to Reports and Accepting Feedback and that these variables affected Capacity Building both directly and indirectly via Improving Self-evaluations. School Effectiveness was in turn strongly affected by the Capacity Building variable.

The path model thus identifies the two factors Stakeholder Sensitivity to Reports and Accepting Feedback as the main mediators of the effects of inspections on the outcomes. These factors are in the conceptual model identified as two main drivers of the effects of inspections on activities to improve learning and instruction, so these empirical results are expected from the conceptual model, and provide support for the model.

Ireland:

All schools were included in the target sample: 3200 primary schools and 729 secondary schools. In the first year a total 182 schools were received which were usable for the longitudinal study, 129 from primary schools and 53 from secondary schools. For years 2 and 3 of the survey the 182 schools who responded in Year 1 were used as the target sample.

The attitude of principles to inspection as indicated by this research is very positive. As can be seen in the data on virtually every variable, principals report strong support for inspection and are clearly influenced by inspection reports. There appears to be a very good relationship between schools and the Inspectorate.

Principals are very positive about external recommendations clearly believing that they result in a faster pace of educational change.

Since this research began the Inspectorate have changed their schedule of inspection visits to a system where inspection visits should be proportionate and based on the change capacity of the school and the required improvement action needed.

The majority of principals in the surveys seem convinced that inspection results in better management, teaching and learning and that inspection has had an impact on the quality of teaching, learning and management in their schools, in particular where adherence to management and teaching standards is required. In addition, as a result of inspection, principals were also of the view that there was an accelerated rate of change in certain elements of practice.

However, this picture as previously mentioned in effect to a very different time. Since 2012 all primary and post-primary schools in Ireland have been provided with procedures and guidelines for the self-evaluation of teaching and learning (DES 2012). Indeed, as stated by the DES 'Over time the guidelines will be further developed to support schools as they evaluate other key dimensions of school provision' (DES 2012, p.8), with these key dimensions being that of leadership and management and support. This move to a type of inspection with a much stronger data base for decision making not only significantly increases the workload on schools but more importantly will provide the type of evidence on which robust judgements of both school and teacher performance

can be based going forward. Inspection may thus become a more judgemental and invasive process and the very positive attitudes and relationships revealed in this research may be tested.

The Czech Republic

TIMSS design used to frame sample of schools: 150 primary schools and 170 secondary schools. Due to low response rates in Year 3 data collection, we report in this part mainly data from year 1 and 2 for which we have proved above to have a quite good coverage. Where analysis allows, year 3 data are included in overall combined datasets of all three years, or results of analyses are reported shortly.

Descriptive statistics

Discussions taking part in relation to Czech Republic about the model being too overtly focused on compliance and not enough on school improvement. Data here tends to support the view. This is now being the main discussion about the move of the mission of Czech school inspection, that newly elected Chief school inspector since October 2013 is declaring (Greger and Simonová 2014). From that point of view, it will be useful to repeat the questionnaire after few years to see, if changes have been reflected also by school principals. Further analysis presented in other parts however shows, that inspection has a rather low impact on school improvement till date.

Differences between inspected and non-inspected schools

No difference between responses to any of the survey scales. We could thus summarise, that having inspection visit in schools or not does not differentiate principals reporting significantly in Year 1 data. No differences in any of the scales between inspected and non-inspected school were found even in Y2 data. Year 3 data unreliable due to small sample.

Ehren, M.C.M., Altrichter, H., McNamara, G. and O'Hara, J. (2013). Impact of school inspections on teaching and learning; Describing assumptions on causal mechanisms in six European countries. Educational Assessment, Evaluation and Accountability. 25(1), 3-43. http://dx.doi.org/10.1007/s11092-012-9156-4

Figure 1. Intended effects of school inspections Inspection methods, standards, threshold, feedback

Setting expectations	Promoting/ improving self-evaluations	High improvement capacity
Consequences		Good education/ high student achievement
Accepting feedback	Taking improvement actions	Highly effective school and teaching conditions

Public reporting

Actions of stakeholders

(p.14 – full version with boxes).

Using this above model, surveys were given to Principals in these six countries and to teachers in Primary and Secondary schools in the Netherlands. This was over the period 2011-2013. 408 primary schools were included in the target population sample and in each school three teachers from grades 3, 5 and 8. These represented a spread of 'risk', from basic inspection, to weak schools, to

very weak schools. For the secondary schools, the target population was 454 schools. Findings: there was a positive correlation between principals' accepting feedback and setting high expectations. Also, principals who reported higher scores on promoting self-evaluations reported higher scores on average for taking improvement actions in capacity building and school effectiveness. Higher scores on promoting self-evaluations in year 2 was significantly indirectly related capacity building and improving school effectiveness in year 3. Autoregressive modelling (accounting for changes over time) for years 1 and 2 showed that:

"There are no associations for changes in improvements in self-evaluations or changes in improvements in capacity building between year 1 and year 2. However there is an association between improvements in capacity building in year 1, and promoting self-evaluations for changes in school effectiveness between year 1 and year 2. Higher scores on improvements in capacity building in year 1 are associated with larger changes in improvements in school effectiveness between year 1 and year 2. Whereas higher scores on promoting self-evaluations are associated with decreases in the improvements in school effectiveness between year 1 and year 2, controlling for the influence of the other scales in the model. Again this may be because schools that report high levels of promoting self-evaluation don't need to take as many improvement actions as schools which report lower scores on promoting self-evaluations in year 1. For school effectiveness promoting self-evaluations is also associated with a decrease in changes in school effectiveness over time. Again this likely reflects a kind of ceiling effect where schools who are already scoring highly on school effectiveness, who are most likely school which do promote self-evaluations can't score much higher on school effectiveness so there is less change in the measure over time. Stakeholder sensitivity was also associated with larger changes in school effectiveness between year 1 and year 2." (p30).

Between years 2 and 3: Principals for schools in weak or very weak in year 2 increased acceptance of feedback in year 3. However, those that go from 'basis' to weak/very weak in year 3 reduced acceptance of feedback. Principals in weak report greater increases in capacity building and effectiveness in year 3 (perhaps because more potential for improvement). Differences were found between principals and teachers: e.g. setting expectations was higher for principals, although responses to setting high expectations over years 1-3 followed the same declining, then flattening off effect (see graph on p.53). Principals also showed higher scores on improvements in capacity building year 2, improvements in school effectiveness year 2 and capacity building year 3 than teachers. For teachers, stakeholder sensitivity remains stable between years 1 and 2 but increases in year3 whereas this falls for principals in year 1-2 and then flattens off. Improvements in self-evaluations fall from year 1-3 for both groups. Improvements in capacity building remain higher and fairly stable for principals but decline between years 1-3 for teachers. Increase in school effectiveness declines over time for both groups (overall higher for principals).

Inspected versus non inspected schools: inspected schools report higher scores on setting high expectations (across all years); I year2, inspected schools are higher on "accepting feedback, stakeholder sensitivity, improvement of self-evaluations, improvement of building capacity and improvement of the school's effectiveness. Inspected schools in year 1 and 2 also report more unintended consequences, particularly in discouraging teachers to experiment with new teaching methods and the narrowing and refocusing of the curriculum" (p86). There was no consistent pattern of difference between schools in the different categories (weak, very weak, basis).

Accepting feedback and stakeholder sensitivity are not seen to be linked to improvements whereas improvements in self-evaluation are significant for capacity building and school effectiveness. Central exam scores for students indicate that weak or very weak schools are likely to improve to the level of the 'basis' schools by the 3rd year. Student satisfaction levels initially decline in the

weak/very weakschools but recover in year 3 to match the basis schools. Student numbers in weak/very weak schools decline over 3 years (71 fewer on average). There are no differences on a range of other variables, such as: parental satisfaction, the ratio of pupils to management full time employees, the ratio of pupils to teacher full time employees, the proportion of pupils living in poverty areas, the proportion of sick leave days over time by inspection category (p.89).

Ehren, M.C.M., Gustafsson, J.E., Altrichter, H., Skedsmo, G., Kemethofer, D., Huber, S.G. (2015). Comparing effects and side effects of different school inspection systems across Europe. Unpublished manuscript

In this article different inspection models are compared in terms of their impact on school improvement and the mechanisms each of these models generates to have such an impact. Our theoretical framework was drawn from the programme theories of six countries' school inspection systems (i.e. the Netherlands, England, Sweden, Ireland, the province of Styearia in Austria and the Czech Republic). It describes how inspection models differ in the scheduling and frequency of visits (using a differentiated or cyclical approach), the evaluation of process and/or output standards, and the consequences of visits, and how these models lead to school improvement through the setting of expectations, the use of performance feedback and actions of the school's stakeholders. These assumptions were tested by means of a survey to principals in primary and secondary schools in these countries (n=2239). The data analysis followed a three step approach: 1) Confirmatory factor analyses 2) Path modelling and 3) fitting of MIMIC-models. The results indicate that Inspectorates of Education that use a differentiated model (in addition to regular visits), in which they evaluate both educational practices and outcomes of schools and publicly report inspection findings of individual schools, are the most effective. These changes seem to be mediated by improvements in the schools' self-evaluations and the school's stakeholders' awareness of the findings in the public inspection reports. However, differentiated inspections also lead to unintended consequences as Principals report on narrowing the curriculum and on discouraging teachers from experimenting with new teaching methods.

"Additionally we used the year 1 data to test for differences in impact between countries with different inspection models. Differences in inspection models we studied include the use of differentiated inspections, where potentially failing schools are targeted for increased inspection visits (in contrast to regularly scheduled visits to all schools), in inspections of school outcomes (in addition to only evaluation of educational practices in schools or compliance to legislation), in inspections with or without punitive sanctions, and in the presence or lack of public reporting of inspection outcomes of individual schools.

These different inspection models may 'work' differently in generating impact and may vary according to the degree of influence on school improvement as well as on various unintended consequences. This paper presents the results of a survey of principals in primary and secondary education in six European countries (the Netherlands, England, Ireland, Sweden, the Czech Republic and Austria) on the impact and mechanisms of impact of these different school inspection models. We used multiple indicator and multiple cause modelling (MIMIC) to compare the impact of various inspection approaches as described above.

The results of our study indicate that Inspectorates of Education that use a differentiated model (in addition to regular visits), in which they evaluate both educational practices and outcomes of schools and publicly report the inspection findings of individual schools, are the most effective. Principals in these systems report the most changes in capacity-building and in improved school and teaching conditions. These changes seem to be mediated by improvements in the schools' self-

evaluations and the awareness of the school's stakeholders of the findings in the public inspection reports. However, differentiated inspections also lead to unintended consequences since Principals report a narrowing of the curriculum in the school and the discouragement of teachers from experimenting with new teaching methods.

An interesting issue emerges when it comes to the 'causal mechanisms' by which inspection systems intend to produce their effects. Three of the four inspection models we studied (differentiated inspections, outcomes-orientation and sanctions, which may well be taken to indicate 'high stakes' inspection approaches) influence 'setting expectations'. These models, at the same time, reduce the likelihood that principals pay attention to the inspection feedback and derive action strategies for school improvement based on this feedback. " (p.28/29 Ehren et al 2014 technical report)

Additionally, the paper compares the soft and hard governance approaches of Austria and England respectively. The soft approach of Austria led to fewer self-evaluation and development activities, however, feedback was more accepted and led to fewer unintended consequences.

Ehren, Melanie C. M. & Shackleton, Nichola (2015): Mechanisms of Change in Dutch Inspected Schools: Comparing Schools in Different Inspection Treatments, British Journal of Educational Studies

This article aims to enhance understanding of the connections between school inspections and their impact on school improvement, using a longitudinal survey of principals and teachers in primary and secondary education. Random effects models and a longitudinal path model suggest that school inspections in particular have an impact on principals, but less so on teachers. The results indicate that the actual impact on improved school and teaching conditions, and ultimately student achievement, is limited. Schools in different inspection categories report different mechanisms of potential impact; the lack of any correlation between accepting feedback, setting expectations and stakeholder sensitivity and improvement actions in schools suggests that the impact of school inspections is not a linear process, but operates through diffuse and cyclical processes of change.

Data and sample: This article reports findings from data collected using methods reported in an earlier technical report (Ehren and Shackleton,) . A survey was completed by principals and teachers in Dutch primary and secondary schools in three consecutive years (September–December 2011, 2012 and 2013) with schools sampled according to the categorisation of the inspectorate as 'basic', 'weak' and 'very weak'. This looked at:

setting expectations and institutionalisation of norms;
accepting and using feedback;
sensitivity of stakeholders to inspection reports (voice, choice and exit).(p 6)
and to identify the mechanisms linking school inspections to the improvement of schools.
The findings lead the authors to the following conclusions:
(1) The results indicate different mechanisms of potential impact for schools in different inspection categories: potential improvement from school inspections in the 'basic' inspection category appears to result from the setting of expectations and the preparation and improvement in self-evaluations, openness to inspection feedback and sensitivity of stakeholders to inspection reports in the year of, and after, the early warning analysis. Weak and very weak schools show a pattern of impact through an increase in openness to, and acceptance of, inspection feedback and increasing changes in the schools' self-evaluations and capacity building over the years.

(2) School inspections appear primarily to have an impact on principals rather than teachers.

(3) The results indicate that the actual impact on improved school and teaching conditions is limited. However, as such effects are more likely to take effect after a period of time longer than the three years of data collection, these improvements may potentially be beyond the scope of the study.

(4) We find few unintended consequences resulting from school inspections. Only principals in 'weak/very weak schools' reported sending documents that presented a more positive picture of the school to principals.

 (5) The lack of any correlation between accepting feedback, setting expectations and stakeholder sensitivity on the one hand, and improvement actions in the schools on the other, also suggests that the impact of school inspections is not a linear process, but operates through diffuse and cyclical processes of change. (p25)

Ehren, M. C., & Swanborn, M. S. (2012). Strategic Data Use of Schools in Accountability Systems. School Effectiveness and School Improvement, 23(2), 257-280.

Ehren and Swanborn investigated the impact of the use of pupil performance data in order to estimate the extent to which schools manipulate or 'cheat' the measures of reported performance. The first comparison was made by comparing data sets taken from just before and just after the inspection system in the Netherlands was changed. Prior to 2007, all schools were inspected with equal regularity, but after this date, pupil performance data was used to identify those schools which were most likely to need improvement in educational quality and these schools were more likely to be inspected. The second comparison was between schools judged in inspection to have high educational quality and those with low educational quality and the third comparison was between high and low scoring schools (schools with low performance data from two successive years are considered to be 'at risk').   Data was available on the extent of cheating in the administration of tests and on the numbers of pupils excluded from published results (reshaping the test pool). However, although Ehren and Swanborn found that 14 schools (approximately 5%) did not comply with guidelines for test administration and one third of schools excluded one or more pupils from tests, there did not appear to be a relationship between test manipulation and either the increase in likelihood of inspection or measures of school quality.

Ehren, C. M. & Shackleton, N. 2014. Impact of school inspections on teaching and learning in primary and secondary education in the Netherlands. London: Institute of Education. (not in spreadsheet as data is detailed in Ehren et al (2013) Impact of school inspections on teaching and learning; Describing assumptions on causal mechanisms in six European countries.)

This study focuses on the effects and negative consequences for school inspections in six European studies (England, Ireland, Sweden, Austria/Styearia and the Czech Republic), and three years of data collection of primary and secondary schools in the Netherlands. The mechanism for improvement of teaching and learning was taken from each country's inspection framework, and this was analysed through interviews with inspection officials and analysis of policy and inspection documentation. The conceptual model derived from this analysis was presented in more detail in Ehren et al (2013) Impact of school inspections on teaching and learning; Describing assumptions on causal mechanisms in seven European countries.

Ehren, M., Perryman, J. & Shackleton, N. 2014. Setting expectations for good education: how Dutch school inspections drive improvement. School Effectiveness and School Improvement, 1-32. (not in spreadsheet as data is detailed in: Ehren et al (2013) Impact of school inspections on teaching and learning; Describing assumptions on causal mechanisms in six European countries.)

Discussion sections adds: Setting expectations and stakeholders sensitivity are important factors in school improvement from inspections: the inspections that set expectations about standards in education can drive school self-evaluation, and the capacity of a school to improve, through cooperation with teachers, transformational leadership and teacher participation in decision-making. Changes in self-evaluation lead to improvements in capacity building which leads to improvements in school effectiveness. Schools that use inspection standards are also more able to accept feedback however the latter is not a major factor in school improvement. Setting expectations and stakeholder sensitivity are indicators of accepting feedback. Unintended consequences including "gaming and window dressing and teaching to inspection" (p. 24) were relatively infrequent according to principals and teachers. Schools that were high on Setting expectations, accepting feedback, and stakeholders' sensitivity were also more likely to narrow the teaching strategies and curriculum; send a more 'rosy' picture of the school in documents sent to the inspectorate and to have special protocols I place for inspection visits.

The results suggest confirmation of 'neo-institutional' theories to explain how schools seek legitimacy from external inspections as keys to their survival and as a benchmark of their high performance. One risk of this is to increase extrinsic and decrease intrinsic motivators of improvements. Also: "Schools copy practices of other schools who successfully meet the expectations of the Inspectorate to increase their legitimacy and prevent potential action and pressure by external actors, such as the Inspectorate or the school's stakeholders" (p. 25). However, the govt and inspectorate in this model are also seen as only one of multiple actors, and a 'polycentric' form of steering is taking place. Within this network of actors, the principals are key 'linking pins' (p. 26). More research is needed to understand the process by which inspection standards are translated into school (structure and processes) and education systems.

Erdem, A. R. & Yaprak, M. 2013. The Problems That The Classroom Teachers Working In Villages And County Towns Confront In Educational Inspection And Their Opinions Concerning The Effect Of These Problems On Their Performance. Educational Research And Reviews, 8, 455-461.

This is taken from a doctoral thesis. A survey was administered to 321 class teachers working in official primary schools in townships of Denizli and 272 class teachers working in official primary schools in villages of Denizli. Likert scale items invited responses about the scale and frequency faced by teachers in terms of educational inspections. The most significant finding was that teachers 'abstained' from telling inspectors what professional attributes they needed to improve. This was the case in both town and village settings. The failure of inspectors to engage in effective and open dialogue or to show an interest in teachers' development were seen as key aspects. Criticising teachers in the presence of students was given as one example of how the negative relationship prevented such dialogue.

Follows, M. 2001. After the storm the tale of an Ofsted inspection. Management in Education, 15, 24-28.

Follows, M. 2001. Learning from the tale of an Ofsted inspection. Management in Education, 15, 32-34. (not in spreadsheet as same data set as above).

An imaginative piece and a follow up piece (that helps explain the first one!) using a fictionalised narrative, where characters are referred to as animals in a childrens' story. The author is a headtacher of an infant school who explains the feelings of betrayal and powerlessness that she feels as a result of an Ofsted inspection. She later points out that others had contacted her about

similar experiences, suggesting that it is not uncommon to feel this way. Does not add much empirically but highly original!

Francis, B. (2011). (Un)satisfactory? Enhancing life chances by improving 'satisfactory' schools. London, Royal Society for Arts and Manufactures (RSA).

Francis produced her report with the cooperation of Ofsted, who supplied data about inspection judgements in successive inspections for all secondary schools in England and a sample of inspection reports from secondary schools that had received 'satisfactory' ratings in two successive inspections with 'satisfactory' capacity to improve. The data from all inspections showed that such schools were more likely to be found in areas where children faced multiple disadvantages, with implications for educational and social equity. Of the 937 schools judged to be 'satisfactory' in a previous Section 5 (2005-2009) inspection and that had been inspected since, 50% remained 'satisfactory' and 8% had declined to 'inadequate'. (i.e. 42% had improved to good or outstanding). The 1034 schools previously judged to be 'good' or 'outstanding' tended to stay there, with 19% declining to 'satisfactory' and 2% to 'inadequate'. Their analysis of inspection reports for the sample of 36 schools from those stuck at 'satisfactory' showed that the strongest factor was inconsistent quality of teaching and assessment practice within 'satisfactory' schools, although contextual challenges, such as falling rolls or difficulties in staff recruitment, together with weaknesses in leadership, governance and school self- evaluation were also common. Francis comments that although the inspection reports highlight what needs to be done, there is little guidance on how to do it and the report recommends that a range of support measures are implemented to help such schools improve.

Gaertner, H., Wurster, S. & Pant, H. A. 2014. The effect of school inspections on school improvement. School Effectiveness & School Improvement, 25, 489-508.

The introduction to this study has a good overview on the weaknesses of the empirical evidence about the effect of inspections on school improvement. They try to address this by using repeated measures on schools that are inspected compared to those not inspected. Both process aspects and student attainment measures were used. The former included perceptions of heads and teachers, that the authors cite (based on MacBeath 99 and 2001) as important aspects that determine self-directed school improvement. The introduction of school inspections in Berlin and Brandenbury formed the focus of the sample. These inspections look at a full range of data and include visits (similar to Ofsted) but an overall grade is not given and the report is used to set agreed targets (between the local authority and the school) although the latter can include aspects other than the inspection report. The next inspection is after 5 years or 2 if the standards of the school are seen to be low. The report was not made publicly available at the time of the article. The sample:

● Schools that were inspected in the 2006/7 school year; that is, 2 years before the

first survey (S1).

● Schools that were inspected in the 2007/8 school year; that is, 1 year before the first

survey (S2).

● Schools that were inspected during the survey period (S3).

● Schools that had not yet been inspected (Control Group; CG).

Thus the time frame covered a three year period. Schools are inspected randomly, therefore the control group can be seen as random too. Surveys were sent to Principals and teachers who were

members of the school council (these were involved in school decision making). Primary school pupil attainment came from standardised tests in maths and German in grade 3 and for secondary schools it was school-leaving examinations at the end of grade 10. T tests for independent samples (and effect sizes) showed no overall difference in school achievement.

There were 16 dependent variables related to school quality. These are based on evidence on school effectiveness by Scheerens and Bosker, 1997 and the inspection framework is based on these. The questionnaire and inspection judgements were highly correlated. Principals and teachers had a stable view of the quality of their schools over time (survey repeated over 1 year period) and this did not vary according to whether they had been inspected. They conclude that "an inspection does not produce fundamentally new knowledge, but rather officialises known problems" (p 503, referring to Landwehr 2011). The writers also speculate that (based on other research, e.g. Plowright 2007) that school improvement comes in preparation for the inspection not after. They state that this would mean the inspection takes the fourth function of Lanwehr's model, i.e. the enforcement of standards. However, the process of target agreement with schools in the sample was new and may not have had time to lead to improvements.

Gray, C. & Gardner, J. 1999. The Impact of School Inspections. Oxford Review of Education, 25, 455-468.

This study surveys principals at 38 secondary and 38 primary s schools on the impact of pre/during and post inspection. These were taken from a stratified random sample of recently inspected schools (n=130), therefore 58% response rate. The survey was initially generated from interviews with 5 principals. Findings:

Pre-inspection: 34% primary and 21% of secondary school principals (probably lower since there were more focused inspections in this group) felt that most staff experienced high levels of anxiety. There was a positive correlation between levels of anxiety and length of time spent preparing for the inspection. 63% of Primary and 42% secondary complained that the inspection disrupted the normal running of the school. After the visit: 59% said the most valuable feature of the inspection was to externally validate the good practice provided by the school (affirmation). 12% said it highlighted areas to change and 11% said the process enhanced cooperation and team building. 13% felt that stress and anxiety had a negative impact and months later teachers were complaining of tiredness. 8% reported they would implement all the recommendations of the report, 28% said no change and 64% identified one area of change, of which, whole school planning (30%) and classroom teaching (22%) featured the highest.

Relationships with inspectors were crucial, those with inspectors described as aloof and imposing experienced high levels of anxiety.

Greger, D., Huber, S., Jones, K,L., Kemethofer, D., Myearberg, E., O'Hara, J., Skedsmo, G., and Tymms, P. (2015) The unintended effects of school inspection: The prevalence of inspection side-effects in Austria, the Czech Republic, England, Ireland, the Netherlands, Sweden and Switzerland. Unpublished manuscript

"In order to contribute to the limited empirical research in this area, we analysed responses from school principals in seven countries, with differing school inspection systems to measure the prevalence of these unintended consequences and to investigate the part pressure (to do well in inspections) has to play.

Findings from the survey are mixed, with varying prevalence for each unintended effect and varying association with pressure. As expected, the results showed a clear association between increasing pressure in a school inspection system (in terms of pressure to do well in an inspection) and an increase in the narrowing of the curriculum and instructional strategies in the school. Similarly, as the inspection systems increase in pressure there is an associated increase in principals claiming that school inspections have resulted in a refocussing of the curriculum and instructional strategies. It is just as interesting to find that unexpectedly some unintended consequences do not seem to be linked to how pressurised the inspection system is. It is not clear why this is.

Some unintended consequences, such as the discouragement of experimenting with new teaching methods and the misrepresentation of the school in the data sent to the inspectorate, do not seem to be a prevalent as previously thought, although a number of principals do admit to these practices. It is possible that this is due to principals not wanting to admit to behaviour that is perceived as negative. The prevalence of formalisation and proceduralisation is much more evident across all countries whereas the prevalence of inspection narrowing and refocussing of the curriculum and teaching strategies is, as already mentioned, dependent on how pressurised the inspection system is.

A key finding is that increasing pressure in school inspection systems is associated with a narrowing and refocussing of the curriculum and instructional strategies. Unexpectedly we find no evidence that pressure is related to the prevalence of misrepresentation of data, formalisation and proceduralisation (excessive focus on records) and ossification (fear of experimentation in teaching)." (taken from EU technical report Ehren and Shackleton 2014, p31/2).

Grek, S., et al. (2013). "Governing by inspection? European inspectorates and the creation of a European education policy space." Comparative Education 49(4): 486-502.

Data collection: first project phase, which was focused on transnational influences on inspection, with particular attention to the agenda setting and policy learning capacities of SICI. Our data included interviews with key 'system actors' at the international level to explore how the inspectorate responds to, and influences, transnational agenda-setting by OECD and the EC (10 interviews). There were also interviews with actors at the national level in all three systems (30 in total). Compares Sweden, Scotland and English models. Paper suggest growing cooperation and learning across states in Europe, shown in growth of activity of SICI. Scottish system most aligned with learning and meditative intention (Jacobsen) while English and Swedish still quite inquisitional (audit) focus.

Hardy, I. (2012). "Managing" Managerialism: The Impact of Educational Auditing on an Academic "Specialist" School. European Educational Research Journal, 11(2), 274-289.

Hardy reports on a case study of one academic, specialist school in England in which he conducted interviews with 18 members of staff about performance management processes, as well as Ofsted inspections. The interviews followed shortly after an inspection. Positive effects were that it reinforced expectations about teaching practice and the need to be ready for Ofsted inspection meant that administrative frameworks (for example for lesson planning) and record keeping were kept up to date. The principle of being open and accountable was seen as fair and reasonable. Concerns were expressed about the time- consuming element of inspection and preparation for this, and with regard to 'performativity' in lessons taught during an inspection visit. However, Hardy quotes one teacher as saying 'And the fact that it was a show for Ofsted is not to say that the quality of the learning that was happening was anymore; I don't necessarily think it was. But it was

definitely being presented in a different way, just to jump through hoops.' (p281). Concern about exam results constrained pedagogical practice and 'substantive student learning in general' (p280). There was anxiety among staff and particularly among students during inspection, with students feeling that they were responsible for success or otherwise. Ofsted influenced target setting processes so that highly aspirational targets were set for student attainment. A sense of failure was experienced by teachers interviewed whose students did not meet these. There was also a tendency for teachers to focus on students near pass/fail borderlines and to focus on practices to improve results rather than student learning overall.

HMIE. (2010). Learning Together: Lessons about School Improvement--An HMIE Report on How Schools Get Better. Second Edition: Her Majesty's Inspectorate of Education.

This report is mainly based on an analysis of the progress found in over 300 HMIE follow- through inspections on which reports were published between 2005 and 2008. Follow- through inspection in Scotland follows the identification of weaknesses in key aspects of the school and in the research period 20% of primary schools, 29% of special schools and 33% of secondary schools were identified as in need of follow through. In most cases, the schools undergoing follow-up inspection had improved and the report presents the factors which led to improvement. The report identifies the quality of leadership and the quality of the school's self- evaluation processes as most significant. Weaknesses in these aspects indicated low capacity for improvement. Other important factors were the extent to which pupils' learning needs were met and the quality of their learning experiences, together with the quality of curriculum and learner attainment. The report found that the process of inspection by HMIE had prompted actions to improve the schools and that staff had felt that inspectors had helped in identifying improvement priorities and an increased sense of urgency about the need for change. Acceptance by staff of the weaknesses identified either self- evaluation or inspection was essential for improvement and the report suggests that it is therefore essential that all staff are involved in gathering and analysing evidence and in using benchmarks and comparisons to identify strengths and weaknesses. Challenge and support, from external stakeholders and from the local authority and their quality improvement officers are important, as is training and development for senior leaders. They conclude that increased effectiveness of leadership at all levels in the school is a key factor in improving the outcomes for learners.

Hogenbirk, P., & Braak, P. (2013). ICT action school development at Helen Parkhurst Dalton School. Education and Information Technologies, 18(2), 233-243.

Hogenbirk and Braak provide a detailed case study from a school in the Netherlands that underwent inspection from a group of European inspectors in relation to their use of ICT in 2007, using a European framework devised for this purpose. The practitioner case study provides a detailed description of how action plans were formulated and implemented as result of findings from the inspection, leading to improvements in student learning, indicating a clear sense of ownership among staff at the school.

(interesting issue in England with declining role of LEA due to Academies).

Hopkins, D., Harris, A., Watling, R. & Beresford, J. 1999. From Inspection to School Improvement? Evaluating the Accelerated Inspection Programme in Waltham Forest. British Educational Research Journal, 25, 679-690.

This covers a period in two London LES's (Waltham forest and Lambeth) who were subjected to a period of Accelerated Inspection (AIP) on the basis of poor performance from an initial sample of

inspected schools in a report published in 1995 by Ofsted. 40% of school in these Boroughs were identified as having 'serious weaknesses'.

Research questions:

What has been the effect of the decision to conduct an accelerated inspection upon schools and the LEA?; and

• How effectively have schools and the LEA responded to this decision?

The evaluation also looked at the mechanism for individual school improvement.

Five data collection methods:

postal questionnaires (3 times) these were sent to all schools affected by the AIP 1 week before the inspection, the week following the inspection and when the school's action plan had been completed;
interviews conducted with staff at 13 schools [3];
interviews with staff from OFSTED and the LEA;
an independent analysis of data generated by OFSTED, the LEA and a sample of schools in the borough; and
an independent rating of schools' self-evaluation and capacity for improvement
(Hopkins et al., 1996).

Pre-inspection effects:

School-Ofsted - The suspension of many aspects of school activities
Media coverage 'a blitz on poor teaching' etc., ramped up the pressure on schools of AIP
LEAs – Schools: found it difficult to work with schools under so much pressure and previous relationships had suffered, partly due to funding problems. They carried out health checks of all schools which were seen as imperfect but useful
Too little help, too late (or too much too late)
During:

Schools- Ofsted:
Great pressure on inspection teams to cover so many inspections in a short time
The school had less time to worry
Inconsistencies with report writing and turnaround times were slow for reporst (some waited 20 weeks)
Post:

Schools- Ofsted: immediately after, great relief and reduction in stress, positive atmosphere to school and revision of school development plans. Reaction to report depended on how critical it was: confirmation or demoralisation.

However, even positive reports led to concerns about the direction that Ofsted were steering the school and the type of education it provides.

LEAs – Schools: LEA took up new critical friend role and revived 'lapsed' relationships with schools.

Recommendations:
preceed inspections with self-evaluation, Ofsted should advise on how to draw up an action plan and promote SSE
LEA: Use Ofsted databases to provide benchmarks for schools.

Schools: Find links between Ofsted action plans and broader SDP; work with LEA to establish networks to share good practice

Conclusions:

Post Ofsted there was less energy for change. Ofsted provides good initial diagnosis but this is not enough for SI.

Need to adopt differential strategies depending on the growth state of the school

The role of the LEA was seen as key mediator to support SI, this was:

interpreting Ofsted reports and facilitating post-Ofsted action planning;
target setting, including being able to frame targets in terms of pupil learning outcomes;
monitoring, analysing and interpreting outcome and process data;
benchmarking;
negotiating, priorities, time-scales and participation; and
developing and refining differential strategies for schools in various growth states

Hussain, I. (2012). Subjective performance in the public sector: evidence from school inspections: London School of Economics and Political Science. Centre for Economic Performance.

Husssain analysed data from inspection reports and compared these with data on student attainment and student socioeconomic composition. He also used data from longitudinal student surveys and parent surveys. He found a high degree of agreement with student and parent judgements about teacher practices and with the inspection gradings given to schools with the same test rankings and socioeconomic composition of students. He suggests that this indicates that the inspection ratings provide reliable information about school quality over and above that already contained in publicly available information on test results. Measures used for comparison were that 'students at higher rated schools experience an environment where teachers are more likely to: take action when a student breaks rules; make students work to their full capacity; keep order in class; set homework; check that any homework that is set is done; and mark students' work' (p12). To test for the impact of a 'fail' grading on students' test results, Hussain looked at the performance of primary schools that failed an Ofsted inspection in September, 2005 in tests undertaken by pupils in May 2006. These schools have thus nearly the whole of the school year to respond to inspection findings. A control group was provided by looking at the test results for tests taken in May 2006 for those primary schools that failed an inspection in June 2006. These schools were thus inspected before test results were known and whose activities during the school year 2005-6 were not influenced by inspection. The exercise was repeated for the years 2006-9, with approximately 120 schools in each year. Hussain found improvements in test scores for those schools that had a 'severe' fail (special measures) resulting in external intervention and support and those schools that had a 'mild' (notice to improve) report, with no external intervention or support. Hussain also examined improvement in relation to prior attainment, to control for 'gaming' by schools, for example by failing to enter pupils less likely to perform well or by targeting borderline pupils. He found no evidence to suggest such gaming and found improvement for all pupils in the schools studied. Furthermore, the improvement in student attainment was found to be maintained in student data for the following three years.

IpsosMori (2008). School inspections - final report 2008: Office for Standards in Education, Children's Services and Skills.

Surveys of parents of school- age children were commissioned by Ofsted in 2006 and 2008 from Ipsos MORI and indicate that, although parents are in favour of school inspection, only a minority think that it provides them with useful information. The 2008 reports states that:

As in 2006, the overwhelming majority of parents (92%) say they are in favour of school inspections, with just 4% who are not in favour... One in three (34%) say that inspections help schools to improve, 30% mention that it keeps schools from becoming complacent, 16% think it provides them with useful information, and 12% say it helps to protect children's interests. Just three in ten parents (29%) do not give any reason. In general, those whose child's/children's school has been inspected recently are more likely to be in favour of inspections (96% versus 92%). (Ipsos MORI, 2008, p3)

Jeffrey, B. & Woods, P. 1996. Feeling Deprofessionalised: the social construction of emotions during an OFSTED inspection. Cambridge Journal of Education, 26, 325-343.

This ethnographic study followed three months prior to an inspection and up to a year afterwards. Observation, semi and informal interviewing of staff, examining documents teachers consulted approx. 10 times. The aim of this approach was to uncover the meanings of emotional responses to the inspection process rather than just to state what these were. In terms of its relevance to school improvement there are a number of issues: the overwork and stress that lead up to the inspection, the inability to leave the threat of Ofsted behind at home, and anxieties about iandequacy. Several meanings emerged about these emotional responses, including: the dehumanisation process (e.g. whether to shake the inspector's hand when he left); the invasion of 'Ofsted' work into home life; the feeling of professional iandequacy- feeding into teachers' pre-existing thoughts that you could always do more; the lack of humour in lessons and between staff on the day of the inspection. Also, there was a 'loss of pedagogic values' – described as a grieving process, teaching being reduced to numbers, percentages that were 'satisfactory'; people feeling persecuted and guilty through the exercise of bureaucratic controls. After the inspection there was a feeling of emptiness and irritation. When the report came it made staff re-live the experience (negative) of the inspection process. The loss of confidence in teachers' professional role is mentioned and that teachers may need to 'redefine the self' (p.341) to fit in with this new version of professionalism; see teaching as more instrumental (rather than moral) or to leave the profession entirely.

Jones, K. & Tymms, P. 2014. Ofsted's role in promoting school improvement: the mechanisms of the school inspection system in England. Oxford Review of Education, 40, 315-330.

This article explores the underlying mechanisms for school improvement in Ofsted's inspection system and proposes a program theory (Leeuw, 2003), specifically the policy scientific approach (Ehren, Leeuw and Scheerens, 2005) to explain this (also to allow evaluation). The study here, forms part of the European Commission's Lifelong Learning Project, 'The impact of school inspection on teaching and learning'. Six countries are compared, in order to maximise the effect on school improvement and to minimise the unintended consequences. Focuses on inspections in maintained schools in England in the age range of 4–16 years (Reception through to Year 11). Three research questions explored:

What are the mechanisms by which Ofsted's actions are expected to improve the education of children and young people?
How consistent, precise and realistic are the mechanisms regarding the expected effects of inspections on promoting school improvement?

How likely are inspections to successfully promote school improvement

The programme theory created (from documents and interviews with Ofsted) was for the inspection framework from September 2009 until August 2011. The programme theory (see p325 for whole chart), sets out 5 aspects: 1. Setting standards, 2, giving feedback, 3. Sanction and rewards, 4 collecting information and 5 public accountability.

Evaluation of the programme theory:

The theory was evaluated in terms of logical consistency and completeness, and also 'realism' ie. Finding evidence to support assertions in the theory.

Consistency: the authors point to the 'nudge' mechanism implied by Ofsted which 'nudges' actors (schools, parents and policy makers) into improving schools. The authors suggest that market mechanism may not lead to schools working on the feedback given and that some may improve and others not. However, generally this is plausible to improve most.

Completeness: the authors suggest that it has no significant gaps.

Realism:

The evidence for mechanisms in the pr.theory is too vast to undertake in the review. The authors call for more research on the mechanisms.

E.g. the conditions underlying whether feedback leads to improvement. This example shows the complexity and no simple solutions. However, he points to one review by Coe (1998) that says that there is limited empirical evidence on the link between feedback and improved performance. Also, Helen and Timperley (2007) show how feedback can lead to improved pupil learning but also can have negative outcomes. More is needed on feedback from inspectors, the authors suggest and also feedback on school level. The report also highlights the mixed findings from key research into the effect of inspections on pupil attainment (mixed findings) and also evidence of unintended effects, such as teaching to the test and short termism.

Keeble-Allen, D. 2004. Inspection at Summerhill: did Ofsted inspection result in improvement? U210580 Ed.D., University of Leicester (United Kingdom).

Author questions "whether an OFSTED inspection of an atypical independent school (Summerhill) is able to make appropriate judgements about that school given an inspection methodology based in modernist constructs and systematic observation." (p.3). Summerhill was inspected in 1999 and the report suggested there were repeated failure to address issues, such as lack of attendance to English and Mathematics classes (attendance was not compulsory), in relation to previous inspections in 1990, 93 and 98. Ofsted had concerns over levels of literacy and numeracy and lack of supervision in overnight accommodation. The idea of assessing students' needs clashed also with Summerhill's philosophy. In a legal appeal, based on Article 12 of Convention on the Rights of the Child, parents stated that they were against traditional schooling but did not want to deny children the right to participate in a community, thus sent them to Summerhill; this right was being denied. Methodology: Case study. Researcher attended 'visitor days' to get a view of the school (similar to Ofsted); documentary analysis (Ofsted reports) and interviews with 2 members of current staff, 4 parents and past/current students.

Findings: themes: curriculum responded to the needs of each individual therefore the evaluation of the curriculum was not possible in the sense that Ofsted required it. Some parents felt that

inspection gave a much needed external critical appraisal of the school that did not happened internally (and cited improvements to facilities such as toilets from earlier inspection). Others felt that the inspection simply reinforced to Summerhill, the Summerhillian approach (ie. better than standard schooling). The view of Ofsted reports that there had been 'no change' in relation to their criteria clashed with views of parents, staff and students that is was a place of constant change. Staff felt demoralised, de-professionalised (their commitment and hard work not recognised) and drained by the frequency of inspections. Preparation for the Ofsted inspections was seen as robbing the school of its unique philosophy and integrity. Inspectors view the curriculum as 'fragmented' and narrow (due to non-attendance) while Summerhill saw the curriculum as the broadest one imaginable (and highly responsive to needs). Mixed dormitories were frowned on by Ofsted but viewed as part of an effort to promote equality and integration (including play and learning) by Summerhill. Ofsted's notice of complaint forced Summerhill to segregate but parental complaints and a legal appeal later allowed them to revert back. Staff involvement in decision making (school democracy) was also an issue AS Neils' philosophy allowed for fluid staff attendance as well, but some staff equally felt left out of decisions. Summerhill provided education for a lot of international students and purposefully cut itself off from the rest of the school system in England, hence could be accused of being stuck in their ways (ambiguities). The collegial model of Summerhill also made accountability difficult for Ofsted and may imply a lack of leadership, since decisions were meant to be made collegially (laterally).

Improvements: no evidence in changes to teaching; no changes to educational attainment measures; the school did not become embedded in UK system but the appeal process was seen to lead to improved networks with other external links; no improvements to priority goal setting and implementation (the latter was seen to be inappropriate to Summerhill's collegial outlook.

Author suggests that Ofsted follow a traditional, neo conservative model of school effectiveness that suggest that teachers following one particular model are the best teachers. A successful court appeal was seen to lead to a defence against future Ofsted inspections as well as reaffirming Neil's values.

Kelchtermans, G. (2007). Macropolitics caught up in micropolitics: the case of the policy on quality control in Flanders (Belgium). Journal of Education Policy, 22(4), 471-491.

Kelchtermans refers to a number of studies on the impact of inspection in Flanders conducted between 1995 and 2002 which are not available in English (Kelchtermans et al., 2000; Vandenberghe et al., 1997; Devis, 1998; Geerts, 2002; Daniels and Kemps, 2001; Daniels et al., 2002). Teachers wanted individual feedback. However, inspectors with an authoritarian attitude triggered reactions of resistance and rejection. More positive, constructively critical approaches made it more likely that teachers would use advice for improvement. Teachers felt that context was not fully taken into account by inspectors and, if they were not able to recognise themselves in reports, they were more likely to react superficially rather than by making any real change in classroom practice. Inspection was sometimes used tactically, with one example of teachers bringing their concern about toilets for nursery- age children to the attention of inspectors with the intention of getting the school board to take action. In a further example, a new principal used a negative inspection report to justify controversial decisions about reassignment of staff and to strengthen their personal authority. Kelchtermans notes that the audit procedure communicated a particular view of a 'good' education, so that schools reacted either with compliance or through an awareness that they needed to justify a different approach. Kelchtermans claims that individual teachers reacted to policy measures from either inside or outside the school so as to try and safeguard their beliefs about the kind of teacher they want to be, their 'professional self- understanding' (p485).

Klein, G. (2000). Improving Inspection for Equality. Improving Schools, 3(2), 38-43.

Inspectors felt more training was required and Headteachers also wanted some guidance on how to monitor race issues in schools. The article mentions that Ofsted has long had the powers to look at the access to curriculum and learning opportunities for all students. However, few inspectors comment on, for example, how setting and streaming my limit such chances for some students and reinforce stereotypes of what some students can achieve. Klein suggests that Ofsted should engaged in dialogue about the self-evaluation tools that schools use and learn from existing examples of good practice that support the idea that measures to improve equality can also improve overall quality.

Kogan, M. & Brunel University, Centre For The Evaluation Of Public Policy And Practice 1999. The Ofsted system of school inspection: an independent evaluation, Uxbridge, Brunel University Centre for the Evaluation of Public Policy and Practice.

Several aims to study, including, the extent to which Ofsted can provide valid and reliable data about the school and system; stimulate a culture of improvement in schools; be fair and supportive, cost effective, ensure public and professional accountability.

Methodology: Lit search of primary and secondary sources;

707 Head teacher questionnaires to samples of schools. These included primary, secondary, middle, nursery, special schools, different types of funding (LEA, Grant Maintained etc.), selective and non-selective.

Case studies conducted through 135 interviews in 26 schools (19 primary, 4 secondary, 2 special schools). These were mixture of heads, governors, teachers, partents, lea advisers and one secretary.

17 interviews with: teacher, subject and local authority associations; financial analysis.

Fieldwork conducted between March and October 1998.

Findings: Impact of inspection: preparation increased workload and led to deferment of other planned activities, was very bureaucratic. There was a sense of preparing a show for Ofsted and narrowing of activity to those that the inspectors were going to value and measure. LEA involvement was critical and it helped if the person was also Ofsted trained. In one case an acting head was appointed by LEA to take a school through inspection. Teachers felt the preparation period was stressful and worse than the inspection itself; trauma, stress, relief when over. In some cases this galvanized staff and had increased their sense of purpose. A lot of additional costs, much 'window dressing; gearing teaching specifically towards the inspection. The relationship established by Lead inspector and staff was very important to the process and reaction to it; it helped establish trust and also credibility of the team. There was also concern about lack of specialist knowledge of some inspectors. Over half of Head teachers questioned the reliability of Ofsted judgements about their school. Teachers who were observed wanted feedback and most found it useful to some extent at least, and those not given feedback were disappointed. After the inspection there was a period of "post Ofsted trauma' (p57).

The survey showed wide variations in the perceived usefulness of the report or action plan and also the objectiveness of the overall judgement of the school. Governors were generally happier than headteachers although the majority found the report to be objectively fair. Case study schools were not generally optimistic about the extent to which the report or action plan led to substantive improvements and changes at the school.

Impact following inspection:

The vast majority felt that changes to school activity would have happened without the inspection

Over 40% of headteachers reported implementing organizational changes as a result of the inspection. Examples of changes: School timetable; management restructure, governors, organization of classes, teaching style/curriculum (58%, the biggest). From 198 schools that reported staff changes, 80 instances of staff leaving were attributed to the inspection itself. There were also many reports of increased illness and absenteeism due to stress.

Positive impact:
Focus on development planning increasing over time
Increased teacher focus on the effectiveness of their teaching
Increased focus on Ofsted criteria (internalizing for practice)
Depends on quality of report and quality and expertise of inspectors
Reservations of national associations:
Ofsted helps failing schools to improve but there is less evidence of helping good schools. Concern was expressed that money could be best spent elsewhere.
Narrowing of teaching to exams and disruption of inspection process
Declines in performance due to stress caused by process of inspection
Decline in performance, e.g. one school SAT results decline because school could not focus on this
Head teacher surveys on improvement attributed to inspection:
Only 28.5% felt that Ofsted was leading to improvements in the school
Asked about improvements to SAT and GCSE results, the majority felt that no improvements could be attributable to the inspection (varied between 62 – 78%).
Helping the school to focus on areas for development – 44% agreed
Some improvements were attributed to Ofsted inspection in terms of: increased monitoring of teaching, target setting, pupil assessment and staff development.
Also, some improvements to pupil attendance and lowering of exclusions (p76 for figures).
43.1% Governors in the survey felt that the Ofsted system would lead to an improvement in standards.
However, only 33.6% governors felt that the inspection has led to improvement in the school (the rest were either neutral or disagreed – 33.8). Around two thirds felt that the inspection had helped them to focus on development areas.

Costs:
Average additional costs to schools is 1-3% of annual budget. Indirect costs great, e.g. 20% of additional teacher time in Ofsted preparation in 3 months prior to inspection.
When adding the costs of LEA assistance: approx. 4.5% for average sized primary school and 3.2% of average secondary school, excluding opportunity costs, more than is spent on staff development.
Some conclusions and recommendations:
Too strong an inspectorate leads to 'infantilism' and schools being too passive, rather than having a strong engaged profession.
Great role of SSE, self-appraisal, peer review and evidence–based practice.
Parents and governors are happier with Ofsted than teachers (they appreciate the information they can gain about the school and pointing out any key failings).
Shorter build up to inspections and less preparation needed
Relationships between inspectors and schools is crucial

Credibility of inspectors is important, especially with regard to classroom observations
Action plans were mostly a bureaucratic process; these were already in place at the school, or disrupted by the inspection or seen as a formality.
In terms of Ofsteds own criteria: e.g. quality of teaching, and of spiritual, cultural and moral development, Ofsted was perceived to have had an impact on only a minority of schools.
Schools' reflections:
Agreed with being held accountable, but may disagree with process; inspection leads to self-examination, sometimes increases mutual support among staff and gives an external perspective and clarity about roles; however, too punitive and too many adverse effects

The framework helps promote self-examination and can be applied across schools
It promotes an outdated idea of management and is inflexible to alternative approaches to school improvement
No support for naming and shaming as this would be counter productive
Different arrangements were need to identify underperforming teachers
More developmental and dialogue approach to inspections.
Teacher associations need to establish clear professional basis for teachers so that they would be stronger in relation to inspectorate

Lee, J. & Davies, J. D. 2000. Bereavement and Loss: the Emotional Consequences of Special Measures for a Sample of Special Schools. Improving Schools, 3, 44-50. (not in spreadsheet as based on same data set as reference below)

(based on DATA FROM: Book: Thomas, G. Davies, J.D., Lee, J., Postlethwaite, K., Tarr, J., Yee, W.C. and Lowe, P. (1998) Best Practice Amongst Special Schools On Special Measures: The Role of Action Planning in Helping Special Schools Improve, Research.

Research was DfEE funded in response to the identification of a disproportionate number of special schools judged as failing. 15 Special schools in SM were looked at.

Authors apply Kubler-Ross (1973) five stages of bereavement to comments of interviewees about the inspection process. Shock: to have worked so hard trying to help students with special needs and then to have 'failed' was 'devastating'. Guilt: This followed, particularly by leadership who in some cases considered resignation. Anger (and resentment): This was by teachers of leadership and also at inspectors whose interaction during the process suggested they had not understood the school well. The extra work caused by the inspection added to the resentment. Depression: many talked of a 'dip' in morale after inspection and suggested that even a good judgement may lead to this. Resolution: overcoming 'denial' was an issue in the schools that got out of SMs, followed by resolve and galvanising around a drive to work even harder as a 'professional'.

In addition, one head mentioned the authority that the SM by Ofsted had given him new clarity of purpose.

In terms of costs; professional (loss of values and cherished practices) and personal, insecurity. Also one head talked of the increase in absenteeism due to minor illnesses.

Successful schools (who get out of SM) need to recognize and explain their loss; secondly, accept emotionally what had happened and thirdly achieve a new identity.

Lee, J. & Fitz, J. 1997. Hmi And Ofsted: Evolution or Revolution in School Inspection. British Journal of Educational Studies, 45, 39-52.

Compares the Ofsted system to the old 'HMI' . The paper draws on interview transcripts with HMI staff and recently inspected school heads; documentary evidence as and national surveys and annual reports from HMI and Ofsted. The principle arguments are that Ofsted now has a vast wealth of information and data about schools. This information adds to its authority and power and creates the notion of 'steering' the system via this knowledge of what constitutes 'good practice' and good pedagogy.

Lee-Corbin, H. 2005. Under the microscope: A study of schools in special measures and a comparison with general characteristics of primary school improvement. Education 3-13, 33, 51-57.

Author looks at five primary schools in England that have successfully come out of SM and look at the processes that lead to improvement, using semi-structured interviews conducted with head-teachers, deputy headteachers and teachers. They compare these with general ones that come out of the literature and also their own previous research from the Sustained Primary School Improvement study (SPSI) (Southworth and Lee Corbin, 1999). Overall, they find considerably overlap, i.e. The factors that lead to improvement: Leadership: Educational leadership factors which centre on:

the head's knowledge of teaching and learning in the school, the head teaching for some time each week, the development of shared leadership within the school. - Staff relationship factors include teacher collaboration and opportunities for focused professional talk. - Teaching and learning factors include a concerted effort to improve the following: planning procedures, monitoring and evaluation of teaching and learning by senior staff. the development and implementation of teaching and learning policies, teacher expectations, using pupil learning data to assess progress and inform target setting. Within this, LEA support was considered particularly important alongside the continuing involvement of an HMI with each school (whose judgements were considered to be more accurate than Ofsted Inspectors. Author found that leadership was a particularly strong factor – inexperienced leaders, high turnover, inheriting many difficulties. The damage to morale of SM was seen as much greater than anticipated and also the increase in illnesses thought to be brought on by stress. Teaching and learning improved, through sharper objectives in lesson planning, monitoring by heads who usually taught a little too and supported by colleagues, the culture of professional learning was also considered to be important. Important for Heat to know staff well. Leadership and staff collaboration/support were seen as enablers for changing key aspects of teaching and learning that improved the schools.

Lindgren, J. (2014). The Front and Back Stages of Swedish School Inspection: Opening the Black Box of Judgment. Scandinavian Journal of Educational Research, 1-19. doi: 10.1080/00313831.2013.838803

highlights the official face (front stage) of inspectors, which uses a juridical and official presentation to appear 'professional and reliable' and contrasts with the more nuanced 'back stage' where judgements are more of a struggle, intuitive etc.. This is where the school development interests for staff and inspectors most come together. Inspection judgements are constructed in the light of policy context; deficiencies are 'made' for example. interviews with 8 inspectors, observations of meetings and classes and document analysis

Lindgren, J., Hult, A., Segerholm, C., & Rönnberg, L. (2012). Mediating school inspection – Key dimensions and keywords in agency text production 2003–2010. Education Inquiry, 3(4), 22.

analysis of 15 key documents, some from 2003-7 and others 2008-10. Looks at shift in the style and language of inspection documents. Improvements are part of a new system of control in which inspections need to be seen to be more impartial and technocratic in a deregulated school system. Post 2008 inspection language is less supportive and less recognition given to the local dimension.

Lowe, G. (1998). Inspection and change in the classroom: rhetoric and reality School Improvement after Inspection (pp. 97-109). London: Chapman.

7 case study schools. Interviews with 60 teachers over three years (first year interim data). Only 1 in 7 case study schools had 'substantially implemented inspection recommendations related to teaching and learning' (p.45); 3 had made some changes, and 3 none or very little change.

Luginbuhl, R., Webbink, D., & de Wolf, I. (2009). Do Inspections Improve Primary School Performance? Educational Evaluation and Policy Analysis, 31(3), 221-237.

720000 pupils in 6230 schools. Comparison of the effects of one type of inspection with another, more intensive. More intensive inspections are responsible for larger increases in the Cito test scores than the less intensive ones. In the first two years following an inspection test scores increase by 2% to 3% of a standard deviation

Lupton, R., & Hempel-Jorgensen, A. (2012). The importance of teaching: pedagogical constraints and possibilities in working-class schools. Journal of Education Policy, 27(5), 601-620.

Lupton and Hempel- Jorgenssen considered the impact on pedagogy of externally- imposed programmes for improvement and local authority pressure, including on two schools which had had poor Ofsted reports. Their article is based on case study data relating to four schools, from a research project on the effect of school composition on school processes and pupil attainment. The consequences for these two schools of not improving were those of loss of local authority confidence, increased scrutiny by both the local authority and the inspectorate, and the potential loss of the headteacher's job. Lupton and Hempel- Jorgenssen argue that it is not simply the external pressures of test league tables and Ofsted that encourage the narrow forms of pedagogy and 'teaching to the test' seen in these two schools, but the combination of this with their disadvantaged circumstances and the consequent ways in which both teacher and pupil identity are constructed and institutionalised. They present evidence from interviews with children and teachers to illustrate their claim that the 'ideal' pupil in such schools is seen as passive, as either 'able' or otherwise, and that the preferred model of teaching is didactic and in conformity with externally-prescribed models. In contrast, teachers and pupils in the other two case study schools, which serve more advantaged communities and which have not had recent negative inspection reports, are more autonomous and report fewer feelings of pressure.

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.

A range of data from different sources looking at improvement in these categories of schools. Schools identified as least effective in England (in special measures) are more likely to sustain the improvement they make after inspection than those that are relatively more effective, although still causing concern (identified as having serious weaknesses). Weaknesses identified by inspections need to be explicit and leadership focused on SI.

Matthews, P. & Smith, G. 1995. >Ofsted: inspecting schools and improvement through inspection. Cambridge Journal of Education, 25, 23-34.

Written by Peter Matthews, Head of Quality Assurance and Development, OFSTED (George Smith is research consultant). Looks at aspects of quality review and also improving inspection, and importantly, improvement through inspection. Regarding the latter: authors mention 'growing evidence' that Ofsted was contributing towards improvement. Surveys of secondary schools revealed:

"the value of having an external audit of achievements, strengths and weaknesses, providing information for parents and accountability for the expenditure of public money;
the growth in confidence and morale resulting from affirmation of a school's quality and direction;
the major impetus provided to focus thinking on aspects of the school which did not meet the Framework criteria and its power to act as a catalyst to accelerate policy review and staff development;
the identification of areas for improvement, although some inspection
reports still need to make these more clear. (p.30)"
They also assert that preparation for inspection has been shown to lead to improvement. The fact that the majority of schools had met their targets vis a vis action plans (6 month period) drawn up from the inspections is seen as evidence that inspection itself has led to improvement. The article also point to the role of Ofsted in improving the national system, not just at school level and mentions its work in looking at difficulties of teaching in particular social contexts and producing data on this to focus the government's mind on this issue (indirect improvement).

McCrone, T., Coghlan, M., Wade, P., & Rudd, P. (2009). Evaluation of the impact of Section 5 inspections - strand 3. Final report for Ofsted. Evaluation of the impact of Section 5 inspections - strand 3. Final report for Ofsted, 58.

McCrone et al., in an independent evaluation report, visited 18 schools that had recently been inspected in England. They found that inspections were seen to have a direct, positive impact on school improvement, particularly for improving assessment practices and the quality of teaching and attainment. Inspection encouraged the sharing of leadership responsibilities, with schools believing that the quality of their own self-evaluation had significantly improved since September 2005. Where teachers understood the rationale behind lesson observations, they were seen as fair, appropriate and important for professional development. Feedback directly to the teacher from an inspector was crucial to improving the quality of teaching. If the quality of feedback on lesson observations was good, the schools were more positive about the inspection. Specific and clear recommendations were most helpful, to refocus leadership and to have an impact after the inspection. Schools sometimes felt that there was insufficient observation of teaching or that inspection judgements were purely data based. Schools in particularly challenging circumstances also sometimes felt that inspectors ought to be able to take more account of these.

Memduhoglu, H. B. (2012). The Issue of Education Supervision in Turkey in the Views of Teachers, Administrators, Supervisors and Lecturers. Educational Sciences: Theory and Practice, 12(1), 149-156.

Semi structured interviews with 30 participants, academic, teaching and supervisors (inspectors).

• External 'auditing' was seen as a way of avoiding 'entropy'.
• Apart from supervisors, most participants were negative about the aim of supervision, stating that it tended towards error detection and evaluation rather than being a developmental process.
• a strength was seen as the experience of supervisors, most had at least 7/8 years of teaching experience and it was felt that such supervisors were more able to offer constructive advice.

• A weakness was the supervisors were viewed as failing to spot issues in teaching practice and teachers were unwilling to share these aspects with them.
• The official role of the inspectors was seen as a hindrance to trust and in terms of guiding and developing teachers. Supervisors were seen as lacking in training and not up to date.

Morrison, K. (2009). School inspection in small states and territories: An overview and case study of Macau. Compare: A Journal of Comparative and International Education, 39(6), 751-767.

This paper reviews school inspection systems in small states and territories (SSTs), indicating tensions, challenges and sensitivities that relate to their small size. A qualitative case study is presented of the new school inspection system in Macau, a post-colonial, small territory facing issues experienced in other SSTs. Macau has 13 government schools and 73 private schools (in 2006/7). Macau's inspection system is contextualized within its decentralized, developing schooling system and local sensitivities, potentiated in the Chinese culture where relationships, harmony and 'face' feature highly. As a result, school inspections became a high stakes exercise. Implications are drawn for policy and practice, including the need for bespoke inspection systems in SSTs, the development of resident expertise, attention to economic, personal and interpersonal costs of inspection, the need for transparency, the time cycle of inspections, the need for sanctions following inspection, and the need for experience of other schooling and inspection systems in order to make informed judgements about quality. Though large-scale inspection systems often separate inspection from development, support and advice, in SSTs that separation may be blurred. It is argued that the content and methodology of comparative study of school inspection in SSTs has to take account of their unique economic, cultural and professional features.

Nees, P. (2006). SCHOOLS AND THEIR E.R.O. RECOMMENDATIONS: a study of six Wellington area schools – May / June 2006 Sabbatical Report.

Nees's practitioner study of six primary schools in Wellington, New Zealand reported concerns about the sustainability of improvements made in response to external review, including recruitment challenges and barriers to learning among their students.

Nusche, D., Halász, G. G., Looney, J., Santiago, P., & Shewbridge, C. (2011). OECD Reviews of Evaluation and Assessment in Education: Sweden 2011.

As part of the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, the research team led by Nusche et al. (2011) in Sweden visited schools and interviewed representatives of stakeholder groups. They concluded that Swedish schools get comprehensive, high quality, feedback on performance via SSE, student and parent surveys, municipal evaluation, publication and ranking of student attainment data and external inspection, stating that, "the quality of feedback given to them (schools) about their performance, as well as their capacity to improve their own work using this feedback, have become a key success factor in the Swedish system." (p78) Inspection reports were considered to be detailed and specific with actions identified for improvements needed and schools were required to submit a plan to say how they would make improvements within 3 months. The structure of inspection reports allowed for progress to be seen over time and they considered that the SSE was well developed to enable improvement.

Ofsted (2015) Investigation: School leaders' views on the impact of inspection.

This report analyses responses to the post- inspection survey provided to inspected schools in England. The survey asks school leaders, staff and governors for their views on a range of issues,

including the quality and conduct of the inspection, its likely impact on their school and what changes they intend to make. It has an annual response rate of approximately 75%. 22 800 responses from between 2009 and 2014 were analysed for this section of the report.

In 2013-14, Ofsted also surveyed school leaders four months after the inspection had taken place. This impact survey asked leaders about the changes they were making, or had already made, as a result of their school's most recent inspection. The report is based on 829 responses to the online survey.

Key findings are reported as follows:

'The post-inspection survey 2009–14

Almost all respondents said that they would use the inspection recommendations to improve their school (98%).

Around nine out of 10 respondents (92%) reported that the demands of being inspected were reasonable and that the judgements were fair and accurate.

Most respondents agreed that the benefits of inspection outweigh the pressures of being inspected (82%).

The impact survey 2013/14

Nearly nine out of 10 school leaders (88%) reported that they had made changes to their school as a result of inspection.

Most leaders (81%) said that inspection helped them to improve by providing an accurate analysis of their strengths and weaknesses.

A large majority of leaders (79%) had found inspection helpful in confirming that they were taking the right actions.

Around seven out of 10 school leaders (73%) agreed that the inspection report would help their school to improve.

Over half of school leaders (56%) identified that ongoing professional dialogue with inspectors was the most useful aspect of the inspection process. ' (pp1-2)

Examples of changes implemented are provided in the report 'improvements to feedback and marking strategies, specific mathematics interventions and enhanced programmes of professional development.' (p 15) (in outstanding schools). Where schools needed to improve most changes were identified in areas of management and teaching and learning, for example, in monitoring and evaluation, use of data and tracking, improving professional development and in behaviour management. Very detailed responses, involving significant changes in leadership and governance were provided by six schools which had been judged to have been iandequate.

Osler, A., & Morrison, M. (2002). Can Race Equality be Inspected? Challenges for Policy and Practice Raised by the OFSTED School Inspection Framework. British Educational Research Journal, 28(3), 327-338. doi: 10.1080/01411920220137421

60 Inspection reports from 1999, 30 analysed in detail. Content analysis for terms on race equality of 10,623 inspection reports. 22 interviews with inspectors, headteachers and LEA staff. Framework for inspection was sufficiently robust to look for racial equality issues but was not prioritised in practice.

Only 34% reports contained references to race equality issues. Schools not required to monitor by ethnicity; inspectors of this area often lacked sufficient expertise; interviewing of small school-selected sample of students did not uncover issues of racism sufficiently. More priority and training needed on this area and school needs to monitor with community; not enough for inspectorate alone.

Ouston, J., et al. (1997). "What Do Schools Do after OFSTED School Inspections--Or Before?" School Leadership & Management 17(1): 95-104.

Postal Surveys sent out to secondary school headteachers in four phases 1993, autumn term; those inspected a year later, autumn 1994; third study was two years after the 1993 inspected schools and the fourth, was the equivalent for the 1994 inspected schools (i.e. also two years later). Findings:

Nearly a quarter employed a consultant to guide the school on the state of the school prior to the inspection. 48% in 1993 inspections rated this highly valuable compared to 38% in 1994. Many had used the inspection framework itself for this and found it useful.

Two thirds found the report to be fair. Most were encouraged by the report, 21% were dispirited and 10% neutral.

Speed of development: stopped 4%; 24% slowed; 34% speeded up and 38% unaffected. Some said that development slowed due to the need to prepare for inspection while others said that it made them work on things more quickly.

17% said SDP was different to the report action plan I 1993, only 5% said it was different in 1994 – perhaps learning to set own priorities based on the framework?

Follow up:

Most schools were slightly more positive about the report two years later compared to 1 year after

Schools in the 1994 group were more likely than 1993 group 30% compared to 25%, to say the effects of the inspection had been 'mixed'.

Other positive outcomes:

Confirmation that the school was good

Additional audit info

Sharpening of SDP

Negative:

Lack of confidence in judgements

Stress and demoralisation of staff

Negative impact on the community

Progress on action plans was mixed, depending on the type of action recommended. Only a small number said no progress was planned (some said it would take longer than 2 years).

Ozga, J. (2013). Accountability as a policy technology: accounting for education performance in Europe. International Review of Administrative Sciences, 79(2), 292-309.

Ozga in an interview with an Ofsted inspector, collected as part of a research project investigating the relationship between inspection and the governance of education in England, Scotland and Sweden, reports his claim of a large number of hits on the Parentview website, particularly by estate agents who use Ofsted judgements in their literature, as evidence of the belief that this is widely used by parents for choosing schools.

Penninckx, M. (2015) Inspecting School Inspections. Doctoral dissertation. University of Antwerp

"The aim of this dissertation is to investigate the (side) effects of inspections on schools and their staff, as well as to examine the features that explain the differences between schools or school staff in terms of these (side) effects. The study sets out on the following four research questions:

RQ 1. What are the expectations of different stakeholders with regard to the effects of school inspections?

RQ 2. Which indicators enable the evaluation of the effects that inspections have on schools and on their staff members? How can these indicators be made operational?

RQ 3. What are the effects of inspections on schools and staff members, as perceived by staff members in those schools? What side effects can be discerned?

RQ 4. How can differences in the perception of effects and side effects of inspections on schools and staff members be explained?" (p.3)

Methodology:

Mixed methods design. Qualitative and quantitative data were collected concurrently and the smaller qualitative studies were taken from a sub set of the larger quantitative group. Sample: Every primary and secondary school in Flanders inspected between September 2012 and February 2013 was asked to participate in the study. In those schools participating, the questionnaire was administered with principals, teachers and members of the management team. In total, 2,718 respondents from 148 schools (out of 239 schools inspected during the predefined period, equalling 61.9%) filled in the questionnaire. In 11 schools the response rate was considered too low, and was therefore excluded. Therefore, 106 Primary schools and 31 secondary schools were included and 2,668 staff members. Surveys were administered approximately 8 weeks after the inspection, to allow for the report to be received by the school. In the qualitative sample, five case study schools were selected at random. Interviews were conducted with the principal, with the pupil care coordinator, and with the teacher of grade 7. Interviews were undertaken two weeks prior to the inspection, two weeks after the inspection and finally four months after the inspection.

Summary (from p183):

Effects of inspections

A minor, increased reflection by teachers on the accordance between their teaching and the attainment targets and development goals, during the notification period prior to the actual inspection. An instrumental effect at classroom level (changes to teaching practices) is more common than an instrumental effect at school level (changes to school policies). Symbolic and strategic effects are rather rare, although some examples have been documented. On average, inspections have a moderate positive impact on staff members' feelings of personal efficacy, and a stronger positive impact on collective efficacy. There are, however, large differences between respondents with regard to their perception of these latter effects.

Side effects of inspections

The most pronounced side effects of inspections are the increased levels of stress, anxiety and tiredness, and the decrease in professional enthusiasm amongst staff members prior to, and during, the inspection. Although the stress levels generally drop to regular levels immediately after the inspection, the weeks after the inspection are still characterized by feelings of anxiety and tiredness, and predominantly a decreased level of professional enthusiasm. Furthermore some staff members report a severe impact on their personal lives, while hardly any impact is experienced by others. The inspection has no significant impact on the number of conflicts in the school team. We found only limited evidence for misleading behaviour by schools before and during the inspection. In these cases, it concerns mostly the updating or creating of new documents, or small adaptations to the physical environment (e.g. putting up new educational materials in the classrooms). On average, apart from the meticulous preparation, the lessons to be observed provide the inspectors with a good idea of the conduct of regular lessons: misleading behaviour by individual teachers is rather rare. Also disturbing effects on normal school life are rather uncommon, although there is a large diversity between the responses of different staff members in this regard.

Explanatory features

The inspection judgement has a great impact on the occurrence of effects and side effects. Fewer effects are reported in schools which have received a positive judgement, but also fewer side effects are perceived in these schools compared to schools which have received a 'restricted positive' judgement.

The perception of school staff about the inspection quality (inspector's behaviour, psychometric quality and transparency) is even more important in explaining effects and side effects. When the inspection is perceived as a high quality process, the inspection has a larger developmental effect on the school, and the decrease in professional enthusiasm after the inspection is less apparent.

The school's policy-making capacities mainly affect the extent to which stress and anxiety are experienced during the inspection: teachers in schools with strong innovative capacities, shared leadership and good professional relationships, report a higher increase in levels of stress and anxiety. This latter finding needs to be the subject of further research.

Penninckx, M., Vanhoof, J., De Maeyer, S. & Van Petegem, P. 2014. Exploring and explaining the effects of being inspected. Educational Studies, 40, 456-472.

Looks at the effects of inspection on five primary schools in the Flanders region of Belgium, just before, during and 4 months after the inspection. These were randomly selected from 64 schools that were being inspected during this period. Interviews of principal, pupil care coordinator and grade 7 teacher in each school, thus 45 interviews (5 schools, 3 respondents, three times). Analysis of conceptual (its influence on the understanding or reflection of principals and teachers in schools)

and instrumental effects (decisions taken as a result of the inspection and the actions that are based upon these decisions). 4 schools were given positive and 1 school 'restrictive positive' (building In need of repair). Few instrumental or conceptual effects were found overall. Conceptual effects were reflection on lessons and general quality of education before the inspection and some members of staff became more aware of the value of their profession or of policy matters (in own school). Only in the school with restrictive positive, were changes made, particularly to the school building (infrastructure). Authors conclude that several matters made this (lack of effects) more likely: some perceived errors by inspectors; lack of specific advice on how to tackle issues (where teaching tips were given one teacher reported trying these out). Also, aspects peculiar to the Flanders system: no advice is given to the school (just strengths and weaknesses); few schools get negative judgements (1.6%) and only 29.1% get restricted positive, thus the majority need make no response to the inspection judgement; also inspections and not 'full' inspections and focus on only key aspects 'relevant to the school' (thus some teachers and principals felt that it was not enough info' on which to base judgements). Other actors called 'school counsellors' are allowed to give advice to schools – so perhaps not entirely surprising. Also, after the inspection, four month period may not be enough, since this is sometimes seen as a time to reflect (and rest) from the stressful inspection process, so it is possible that some aspects are worked on later.

Perryman, J. 2002. Surviving Special Measures: A Case Study of a 'Fresh Start' School. Improving Schools, 5, 46-59.

This is a detailed case study of a school going through 'Special Measures' from 1999-2002 in an Inner-city setting. Documentary materials (particularly HMI), interviews with 13 teachers and surveys of students in year 10, carried out in 2002. Placed in SM in June 2000, new Head appointed in Sept 2000, six inspections until Nov 2001. March 2002, 7th inspection, shortly after, came out of SM. Fresh start scheme allowed school to restart in July 1999 with new management and new name. Many staff encouraged to retire or relocate. Only five staff remained! The turbulence and instability after Fresh Start plunged the school into SM (and this was not unusual) – high staff turnover, pupil enrolment problems, funding and lack of (inexperienced) leadership. First inspection after FS reported some of these problems that were caused by FS! The SM label was punitive and stigmatising and most staff felt the improvements could have been made without the inspections and SM, more support needed from LEA; however it did help give a 'mandate' to new school head to make difficult decisions. Losing the tag helped student enrolment which contributed to further improvements.

Perryman, J. (2005). School Leadership and Management after Special Measures: Discipline without the Gaze? School Leadership and Management, 25(3), 281-297.

1 secondary school. Analysis of semi-structured interviews with teachers just over a year after the school came out of Special Measures (research period 1) and the following year in the weeks leading to the school's next Ofsted inspection (research period 2). 12 teachers in research period 1 and 17 in research period 2.

The school management behaved as if the inspectors were still there by establishing other disciplinary mechanisms, and there is clear evidence that this was successful in the first nine months following Special Measures. However, there was a definite weakening of the panoptic pressure, and it took the return of inspectors to restore the disciplinary regime

Perryman, J. (2006). Panoptic Performativity and School Inspection Regimes: Disciplinary Mechanisms and Life under Special Measures. Journal of Education Policy, 21(2), 147-161.

1 secondary school. Interviews of teachers in Year A whilst the school was still under special measures, then in the following year (Year B). 13 in year A: two were senior management, two were heads of faculty, two were heads of year, four were heads of department, and three were main scale teachers. This reduced to 10 in year b.

Discusses the constant observation and monitoring of staff, to keep up the inspection performativity at the school

Perryman, J. (2007). Inspection and Emotion. Cambridge Journal of Education, 37(2), 173-190.

1 secondary school. The first research for this paper was carried out just over a year after Northgate came out of special measures (summer, Year A), a term later (autumn, Year B), in the subsequent spring in the weeks leading to the school's next Ofsted inspection (Year B) and in the summer following this Ofsted inspection. 12 teachers (including senior and middle managers) in Research period A, then 17 in Research period B and the following summer. Interviewed to see the impact on these individuals of the SM designation. Teachers experience a loss of power and control, and the sense of being permanently under a disciplinary regime can lead to fear, anger and disaffection.

Perryman, J. (2009). Inspection and the Fabrication of Professional and Performative Processes. Journal of Education Policy, 24(5), 611-631.

(not in spreadsheet as same data set as study in 2010 by Perryman)

1 secondary school. Five separate periods of interviews were covered with between 10 and 17 in number of teachers, middle managers and senior managers to look at period before, during and after an Ofsted inspection, designation of Special Measures and the aftermath and school follow up. Teachers and managers 'perform' during inspection.

Perryman, J. (2010). Improvement after Inspection. Improving Schools, 13(2), 182-196.

1 secondary school. Five separate periods of interviews were covered with between 10 and 17 in number of teachers, middle managers and senior managers to look at period before, during and after an Ofsted inspection, designation of Special Measures and the aftermath and school follow up. Although in many respects the school was maintaining its improvement, some middle and senior managers were suspicious about the long-term effects of becoming an institution so seemingly built around passing inspection.

Ritchie, R. 2002. School improvement in the context of a primary school in special measures. Teacher Development, 6, 329-346.

Case study of one primary school (large village in W. England) following a SM judgement in November 1999. In April 2000 three teachers and Deputy and Head teacher were interviewed (semi structured). In Feb 2001, just prior to follow up inspection, three teachers (one from last sample) plus deputy and head again. The author also kept a detailed research journal of meetings and notes of discussions. School came out of SM in March 2001. The author was a governor and also appointed as a consultant by the LEA to support the school to come out of SM. In Jan 2000, a senior management team was formed for the first time, including an LEA adviser, the author, new deputy, charged with putting action plan in place by March 2000. Head teacher resigned in April 2000. Systematic professional development cycles were introduced (collaborative discussion, observation (using Ofsted framework) and then feedback and action planning, aided by mentor). Teacher

recorded progress over time and then later there was a follow up lesson ob. Another strategy was 'better use of assessment data and to set targets for individuals and groups of pupils' (p. 334). The author and deputy were involved in evaluating the professional development cycles. Elements involved in the school's improvement: more professional dialogue, involvement of senior leadership team in professional development; building an ethos of continuous improvement; exploiting opportunities for peer support; making the culture more collaborative; training the new Deputy Head in how to facilitate the professional development cycles; and introducing the principles behind the professional development cycles to all staff; systematic and structured nature of professional development cycles. An important aspect was that there was widespread agreement among staff about the SM decision by Ofsted. Author suggests that the self-evaluation which included Ofsted criteria for lesson observations, helped make the school 'self-inspecting' and they would have a wealth of evidence to show Ofsted when they next came to visit.

Rönnberg, L., Lindgren, J. & Segerholm, C. 2012. In the public eye: Swedish school inspection and local newspapers: exploring the audit–media relationship. Journal of Education Policy, 28, 178-197.

This article looks at the interaction of inspections, their reports, the organisational public face and the role of the media in shaping this. Data was collected from a variety of sources, including 12 newspapers and 440 journal, covering inspections mainly in four diverse municipalities. All hits from inspection from 2003 to 2010 were included, approx. 80 articles in total. There were also interviews with high ranking inspectorate officials. Research questions about dog imagery articles were also analysed in relation to the place, function, and main message of the studied article (based on certain aspects from Ekecrantz and Olsson 1991).

The main point is that the inspectorate is able to use the media in order to steer opinion (policy and the public). Many see inspectorate as a kind of 'watchdog' but the media tend to pick up on 'failing' schools. Thus the inspectorate is able to direct attention to such cases and use this to steer local and national policy in particular directions. They also raise the issue of inspection by 'spin' as inspectorates develop an increasingly professionalised media dept. by providing information to the 'consumer' (the public) about schools, they have the potential not only to shape the perception of certain schools (and hence the behaviour of consumers, exercising choice - having knock on effects to the school enrolment, results, recruitment, morale etc.) but also promoting a particular ideology and view about education itself. The article contrasts the media scrutiny of the inspectorate itself with the 'chummy' relationship the media has as 'co-watchdoggers' in the system.

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.

A small but well-determined adverse, negative effect associated with the Ofsted inspection event for the year of the inspection. Regression analysis comparing schools that had been Ofsted visited (section 10) to those that had not. Dependent variable is exam results of 15 year olds.

Scanlon, M. (1999). The impact of OFSTED inspections, London, National Foundation for Educational Research for the National Union of Teachers.

Looked at the impact of inspection and special measures on school life. Based on interviews with Heads and teachers of schools currently on previous in SM. These were compared to schools not in SM. Most were shocked and over 40% teachers and 25% Heads thought inspection was flawed or unfair. 3/4s of non-SM heads and teachers felt it was a fair reflection of the quality of the school compared to less than a half of Heads and a third of teachers at SM schools. Heads more positive

about inspection than teachers overall. Heads at both SM and Non-SM reported more rigorous monitoring of staff since the inspection (especially the former). SM school staff reported significantly greater concerns over workload. Heads at SM reported 63 hrs per week (58 non SM) and teachers 56 at SM (53). Also increase in time off for illness after inspection, esp. for SM schools. LEA support was appreciated and this improved relationship with school but many complained it should have been put in place before to help avoid going into SM in the first place. Staff uniting in adversity was often mentioned as a key source of support. Quality of education improved in SM schools and standards achieved by pupils, but to a lesser extent. Staff morale reduced in SM schools post inspection. Staff retention was worse in SM schools and Ofsted created as many problems as it resolved overall, it was felt. Former SM schools: great relief when SM was lifted, less stress and improved job security. Other areas, e.g. monitoring of staff remained unchanged. Where high quality staff had been recruited, this was seen as key to maintaining improvements as well as positive outlook by staff.

Sebba, J., Clarke, J. & Emery, B. 1996. How can the inspection process enhance improvement in special schools? European Journal of Special Needs Education, 11, 82-94.

The study: 47 special schools (learning and physical difficulties). 2/3 of schools were all through (2-19). 47 school reports were analysed and 22 Headteachers responded to requests to provide comments. Evidence base for the inspection reports was also looked at, particularly to see the extent of pupil interviews.

Early report into the effectiveness of a (at time) relatively new framework to use with special schools. Mostly about headteachers and others' perceptions of the process. With regard to improvement, 3 Heads felt that the inspection was demoralizing and unhelpful, one felt the school had been misjudged. Others felt that it was helpful but would need time 'for the dust to settle' before addressing the issues (problem of time frame for impact). One said the profile of the school had been improved with parents and other school colleagues, which would have a knock-on positive effect. Another felt that the school team had united 'against the enemy' and this had helped collaborative working towards improvement. It was also felt that this model did not have adequate follow up support for improvement to take place.

Shaw, I., Newton, P. D., Aitkin, M., & Darnell, R. (2003). Do OFSTED Inspections of Secondary Schools Make a Difference to GCSE Results? British Educational Research Journal, 29(1), 63.

For kinds of schools where achievement was already much higher or lower than the average (e.g. selective schools), inspection was associated with slight improvements in achievement. For county, local education authority maintained, comprehensive schools (the largest single group), inspection did not improve examination achievement. Reports of 3000 secondary schools, analysis of GCSE results and statistical modelling to take account of a range of school characteristics

Spink, C. How can school leaders increase the highest levels of student attainment? A case study of a schools actions to increase the proportion of pupils gaining A/A* grades at GCSE following an OFSTED inspection. Dissertation MA in Leadership (Teach First) University of London (Institute of Education) 2012.

Case study of one mixed comprehensive school with 1200 students in London receiving an overall grade of outstanding for Ofsted inspection in Feb 2011 with 'good' for teaching and learning. The latter was due partly to the relatively low number of High Achieving (HA) students (A/A* GCSE and A/B for A level). The 18 month period subsequent to the inspection was looked at.

Subsequent to the inspection, in Sept 2011 the school appointed a HA coordinator on the leadership scale (middle), the author's position. The new role was directed to broaden the scope of Gifted and Talented provision in the school. The HA aim was also incorporated into the School Development Plan. A half-termly meeting was held of the HA team with all but 2 curriculum areas represented by teachers, however this dropped to only 5 teachers in the latter part of the spring term. These were to share good practice and to discuss strategies to improve HA.

The study involves interviews with principal, vp and 3 teachers at the school with at least 3 years' experience from a convenience sample. SLT and teachers' perceptions were compared and the impact of methods to increase HA were evaluated.

SLT regarded the HA aim to come under three themes: the need to increase opportunities of students to access HE, the desire to make the school look more favourably in national and borough league table (especially as these provided "one of the easiest ways for stakeholders to make comparisons about schools" (p.25).

Thirdly, the importance of a rounded education. Teachers broadly agree, only one mentioned Ofsted as the impetus.

The recommendation by Ofsted about increasing HA was seen to have been arrived at as part of a two way process (as much by the school) and probably more to do with attainment data than information provided by Ofsted, although may have helped to focus attention.

Inconclusive whether teaching had changed as a result of the school's efforts, one teacher remarking that progress in the lead up to inspection in terms of SI was more obvious than subsequently. Teachers cited greater awareness of focus on HA aim but only one was able to attribute changes in T and L to the post-Ofsted impact.

SLT felt that the HA efforts had led to changes in LO, Lesson planning, improved HW setting and quality produced, and increase in library books lent since the inspection.

Impact was judged to be modest overall and recommendations included:

Improving CPD, including auditing good practice and increasing dissemination

Thomas, G. Davies, J.D., Lee, J., Postlethwaite, K., Tarr, J., Yee, W.C. and Lowe, P. (1998) Best Practice Amongst Special Schools On Special Measures: The Role of Action Planning in Helping Special Schools Improve. UWE, Research Report to the DfEE

Study looked at special schools that had been successful in getting out of SMs (9 at the time) and compared to others that had not, 61 in total had been in SM. Patterns in the quality of action plans produced subsequent to SM by Ofsted were examined to see if there were differences among the successful schools compared to unsuccessful. 14 schools were visited to reflect a geographical spread and there were interviews with SLT, teachers, parents, governors, parents and pupils as well as analysis of SDPs, policies and minutes of AGMs. Action plans of all school action plans was conducted using a 31 point rating scale derived from Ofsted, DfEE and SI literature.

Findings: well-constructed plans were most likely to lead to progress, these were more precise in terms of responsibility, monitoring, review and progress tracking. Target setting and evaluation were weak for all plans.

Consideration of the timeline for implementation of targets was crucial and better plans had a timeline that was not over ambitious.

Drawing on support and guidance from LEAs on a wide variety of areas of expertise was important part of process and was lacking among many of the schools in SM.

There was more resistance from teaching staff in Special Schools that dealt with behaviour and emotional problems as the curriculum aims were perceived as clashing with therapeutic needs of pupils sometimes. These schools also needed more support.

Overall, SM was stressful but perceived to ultimately lead to beneficial changes.

Timmermans, A.C.; I. F. de Wolf, I.F.; Bosker, R.J. & Doolaard, S. (2015) Risk-based educational accountability in Dutch primary education. Education Assessment, Evaluation and Accountability, in press.

The study investigated effects of a risk-based approach to inspection on multiple cohorts of Dutch primary schools. Adverse effects were defined as below average final achievement and/or below average value added. School composition, previous underperformance, insufficient judgments on having a systematic evaluation approach, evaluation of support, and monitoring student performance appeared as factors related to subsequent underperformance of schools. Data sources and samples were:

A student-level dataset from a sample of primary schools derived from the Monitoring and Evaluation system of CITO, the Netherlands Institute for Educational Measurement. The data are collected by schools for their own use, and these data are therefore not available for the Inspectorate of Education to use in the risk analysis. Analysis was based on reading comprehension scores in grades 3 to 5 for three successive cohorts.

Cohort 2003/2004: 15,195 students and 262 schools

Cohort 2005/2006: 17,886 students and 314 schools

Cohort 2007/2008: 22,815 students and 371 schools

Based on the total inspection framework; 3.7 % of the schools in the sample were considered as indequate and 0.4% of the schools as very poor. At the same time, in the population, 6 % of the primary schools in the Netherlands were considered as indequate and 1.5 % as very poor by the inspectorate so the data set is not fully representative. The dataset also contained schools with a relatively high proportion of students from highly educated parents. Student level data was used to estimate the performance of the schools for the successive cohorts.

School-level information up to 2009 derived from the Dutch Inspectorate of Education and currently used in the risk analysis in the Dutch educational accountability system. It consists of variables concerning the curriculum, classroom practices, additional support, monitoring progress, general quality, staff and student population characteristics, school board.

Variables concerning the schools' curriculum, classroom practices, additional support, and monitoring progress were derived from records from the most recent inspection before the risk based approach was introduced. These were from 1 to 4 years old at the time of the study. These were used as predictors in the risk analysis.

Findings

Results imply that the underperformance of schools cannot be predicted very accurately. This relates to the issue of the moderate stability of school performance indicators over subsequent cohorts.

Although differences in school performance indicators across time might indicate actual changes in the performance of a school, they might also reflect unreliability of the estimation of value added.

The number of false positives and false negatives in a risk analysis depends heavily on the decision rules that determine which schools are considered at risk. More inclusive criteria are less likely to miss underperformance but may also result in inspections of many more well performing schools.

School composition, previous underperformance, insufficient judgments on having a systematic evaluation approach, evaluation of support, and monitoring student performance appeared as factors related to subsequent underperformance of schools.

Whitby, K. (2010). School inspections: recent experiences in high performing education systems. Reading, England. CfBT Education Trust.

As well as reports of research in academic journals, Whitby (2010) includes government reports and external evaluations in a systematic review of literature on six high performing inspection systems (The Netherlands, England, Scotland, Hong Kong, Singapore, New Zealand). Her intention was to summarise the existing literature in the field by looking at why education is inspected, whether inspection systems should be self- or externally regulated (or a mixture of the two), who and what is inspected, and the stakeholders in the processes and the products of inspection. She concludes that external inspection is most likely to be effective when there is collaboration with the school, focused on improvement. Both the content and focus of the review should be agreed with the school and inspection criteria should be clearly understood.

Whitby also found in her review that the quality of feedback was important, giving as an example an external evaluation report, based on extensive empirical research with schools, on the impact of Ofsted inspections in England:

 "the recent report by McCrone et al. (2009) found a statistically significant relationship between constructive oral feedback and overall satisfaction with the inspection process (pii). McCrone et al. also found that 'specific recommendations' were most helpful as they provide focus and the appropriate actions needed were easy to identify (piii). Conversely, it was found that very broad recommendations 'did not instigate direct action' (ibid)" (Whitby, 2010, p 14).

Whitby noted that all of the countries she considered used self- evaluation (SSE) to inform school inspection though to varying degrees. In Hong Kong, for example, external inspection is used to complement SSE. Whitby is of the opinion that the Scottish system is of interest for the way in which inspection and SSE complement one another "self-evaluation and external inspection documentation uses 'the same language', this means that 'teachers are much more likely to see external inspection in a developmental perspective rather than a judgmental one' (Livingston and McCall, 2005, p175). " (Whitby, 2010, p 15).

On the other hand Whitby found literature that noted a tension between SSE and inspection and a risk that they may simply be written to comply with expectations of the inspectorates (citing MacBeath et al., 2000; Estonian Ministry of Education and Research, 2005; De Grauwe and Naidoo, 2004; Meuret and Morlaix, 2003). Whitby's overall conclusion to her review is that it is the amount of guidance and support that schools have for SSE and external inspection that affect the impact of inspection systems on school improvement.

Whylie, C. (2012). Challenges around capability improvements in a system of self-managed schools in New Zealand. San Francisco: WestEd.

Whylie in a report which describes changes in inspection and support processes in New Zealand over time, cites data from the Education Review Office that approximately 20% of schools were identified as in need of 'supplementary review' (i.e. return visit in less than 3 years) in 2010. She also cites analysis by ERO staff in 2009 which showed that 18 percent of all schools came into the supplementary review category twice since the mid-1990s, suggesting that, as in England, some schools struggle to improve and then maintain improved performance. Schools serving low-income communities and schools with very small enrolment are overrepresented in this category. While about 70 percent of the schools that were in the supplementary review category twice had returned to the normal (every three years) ERO cycle after their most recent ERO review, only about 43 percent of the 96 schools that were in the supplementary review category four times or more during the period had returned to the regular ERO cycle after their most recent review. Intervention measures to help schools improve are implemented by the Ministry of Education and include specialist adviser support and the use of school improvement clusters of schools. In addition to ERO review reports, the Ministry of Education uses other data, including financial returns, annual school reports and teacher turnover figures to identify 'at risk' schools.

Wilcox, B. & Gray, J. 1994. Reactions to Inspection: a study of three variants. Cambridge Journal of Education, 24, 245-269.

Individual interviews were conducted with inspectors and teaching staff involved with three primary school inspections in three different LEAs. The interviews took place some time after the end of each inspection when the findings were generally known amongst the teaching staff. The inspections differed according to the degree of negotiation allowed to staff, the extent of inspection coverage, and degree of conformity to a full inspection model. Teacher reactions appeared to vary according to the extent and nature of any 'surprises' in the inspection reports and their findings. Some general issues emerged which were concerned with the contextualisation of judgements and the influence of time on the credibility of inspection methods.

Interviews with teachers, heads and Ofsted inspectors. Findings: The inspectors learned a lot about the schools; the staff at the school generally had few surprises though. Perhaps provided function of information for parents. Important issues raised about the need for the judgements to be accepted and to have firm evidential basis. Inspections could be professionally very damaging, especially if not handled the right way.

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.

Willis noted that despite a lack of confidence in the Ofsted judgements, it had focused the attention of the schools placed in 'special measures' on student achievement results. 3 Primary school. Self-reflection of three Head teachers.

Wong, W.L. (2010). Implementation of School Self-Evaluation in Secondary Schools: Teachers' Perspective. (Ed.D.) The Chinese University of Hong Kong.

This doctoral thesis was a qualitative study of how teacher administrators and teachers in three Hong Kong secondary schools experienced the implementation process of School Self-Evaluation (SSE), perceived the effects of SSE and described the implementation approach of SSE. Selected

secondary schools had all experienced a complete cycle of External School Review (ESR) or Quality Assurance Inspection (QAI) and SSE.

Research questions were:

1. From the perspective of teacher administrators and teachers, how was SSE implemented in the three sample schools?

2. From the perspective of teacher administrators and teachers, what were the perceived effects and/or consequences of SSE?

3. Given these implementation experiences and perceptions, how could the implementation of SSE be accounted for from the perspectives of policy implementation within the policy studies in education?

Conclusions: the author argues that for the schools studied:

 The implementation of SSE was a complex interaction of the accountability Policy, the Place where the policy was implemented, and the People who were responsible for implementing the policy.

The context of the school, the biographical and professional background of teachers and their role shaped the way in which teachers perceived the effects of SSE on school improvement or managerial control.

The implementation of SSE was not classifiable as top-down, bottom-up or hybrid approaches, but varied with the complexity of the implementation context including the Policy to be implemented, the Place and the People who implemented the policy.

Wong, M. N., & Li, H. (2010). From External Inspection to Self-Evaluation: A Study of Quality Assurance in Hong Kong Kindergartens. Early Education and Development, 21(2), 205-233.

Wong and Li studied SSE in kindergartens in Hong Kong, which introduced a Quality Assurance Inspection (QAI) model with three stages consisting of school self-evaluation, external inspection, and release of the report to the public to early childhood education in 2000. They analysed external inspection judgements on the quality of SSE and found that few of the 80 kindergartens in the initial phase of their study scored well on this. Case studies in three of the kindergartens with, respectively, excellent, good and satisfactory rankings for SSE found that all schools reported on feedback as providing insight and being constructive, with several specific examples given of school improvement actions cited as resulting from the feedback and advice given. Negative effects reported by all three schools were those of workload and stress associated with the inspection. Comments on workload were greatest from the weakest setting (which had problems with recruiting and retaining suitable staff).

Wurster, S. & Gaertner, H. Handling School Inspection and Their Results: Differentiation of School Types European Conference on Educational Research (ECER), 2011 Berlin.

Surveys of school principals and teachers were used and Latent Class Analysis was used to identify types of schools (391 schools) and a second part of the survey (one year later) looked longitudinally at responses to inspections (185 schools and 70% response rate). This looked at a variety of indicators:

"teacher and student satisfaction; burnout of teachers; classroom management; participation of parents and teachers; leadership; professional development; co-operation among staff; school-self-evaluation; quality management". The first part of the survey identified 25% of schools as 'active' i.e.

Good inspection results, high activity level and extensive communication and reflection; 29% poor inspection results, low activity and negative perception to the outcome; 26% low activity but very good inspection result and 21% reactive schools – low activity apart from high preparation for the eventually substandard result. While the model was accepted little overall change in school quality was shown.

References for empirical studies on inspection

Allen, R. & Burgess, S. (2012). How should we treat under-performing schools? A regression discontinuity analysis of school inspections in England. Bristol, University of Bristol, Centre for Market and Public Organisation, Bristol Institute of Public Affairs.

Altrichter, H. and D. Kemethofer (2014). Impact of School Inspection on Teaching and Learning in Primary and Secondary Education in Austria: Final-technical-report-Austria,

Altrichter, H. & Kemethofer, D. (2015) Does accountability pressure through school inspections promote school improvement? School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice, 26:1, 32-56.

Araújo, M. 2009. A Fresh Start for a 'failing school'? A qualitative study. British Educational Research Journal, 35, 599-617.

Bates, A. (2013). Transcending systems thinking in education reform: implications for policy-makers and school leaders. Journal of Education Policy, 28(1), 38-54.

Baxter, J. & Hult, A. 2013. Different systems, different identities: school inspectors in England and Sweden a comparative study. European Conference for Educational Research: Educational Improvement and Quality Assurance, Network 11. Istanbul.

Baxter, J. A. 2013. An independent inspectorate? Addressing the paradoxes of educational inspection in 2013. School Leadership & Management, 34, 21-38.

Baxter, J. A. 2013. Professional inspector or inspecting professional? Teachers as inspectors in a new regulatory regime for education in England. Cambridge Journal of Education, 43, 467-485

Baxter, J., & Clarke, J. (2013). Farewell to the tick box inspector? Ofsted and the changing regime of school inspection in England. Oxford Review of Education, 39(5), 702-718.

Bitan, Kristin; Haep, Anna & Steins, Gisela (2014): School inspections still in dispute – an exploratory study of school principals' perceptions of school inspections, International Journal of Leadership in Education: Theory and Practice. Prior to publication DOI: 10.1080/13603124.2014.958199

Braun, A., Maguire, M., & Ball, S. J. (2010). Policy enactments in the UK secondary school: examining policy, practice and school positioning. Journal of Education Policy, 25(4), 547-560.

Brimblecombe, N., Ormston, M., & Shaw, M. (1995). Teachers' Perceptions of School Inspection: a stressful experience. Cambridge Journal of Education, 25(1), 53-61. doi: 10.1080/0305764950250106

Brimblecombe, N., Shaw, M., & Ormston, M. (1996). Teachers' Intention to Change Practice as a Result of Ofsted School Inspections. Educational Management Administration & Leadership, 24(4), 339-354.

Brookes, C (Gen Sec of NAHT) Ch. 8 We need an inspection process. But not this one. In *DE WAAL, A. (Ed) & INSTITUTE FOR THE STUDY OF CIVIL SOCIETY 2008. Inspecting the inspectorate: Ofsted under scrutiny, London, Civitas.

Carlbaum, S. 2014. Equivalence and performance gaps in Swedish school inspection: context and the politics of blame. Discourse: Studies in the Cultural Politics of Education, 16.

Case, P., et al. (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 Inspection. School Leadership & Management, 21(1), 59-73.

Chapman, C. 2002. Ofsted and School Improvement: Teachers' Perceptions of the Inspection Process in Schools Facing Challenging Circumstances. School Leadership & Management, 22, 257-272.

Coldron, J., Crawford, M., Jones, S. & Simkins, T. 2014. The restructuring of schooling in England: The responses of well-positioned headteachers. Educational Management Administration & Leadership, 42, 387-403.

Courtney, S. J. (2012). Ofsted's revised school inspection framework: experiences and implications. Paper presented at BERA conference, Manchester.

Courtney, S. J. & Steven, J. C. 2013. Head teachers' experiences of school inspection under Ofsted's January 2012 framework. United Kingdom.

Cross, A. 2006. Looking for patterns in Ofsted judgements about primary pupil achievement in design and technology. Education 3-13, 34, 163-172.

Day, C., & Gu, Q. (2010). The new lives of teachers. London: Routledge.

de Waal, A. 2006 Inspection, inspection, inspection: how Ofsted crushes independent schools and independent teachers. London, The Institute for the Study of Civic Society

Dean, J. 1995. What Teachers and Headteachers Think about Inspection. Cambridge Journal of Education, 25, 45-52.

Dedering, K., & Muller, S. (2011). School Improvement through Inspections? First Empirical Insights from Germany. Journal of Educational Change, 12(3), 301-322.

Dobbelaer, M. J., Prins, F. J., & van Dongen, D. (2013). The Impact of Feedback Training for Inspectors. European Journal of Training and Development, 37(1), 86-104.

Dougill, P. R., Mike; Blatchford, Roy; Fryer, Lyn; Robinson, Carol; Richmond, John. (2011). To the next level: good schools becoming outstanding. Reading: CfBT Education Trust.

Education Scotland (2012). Quality and Improvement in Scottish Education: Trends in Inspection Findings 2008-2011 (pp. 48). Livingston, Scotland.

Ehren, M. C. M. and A. J. Visscher (2008). "The Relationships Between School Inspections, School Characteristics And School Improvement." British journal of educational studies 56(2): 205-227.

M. Ehren., and N. Shackleton (2015) Risk-based school inspections: impact of targeted inspection approaches on Dutch secondary schools Unpublished manuscript

Ehren, M. C. M., & Honingh, M. E. (2011). Risk-based school inspections in the Netherlands: A critical reflection on intended effects and causal mechanisms. Studies in Educational Evaluation, 37(4), 239-248. doi: http://dx.doi.org/10.1016/j.stueduc.2012.02.001

Ehren, M. C., Honingh, M., Hooge, E., & O'Hara, J. (2015). Changing school board governance in primary education through school inspections. Educational Management Administration & Leadership, 1741143214549969.

Ehren, M. C., Shackleton, N., Tymms, P., Jones, K., Anane, E, Gustafson, J, Myearber, E. McNamara, G., O'Hara, J., Conyngham, G., Altricher, D., Kemethofer, E., Schmidinger, E., Greger, D. (2015)

Technical report EU-project. Year 1-3 data (September-November 2011, 2012, 2013) 'Impact of School Inspections on Teaching and Learning'

Ehren, M.C.M., Altrichter, H., McNamara, G. and O'Hara, J. (2013). Impact of school inspections on teaching and learning; Describing assumptions on causal mechanisms in six European countries. Educational Assessment, Evaluation and Accountability. 25(1), 3-43. http://dx.doi.org/10.1007/s11092-012-9156-4

Ehren, M.C.M., Gustafsson, J.E., Altrichter, H., Skedsmo, G., Kemethofer, D., Huber, S.G. (2015). Comparing effects and side effects of different school inspection systems across Europe. Unpublished manuscript

Ehren, Melanie C. M. & Shackleton, Nichola (2015): Mechanisms of Change in Dutch Inspected Schools: Comparing Schools in Different Inspection Treatments, British Journal of Educational Studies

Ehren, M. C., & Swanborn, M. S. (2012). Strategic Data Use of Schools in Accountability Systems. School Effectiveness and School Improvement, 23(2), 257-280.

Ehren, C. M. & Shackleton, N. 2014. Impact of school inspections on teaching and learning in primary and secondary education in the Netherlands. London: Institute of Education. (not in spreadsheet as data is detailed in Ehren et al (2013) Impact of school inspections on teaching and learning; Describing assumptions on causal mechanisms in six European countries.)

Ehren, M., Perryman, J. & Shackleton, N. 2014. Setting expectations for good education: how Dutch school inspections drive improvement. School Effectiveness and School Improvement, 1-32.

Erdem, A. R. & Yaprak, M. 2013. The Problems That the Classroom Teachers Working in Villages and County Towns Confront in Educational Inspection and Their Opinions Concerning the Effect of These Problems on Their Performance. Educational Research and Reviews, 8, 455-461.

Follows, M. 2001. After the storm the tale of an Ofsted inspection. Management in Education, 15, 24-28.

Follows, M. 2001. Learning from the tale of an Ofsted inspection. Management in Education, 15, 32-34. (not in spreadsheet as same data set as above).

Francis, B. (2011). (Un)satisfactory? Enhancing life chances by improving 'satisfactory' schools. London, Royal Society for Arts and Manufactures (RSA).

Gaertner, H., Wurster, S. & Pant, H. A. 2014. The effect of school inspections on school improvement. School Effectiveness & School Improvement, 25, 489-508.

Gray, C. & Gardner, J. 1999. The Impact of School Inspections. Oxford Review of Education, 25, 455-468.

Greger, D., Huber, S., Jones, K,L., Kemethofer, D., Myearberg, E., O'Hara, J., Skedsmo, G., and Tymms, P. (2015) The unintended effects of school inspection: The prevalence of inspection side-effects in Austria, the Czech Republic, England, Ireland, the Netherlands, Sweden and Switzerland. Unpublished manuscript

Grek, S., et al. (2013). "Governing by inspection? European inspectorates and the creation of a European education policy space." Comparative Education 49(4): 486-502.

Hardy, I. (2012). "Managing" Managerialism: The Impact of Educational Auditing on an Academic "Specialist" School. European Educational Research Journal, 11(2), 274-289.

HMIE. (2010). Learning Together: Lessons about School Improvement--An HMIE Report on How Schools Get Better. Second Edition: Her Majesty's Inspectorate of Education.

Hogenbirk, P., & Braak, P. (2013). ICT action school development at Helen Parkhurst Dalton School. Education and Information Technologies, 18(2), 233-243.

Hopkins, D., Harris, A., Watling, R. & Beresford, J. 1999. From Inspection to School Improvement? Evaluating the Accelerated Inspection Programme in Waltham Forest. British Educational Research Journal, 25, 679-690.

Hussain, I. (2012). Subjective performance in the public sector: evidence from school inspections: London School of Economics and Political Science. Centre for Economic Performance.

IpsosMori (2008). School inspections - final report 2008: Office for Standards in Education, Children's Services and Skills.

Jeffrey, B. & Woods, P. 1996. Feeling Deprofessionalised: the social construction of emotions during an OFSTED inspection. Cambridge Journal of Education, 26, 325-343.

Jones, K. & Tymms, P. 2014. Ofsted's role in promoting school improvement: the mechanisms of the school inspection system in England. Oxford Review of Education, 40, 315-330.

Keeble-Allen, D. 2004. Inspection at Summerhill: did Ofsted inspection result in improvement? U210580 Ed.D., University of Leicester (United Kingdom).

Kelchtermans, G. (2007). Macropolitics caught up in micropolitics: the case of the policy on quality control in Flanders (Belgium). Journal of Education Policy, 22(4), 471-491.

Klein, G. (2000). Improving Inspection for Equality. Improving Schools, 3(2), 38-43.

Kogan, M. & Brunel University, Centre For The Evaluation Of Public Policy And Practice 1999. The Ofsted system of school inspection: an independent evaluation, Uxbridge, Brunel University Centre for the Evaluation of Public Policy and Practice.

Lee, J. & Davies, J. D. 2000. Bereavement and Loss: the Emotional Consequences of Special Measures for a Sample of Special Schools. Improving Schools, 3, 44-50. (not in spreadsheet as based on same data set as reference below)

Lee, J. & Fitz, J. 1997. HMI and OFSTED: Evolution or Revolution in School Inspection. British Journal of Educational Studies, 45, 39-52.

Lee-Corbin, H. 2005. Under the microscope: A study of schools in special measures and a comparison with general characteristics of primary school improvement. Education 3-13, 33, 51-57.

Lindgren, J. (2014). The Front and Back Stages of Swedish School Inspection: Opening the Black Box of Judgment. Scandinavian Journal of Educational Research, 1-19. doi: 10.1080/00313831.2013.838803

Lindgren, J., Hult, A., Segerholm, C., & Rönnberg, L. (2012). Mediating school inspection – Key dimensions and keywords in agency text production 2003–2010. Education Inquiry, 3(4), 22.

Lowe, G. (1998). Inspection and change in the classroom: rhetoric and reality School Improvement after Inspection (pp. 97-109). London: Chapman.

Luginbuhl, R., Webbink, D., & de Wolf, I. (2009). Do Inspections Improve Primary School Performance? Educational Evaluation and Policy Analysis, 31(3), 221-237.

Lupton, R., & Hempel-Jorgensen, A. (2012). The importance of teaching: pedagogical constraints and possibilities in working-class schools. Journal of Education Policy, 27(5), 601-620.

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.

Matthews, P. & Smith, G. 1995. >OFSTED: inspecting schools and improvement through inspection. Cambridge Journal of Education, 25, 23-34.

McCrone, T., Coghlan, M., Wade, P., & Rudd, P. (2009). Evaluation of the impact of Section 5 inspections - strand 3. Final report for Ofsted. Evaluation of the impact of Section 5 inspections - strand 3. Final report for Ofsted, 58.

Memduhoglu, H. B. (2012). The Issue of Education Supervision in Turkey in the Views of Teachers, Administrators, Supervisors and Lecturers. Educational Sciences: Theory and Practice, 12(1), 149-156.

Morrison, K. (2009). School inspection in small states and territories: An overview and case study of Macau. Compare: A Journal of Comparative and International Education, 39(6), 751-767.
Nees, P. (2006). SCHOOLS AND THEIR E.R.O. RECOMMENDATIONS: a study of six Wellington area schools – May / June 2006 Sabbatical Report.

Nusche, D., Halász, G. G., Looney, J., Santiago, P., & Shewbridge, C. (2011). OECD Reviews of Evaluation and Assessment in Education: Sweden 2011.

Ofsted (2015) Investigation: School leaders' views on the impact of inspection.

Osler, A., & Morrison, M. (2002). Can Race Equality be Inspected? Challenges for Policy and Practice Raised by the OFSTED School Inspection Framework. British Educational Research Journal, 28(3), 327-338. doi: 10.1080/01411920220137421

Ouston, J., et al. (1997). "What Do Schools Do after OFSTED School Inspections--Or Before?" School Leadership & Management 17(1): 95-104.

Ozga, J. (2013). Accountability as a policy technology: accounting for education performance in Europe. International Review of Administrative Sciences, 79(2), 292-309.

Penninckx, M. (2015) Inspecting School Inspections. Doctoral dissertation. University of Antwerp

Penninckx, M., Vanhoof, J., De Maeyer, S. & Van Petegem, P. 2014. Exploring and explaining the effects of being inspected. Educational Studies, 40, 456-472.

Perryman, J. 2002. Surviving Special Measures: A Case Study of a 'Fresh Start' School. Improving Schools, 5, 46-59.

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.

Ritchie, R. 2002. School improvement in the context of a primary school in special measures. Teacher Development, 6, 329-346.

Rönnberg, L., Lindgren, J. & Segerholm, C. 2012. In the public eye: Swedish school inspection and local newspapers: exploring the audit–media relationship. Journal of Education Policy, 28, 178-197.

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.

Scanlon, M. (1999). The impact of OFSTED inspections, London, National Foundation for Educational Research for the National Union of Teachers.

Sebba, J., Clarke, J. & Emery, B. 1996. How can the inspection process enhance improvement in special schools? European Journal of Special Needs Education, 11, 82-94.

Shaw, I., Newton, P. D., Aitkin, M., & Darnell, R. (2003). Do OFSTED Inspections of Secondary Schools Make a Difference to GCSE Results? British Educational Research Journal, 29(1), 63.

Spink, C. How can school leaders increase the highest levels of student attainment? A case study of a schools actions to increase the proportion of pupils gaining A/A* grades at GCSE following an OFSTED inspection. Dissertation MA in Leadership (Teach First) University of London (Institute of Education) 2012.

Thomas, G. Davies, J.D., Lee, J., Postlethwaite, K., Tarr, J., Yee, W.C. and Lowe, P. (1998) Best Practice Amongst Special Schools On Special Measures: The Role of Action Planning in Helping Special Schools Improve. UWE, Research Report to the DfEE

Timmermans, A.C.; I. F. de Wolf, I.F.; Bosker, R.J. & Doolaard, S. (2015) Risk-based educational accountability in Dutch primary education. Education Assessment, Evaluation and Accountability, in press.

Whitby, K. (2010). School inspections: recent experiences in high performing education systems. Reading, England.: CfBT Education Trust.

Whylie, C. (2012). Challenges around capability improvements in a system of self-managed schools in New Zealand. San Francisco: WestEd.

Wilcox, B. & Gray, J. 1994. Reactions to Inspection: a study of three variants. Cambridge Journal of Education, 24, 245-269.

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.

Wong, W.L. (2010). Implementation of School Self-Evaluation in Secondary Schools: Teachers' Perspective. (Ed.D.) The Chinese University of Hong Kong.

Wong, M. N., & Li, H. (2010). From External Inspection to Self-Evaluation: A Study of Quality Assurance in Hong Kong Kindergartens. Early Education and Development, 21(2), 205-233.

Wurster, S. & Gaertner, H. Handling School Inspection and Their Results: Differentiation of School Types European Conference on Educational Research (ECER), 2011 Berlin.

Annotated bibliography of non-empirical research.

This annotated bibliography summarises those texts consulted for non-empirical articles, books, reports and commentaries on the impact of inspections.

Anonymous journalist (2014). Academy chains under attack. Education Journal, (193), 25

Baroness Hughes of Stretford (Lab) asked the government to act in the light of the removal of 10 academies form the E-Act Academy chain in a House of Lords oral question during a debate in March 2014. She also mentioned another big chain that had falsely claimed money for ghost pupils. She called for the Government to inspect Academy chains, not just ind. Schools. Lord Nash (con) Parlimentary Under Secretary of State for Schools, said that Ofsted could already do 'batch' inspections of academy chains and the DFE had a very tight grip on management of academy chains too, so not needed. Lord Storey (LDP) felt that this policy contradicted the same principle that allows Ofsted to inspect LA children's services and school improvement. In response Lord Nash said that the government had been in discussion with 50 chains to "strengthen their governance arrangements" (p25) and they also had Non-executive directors to support the Academy chains.

Anonymous journalist (2014). Listening to Andreas. Education Journal, (193), 3

This refers to the visit of Andreas Schleicher (the man behind PISA) to the education select ctte for a 2 hr discussion. Michael Gove had described him as "the most influential man in English education" (p3). He said that Academy chains could be part of the solution or the problem. They offered local discretion but a highly autonomous system needed a strong framework in order not to lose central oversight. The article argues that the government's claim to be overseeing 50 academy chains did not take into account the fact that many of these were very small chains. The article mentions concerns raised about the E-Act chain that recently lost 10 of its 34 schools as a result of poor Ofsted reports. they mention that this trust had difficulties at the outset, with Lord Bhatia, the 2009 founder being forced out after financial irregularities and, 2 years later the Director General, Sir Bruce Liddington also leaving for reasons to do with financial mis-management (the latter being on an unusually high salary, raising questions about whether such chains act out of financial gain or the public good. AS argued that a strong framework needed to include inspection and saw no case for separate arrangements to inspect academies (as is being argued by some right wing think tanks).

Anonymous journalist (2014). Damning Ofsted reports on E-ACT academies. (2014). Education Journal, (196), 5.

16 Inspections carried out over 2 weeks on E-ACT schools. The overwhelming number had not provided a good education to their pupils. 11 were less than 'good', 5 in special measures, 10 had not improved since last inspection, (6 had declined); 4 good and only 1 outstanding. Ofsted concluded that the trust had failed to intervene to provide support for those in Special measures. E-Act had top sliced some of the pupil premium funding since Sept 2013 but it was not clear how this money had been spent to help improve the education of disadvantaged pupils. Key weaknesses were noted across many of the 16:

Poor T and L

Weak monitoring of pupils

Poor assessment and marking

Poor QA by middle managers

Weak governance

Lessons that did not challenge the more able

Martin Freedman, Dir Economic strategy and negotiations at ATL said it was astonishing that government had failed to pick these problems up earlier and allowed the chain to expand so rapidly.

Anonymous journalist (2013) Ofqual warns Gove on accountability. Education Journal, (162), 1.

"Ofqual's Chief Regulator Glenys Stacey warned that the current system of school accountability "has too many perverse incentives, and can distort teaching, narrow the curriculum and place undue pressure on individual qualifications."" (p1). Goodhart's law (based on banker Charles Goodhart) suggests that when a measure becomes a target it is no longer a good measure. The article is about Ofqual and gthe government agreeing that the best accountability measure would used the best 8 subjects for each student and the distance travelled from entry to GCSE and separately English and Maths GCSE grades. The former would be the basis for league tables.

Anonymous journalist (2013) New approach to accountability and inspections suggested at NUT national conference. Education Journal, (172), 10

Prof Colin Richards suggested that too much time was spent on ticking off activities and use of quantitative data that was misconceived by Ofsted. He suggested two stage inspections in order to focus on quality and data led performance. This would lead to two published reports, the second of which – data on school performance, should lead to recommendations for school development. He encouraged school self-evaluation as well as abolishing twenty min observations or the requirement to see progress in individual lessons. He suggested a completely independent (of the DfES) Ofsted that was only answerable to parliament and subject to periodic review by teachers and other interested parties. Testing for reading, maths and basic writing skills should be reported to parents but not used as accountability system for schools. He suggested teachers be assessed yearly based on 'a child's progression in relation to their attitudes to learning' (p.10). Accountability would be 3 stages: the system, individual schools, and individual children. He sought a balance between professional autonomy and accountability.

Anonymous journalist (2014) New blueprint for inspection. Education Journal, (195), 1.

IN key note to Ass. School and college leader, M Wilshaw said that schools judged to be good should no longer be routinely inspected. Instead, more frequent, light touch inspections should be carried out by expert inspectors who seek to engaged in discussion with senior leaders at these schools. Ofsted's own evidence gained from reviewing 7000 annual inspections suggest that all schools need not be inspected in the same way. MW also stated the intention to hold a consultation with parents, schools and other parties on what they wanted from inspection and one question for consultation is whether early years and 6th form provision should have separate grades. Mary Bousted (gen sec of ATL) wanted inspections to give a fuller picture of schools, including aspects unrelated to test and exam results and inspection that allowed professionals to provide honest and supportive judgements. Martin Doel, AOC chief exec, said that separate grades for 6th form made sense as post 16 provision was so varied now (UTCs, FE, school 6th form etc.) but that the framework would need to be consistently applied.

Anonymous journalist (2014). Ofsted to bring inspections back in-house. Education Journal, (201), 4.

From Sept 2015, Ofsted announced it will no longer out source for the provision of inspectors. Tribal, SERCO and CfBT contracts will run out in August 2015.

Anonymous journalist (2014) School partnerships. Education Journal, (187), 20

This article shows the 4th report of session 2013-14, HC999 select ctte, 4th special report, published Jan 2014 by the stationery office ltd. the ctte argued that parents should receive more info' on the performance of academy chains and that Ofsted should have the powers to inspect them. The government responded that Ofsted could already inspect groups of Academies if they had an interest or if the Sec of State had asked the chief inspector for advice. Therefore they argued that this would not provide any more information than the DfE already had.

Anonymous journalist (2014) Welcome changes to Ofsted inspections. Education Journal, (195), 3.

(comment by journal). Changes to inspection framework were welcome and having systme more sensitive to local school needs made sense and also inspections that were measuring more aspects that are known to be relevant to what makes schools effective. They back up AS calls for strong accountability in a highly autonomous school sector.

Barber, M. (2004) 'The virtue of accountability: system redesign, inspection and incentives in the era of informed professionalism'. Journal of education, 185: 1-38.

Barber argues strongly that accountability has led to historical improvements in the English education system as a whole and that it has helped inform teachers of the underpinning evidence and standards that mark them out as professionals. He argues that in the 70s and 80s some schools at the 'lunatic fringe' were strongly criticised but that the teaching profession was too vague about its own purpose to mount a sensible defence. He argues that many aspects can be quantifiably measured and externally aggregated, such as literacy, numeracy levels, but that other aspects, e.g. students' self-confidence, community engagement are better assessed by an inspectorate. An inspection system also provides feedback to Government about its own policy interventions, including early indicators of how a policy is being implemented and perceived by teachers (he gives the example of National Literacy Strategy in 1998 being informed by inspections in three primary schools. Inspections can also help identify cases of good practice (what works) and can set up a national database. Inspections can also greatly help (above just looking at the data) focus interventions in failing schools, by giving a more detailed picture of why the school is having problems and what needs to be done. Re-inspection in short succession helps monitor and focus school improvement efforts, in his view. A strong, independent inspectorate also helps hold the government to account too. Barber charts changes to the accountability framework used by Ofsted, particularly the relatively recent approach of having schools meet up with a SI partner once a year to agree targets and the role of SSE. He sees the potential over reliance on SSE as a risk, since schools may not be sufficiently self-critical and other school heads insufficiently challenging to drive the necessary improvements. Barber suggests that the system has 'matured' and that top down improvements are no longer as essential. This marks a change to 'informed professionalism'. He sets out how the system has moved from uninformed professional judgement in the 1970s, to uninformed prescription in the 80s, informed prescription in the 90s and now informed professional judgement in the 2000s. He argues that accountability needs to be even stronger in this system, but sharper and more precise. He suggested the need for a cultural shift towards diversity, choice, customers, best practice, continuous development etc. and that this puts the locus of responsibility more onto teachers and school leaders. He questions wehter 2001 was too early for this cultural

shift and also worries that standards in literacy and numeracy might slip backwards. He also questions whether the emphasis on SE will prove beneficial.

Barton, V. 2005. Inspection and school self-evaluation : a practical guide for school governors, Ely, Adamson.

This concerns the inspection framework (the common framework from early childhood to 19) introduced in Sept 2005. This was characterized by shorter (no more than 2 days) and more regular inspections (every 3 yrs instead of 6); less notice (2-5 days), led by an HMI; more emphasis on self-evaluation, shorter reports, letters to pupils about what the inspectors found. Outside contractors provided non HMI inspectors. The new system was designed to 'raise the bar' and schools would need to conduct rigorous self-evaluation (SEF). Governors would be responsible for the SEF and final reports addressed to them. The 2003 Every Child Matters legislation (2003) would be specifically addressed by inspectors. All local authorities would be expected to coordinate delivery to children and Ofsted would conduct multi-disciplinary inspections, across Las. Individual school inspections would feed into judgements about local provision.

Some inspections would be subject-focused, particularly for colleges and secondary schools. These would not name the school but Ofsted would provide a letter and this would need to be shared with an LA link. Separate inspections would look at inspection of religious worship in voluntary aided schools, looking at the distinctive ethos, collective worship; religious education and leadership and management. These would be graded on a 4 point scale. The SEF would be designed to evaluate the current position of the school and also 'actions designed to improve it' (p.11). Therefore, Ofsted would be able to use this as a basis for assessing whether the school's actions were having the desired effect. This was seen as quality assurance through self-evaluation. Ofsted supply an SEF but schools were meant to apply creatively to suit their own context and the SEF would merely record the process. School with well developed SE should cross reference with the Ofsted framework and link each with evidence. Schools would need to grade themselves on the 4 point Ofsted scale for each of seven sections, which include pupils achievements, personal development, leadership, efficiency (and effectiveness), and taking into account the views of pupils, parents, and other stakeholders. The governing body would additionally have to set out how they were meeting statutory requirements, such as the curriculum, equality; pupils with learning difficulties. The SEF should be completed within the schools normal cycle of review and planning, at least annually. Ofsted make it clear that the SEF is not SE, it is just a record.

Inspectors would access an uploaded SEF prior to the inspection visit. They would use to base their Pre inspection briefing and to prioritise for the visit. The inspectors then follow the evidence trail from the SEF and validate this. They would do this through lesson observations, interviews with staff, pupils and others; tracking SE and performance management; samples of work, joining meetings; analyzing records and policies. After the inspection, the governing body are advised to think about: how to communicate (through themselves or head teacher) the news to pupils, parents and staff. They also need to think about the wider community, including local press. It would be left to the governing body and headteacher to decide what to do about the report and ow to respond to it, apart form those in SM. In the latter case, head and chair of governors would attend a SI seminar provided by Ofsted and to meet with an HMI. These schools would then receive monitoring visits.

Brauckmann, S. & Pashiardis, P. 2010. The clash of evaluations. International Journal of Educational Management, 24, 330-350.

Non empirical, although takes Cyprus as a 'case study'. Looks at conditions in the system in relation to inspection and self-evaluation. Cyprus system is strongly, centralised; principals have no control over funding at all and no say over appointment of personnel in their school. Authors argue that this strongly favours external evaluation over internal. Teachers are required to produce a report on their contribution to school life but in practice this is not used as appraisal or school improvement; merely bureaucratic. System is primarily geared to identifying (by inspectors) who should be promoted. Systems of evaluation: 'whole school' inspections inspect the work of teachers, deputy principals and principals but these focus on the latter and are geared up to promotion prospects.

"Furthermore, inspectors take part in curriculum development activities, the production of textbooks, the identification of other curricular resources, the setting of examinations" (p337). Therefore lack independence from Min. of Education. Inspections of teachers are supposed to happen twice in first year of teaching but thereafter teachers can wait 12yrs for the next. The article goes on to evaluate a new appraisal system for teachers and schools which is meant to balance external and internal controls and reflect process and product.

Findings: Cyprus is governed by bureaucratic processes that do not meet evaluation expectations. Principals' reports do not discriminate between teachers; principals frequently draw on standard descriptors for these reports for all teachers.

Grades awarded by inspectors have a very limited range with all teachers being awarded above 32 points out of 40 and the great majority of teachers being given 35, 36 or 37. This means that age and seniority become the actual discriminants of teacher effectiveness.

The over emphasis on performance management means that summative assessment of teachers hinders the development afforded through formative means. Inspectors make too few observations (4 during a school year) to make adequate summative judgements too and there is no training on how to conduct classroom observations. Authors suggest a new appraisal system should take into account context and should involve participation of all staff. It also recommends the use of ranking of schools (English league tables seen as good practice).

Chapman, C. 2000. Improvement, Inspection and Self-Review. Improving Schools, 3, 57-63.

This article highlights the contrast between a reliance on Ofsted to improve the school and self-review to improve teaching and learning. The assumption is that effective professional development (along the lines proposed by Joyce and Showers: Combining: Exploration of theory through discussion, lectures etc.

. Demonstration or modelling of skill.

. Practice in simulated classroom. e.g. teaching small groups of students or peers

. Peer coaching which includes observing and working collaboratively to develop practice) (see p. 57) will lead to the greatest improvements in pupil attainment. Ofsted inspections and self-evaluation are seen as bi-polar opposites in that the former is external, top down and based on pressure, while the former is internal, bottom-up and supportive. Chapman suggests that self-evaluation can be lacking in that teachers do not always have the necessary expertise; may not be able to identify needs or may not have the necessary challenge in their collaborative partnerships within school.

Some of the conclusions derived from the article were based on a case study of a secondary school in Birmingham that had faced a recent Ofsted inspection and in which 14 of its teachers (out of 50)

had also taken part in a locally funded initiative to systematically drive professional learning through peer observation.

The OFSTED inspection (the school had been recently inspected), aimed to improve the school by providing information about strengths and weaknesses and also to assist in planning, review and improvement through rigorous external evaluation to provide key action points. The author suggests that the effects of inspection occur before, during and after the inspection:

Before: Teachers can feel isolated from planning for the process, so some prof development exercises to help prepare can be useful; secondly, the stress about being observed and extra paperwork needs to be addressed and planned for/supported by senior leaders. In the lead up to the inspection the school staff reported that staff within departments worked extra hard to prepare, understand their classes and to share resources. In previous research by Brimblecombe (1996) however, a quarter of staff reported that they planned 'safer' more teacher-led classes to avoid the possibility of loss of control during the observation. In this case study school, this did not seem to be borne out however, and perhaps teachers were more confident, as many lessons (according to the report) had practical elements and involved independent learning. During the inspection, the author noted that some of the teachers did not want to receive feedback and the receptiveness towards this feedback by the inspector very much depended on prior experiences of inspections. However, he felt that this was an important part of the experience if teachers were to learn from external feedback to improve teaching practice. The author suggested that at best, the inspection could have limited but positive changes to classroom practice and at worst, long term and negative, particularly due to stress and falling morale in the wake of a negative inspection outcome. By contrast, the peer observations, which were supportive and non-pressurized, led to improvements. Where there was one excellent and experienced teacher paired up with an NQT, initial learning was one way, but later as trust was established, the more experienced teacher was able to improve e.g. by using different teaching methods viewed in the NQT's class. The environment of trust was seen as vital in this relationship. The author suggests greater integration between inspection and self-evaluation as a vehicle for improving schools further, as the greater challenge of external evaluation may be complementary to the supportive and bottom up features of self-evaluation.

Coryton, T. (2014) HMCI launches Ofsted's Annual Report. Education Journal, (220), p 14-18)

MW speaking at the lauch of the 2013/14 annaul report for schools, FE and skills. Primary schools had improved but secondary school improvement had stalled, he claimed. Inequality gaps in particular had not improved for secondary students, while attainment for primary school leavers had. MW has commissioned a report into KS3 progress as a result of this and comments from inspection reports about negative culture of learning and low level disruption and low motivation in sec schools. MW argued that many schools were "marooned in partnerships without effective networks" (p.15). in such a situation, school improvement was very difficult. Alan Wood, pres of Assoc, of Dir's of childrens services, argued that autonomy must be partneered by a 'collaborative approach and a local framework of accountability". ASCL Gen Sec, brian lightman, said that report was unfair and did not reflect the large number of changes to that affected the secondary phase. Mary Bousted, Gen Sec of ATL aid Ofsted was part of the problem – too muchy pressure on school leaders which leads to high workloads and a demoralised workforce. Also, long-standing problems with QA at Ofsted, with non-expert inspectors and unreliable judgements.

Cuttance, P. 1995. An Evaluation of Quality Management and Quality Assurance Systems for Schools. Cambridge Journal of Education, 25, 97-108.

This is an outlines the New South Wales quality assurance system. The mechanisms by which the school review system is supposed to lead to improvement are:

A catalyst for schools to develop through collaboration with relevant members of the community

Training up of members of staff in methodology of reviews that should further enhance their quality assurance and management skills to drive improvement in their own schools

The external members of the review add credibility and support for improvements that may be difficult to achieve solely through internal review

Support services for schools are included and these are designed to enable collective information in order to make systemic improvements

The reviews also serve a dual function to aggregate out information about relative strengths and weaknesses in order to audit 'system performance' (accountability)

Cuttance, P. (2005). Quality Assurance Reviews as a Catalyst for School Improvement in Australia. The Practice and Theory of School Improvement. D. Hopkins, Springer Netherlands: 101-128.

Quite a general review of international systems to ensure quality and how they lead to systemic improvements in education. References for this chapter are mid-1990s. Makes the point that the reviewers (inspectorate) need to be 'bureaucratically independent' of the schools that they review. Cites the example of NZ and England. Systems of QA can be first, second or third party: First party assurance criteria are supplied by 'producer' as a statement of quality for the 'purchaser'. These lack transparency needed for accountability. Second party are when the 'purchaser' are able to check out the quality for themselves, e.g. when a parent visits a school. However, such QA lacks the rigour required (and access to information or required expertise) to be effective. Third party QA occurs: "when design standards and specifications are verified through a process of assessment by an external independent body" (p.110). In the latter approach, the quality criteria need to be available and appropriate, with 'professional and technical standards that are specifically relevant to education' (p110). Cuttance suggests that in order to drive systemic improvements all schools need to be included in the inspection system, not just underperforming ones. Cuttance suggests that schools' outputs can be measured in terms of:

Cognitive outcomes - curriculum-based knowledge and skills. The acquisition of propositional knowledge, knowledge application, higher-order problem solving skills and the development of the capacity to construct knowledge from constituent elements and contexts. Aspects of language and mathematics are assumed by most school systems to be an essential focus of such skills, but the wider curriculum includes elements of science, social and human systems, technology, the arts, and health.

• Affective outcomes - the development of personal and social skills and self-knowledge.

- The development of student attitudes, values, self-worth, communication skills, leadership skills, collaborative skills, etc..

• Social outcomes - relating to the role of schooling as a socialising, selection and control function in society.

- Attendance, suspension and expulsion, behaviour in the classroom and other parts of the school, retention rates, post-school destinations, and community participation, all reflect aspects of this function of schooling.

He also says that the literature on school improvement suggests that learning outcomes can be ascertained in three ways:

The level of achievement of students in relation to externally established standards (standards-based performance).

• The relative change over time in cohort differences in student achievement and the differences in achievement among groups of students (cohort-based and equity-based analyses of school performance).

• The progress made over time by students attending a particular school (the value-added by the school). (p117).

Ferguson, N. 2000. Improving schools and inspection the self-inspecting school [Online]. London: Paul Chapman.

Answers various issues on the basis of evidence gathered from a series of Nuffield Foundation funded research projects into the impact of OfSTED inspection in primary and secondary schools since 1993 (Ouston, Fidler and Earley, 1998; Ouston et al., 1998; Ferguson et al., 1999a). Many schools were found to use an LEA consultant to help formulate the school action plan and headteachers felt that this was particularly helpful. Ofsted did not give feedback on the quality of the acton plan though and follow up and support for implementation was largely absent, lessening its impact on school improvement.

Key issues (on inspection reports) have been analysed in this series of studies and mostly these gave impetus for aspects that the school had already considered. In a small number of cases the ki was difficult to meet due to resourcing implications. Most headteachers felt that the inspection told them what they already know but in a 1998 study of primary schools, heads had to predict two Kis before an inspection and only 10% had 2 issues that were in the final report. However, heads of primary (and secondary schools) are very good at predicting the issues that would be high on the inspectors' agendas. This may have a negative implication for the use of self-evaluation. Kis were also criticised for being too general and open to interpretation. The authors suggest including a member of the LEA in the inspection so that the latter can better understand how to help in the support of the school.

Ferguson, N. & Earley, P. 1999. Improvement through inspection: a better system? Management in Education, 13, 22-28.

Refers to Nuffield Foundation funded research at the IOE (as above). Tenet 1: using pre-inspection period to provide impetus for change. The long period of anticipation led many heads to focus only on issues that they anticipated Ofsted would be inteersted in and were reluctant to open up a can of worms by starting (and not finishing) new initiatives. The authors suggest having shorter periods of notice and to have an aspect of the report where Ofsted inspectors could report on areas under development. The functions of accountability and improvement should be separated out. This could include a separate consultancy function in another phase, by inspectors. This is because the high stakes accountability aspects can have very negative effects on teachers (stress, illness, time taken preparing for inspections etc.). Expertise of inspectors is seen as crucial, this should have increased training, use of specialists (not secondary teachers inspecting primary settings); clearer identification

of competency of inspectors; examining inspectors' claims of competency, e.g ability to write clear English. Other recommendations include gving more useful and extensive feedback to teachers and also on the quality of the school's self-evaluation. The greater encouragement of self-assessment

Gray, J. 2000. Causing concern but improving: a review of schools' experiences, Nottingham, DfEE.

The report seeks to address concerns about schools' lack of knowledge about what to do about 'school failure'. They acknowledge that there are other aspects to failure, which concern a lack of resources, personnel or lack of will.

Stats: by summer of 1999 900 schools had been put on SM, or 3% of sec schools and 3% primary; 8% special schools; 6% PRUs (Ofsted, 1999: 54). 40% had been placed in SM due to failures to address previous reports and national indicators (according to same report). Great pressure to improve and a 2 yr window is given. In 1998, 760 schools in SM (SM introduced 5 yrs earlier); 7% closed; 22% out of SM (71% still failing). However, HMI visits suggested that the remaining schools were doing quite well. Thus approx. 91% could be expected to emerge from SM successfully. This would take 20-22 months for special and primary schools; 27 months for secondary schools. Schools in highly deprived areas took longer (about 4 months extra for primaries). Schools in SM were disprorortionately in areas of high deprivation, with many more children on FSM compared to national average, for example. Consistent issues of SM schoolsas judged by Ofsted: under achievement of pupils, unsatisfactory teaching, ineffective leadership. HMI seek evidence of improvement in relation to:

Pupils attainment on national tests and examinations

Levels of exclusion

Attendance

Progress in lessons

Pupils behaviour and attitudes

Proportion of teaching judged satisfactory or better (as observationserved and in SOWs).

Evidence of effective leadership and management

Regarding leadership, SM schools usually have a new Headteacher, either shortly before or shortly after the visit. The report refers to Ofsted evidence that schools coming off SM have improved exclusions from 20% to around 1%. The proportion of satisfactory teaching improved from 50 – 75%. The attendance improved by 2% in secondary and special schools and 1% in primary schools. Attainment evidence harder to ascertain as inspectors need to look for trends, over 3 yrs. Improvement trajectories for 3 special schools were looked at (Gray 98 and Gray et al 99). These show sharp improvements in effectiveness after inspection over 2 years then slower improvement (or plateau) for 2 years after re-inspection.

Key conditions for change:

renewed commitment – by teachers, acceptance of need to change; involving parents and governors in supporting action plan and use of external support

Re-staffing: SLT change; new but experienced teachers

Context: clarification of schools' future (e.g. whether it will close); reduced competition form local schools; lack of extreme deprivation; short history of problems.

Leadership:

New leaders can turn threats into opportunities for change and improvement

Skills: forming a vision; monitoring; collegiate skills; gaining cooperation of staff; developing staff; resource management; ambassadorial (representing schools to LEA, or other bodies (diocese for example) and face of public.

Senior leadership team is important, not just the Head referring to Gray et al 1999, dual leadership for example.

Responsiveness to improvement after SM may depend also on the type of culture at the school (Myers 1998) : 'striving' schools, here, although they accepted serious problems they did not accept Ofsted's simple assessment and were determined to prove them wrong. In Swaying schools, staff morale dropping meant that success was far from certain but new Head was sometimes able to change this round with additional support. In 'sliding' schools, staff no longer believed in initiatives and sought to put out fires and solve short term issues but reverted back to normal routines. Authors suggest that there are other typologies but that improvement strategies need to be tailored to the circumstances of the school.

The report outlines a number of strategies that can be used to improve teaching, teaching support, staff development, behaviour strategies, performance strategies etc.. All of which can lead to improvement. Contextualised analysis of performance is very important. SE requires time and training to get right.

The report raises concerns that strategies to get the school off SM do not necessarily signal the change to an 'improvement culture' in the school. Many schools use tactics, such as concentrating teaching efforts on borderline students, which may not be a good long-term strategy for improving learning of all students. If the school has not managed to make a cultural change this may augur badly in the long term. The same is true of engagement with capacity building.

Gray, J. 2004. Frames Of Reference And Traditions Of Interpretation: Some Issues In The Identification Of 'Under-Achieving' Schools. British Journal Of Educational Studies, 52, 293-309.

Examines language and underlying rationale in the criteria describing under achieving schools. Gray identifies three fallacies: 'low performance' = 'under performance'; improved performance is mainly due to the school's efforts; and that evidence on pupil background having a lasting effect on their performance over time is clear cut (it is mixed according to Gray). Gray points to the fact that socially disadvantaged children attend poorer schools overall, leading to under achievement in some schools. The same children also lack access to privileged classmates, good teachers and other resources (facilities). Gray argues that simply looking at 'low performance' will lead to the mis-diagnosis of many schools as 'under-performing schools'. He goes on to discuss the problem that 'home' factors are very good predictors of subsequent performance, however 'prior attainment' is even more powerful; the two factors are confounded, so although pa adds to explanatory power, home factors may not. The addition of the term 'under achieving schools' is view by Gray as a corrective, which takes into account the starting point of pupils and their progress and is a welcome feature. A judgement based on overall standards, however would conflate low and under performance. Gray is cautious about the value added statistics given the degree of complexity in how to interpret them and how to estimate 'under performance'. Another problem is that within schools there can be differential effectiveness, such that certain groups are under performing while others are not.

Grek S and Ozga J (2012) Governing through Learning: School Self-Evaluation as a Knowledge-based Regulatory Tool. Recherches sociologiques et anthropologiqes 2012/2 83-103

This is a non-empirical paper that discusses the rationale behind Scotland's School Self Evaluation process. Schools are encouraged to use self-evaluation using tools provided to them centrally by HMIE in 'how good is our school (HGIOS)' framework which looks at i) successes and achievements ii) the work and life of the school iii) vision and leadership. This combines: quantitative data; views of various people and direct evidence through observations and documentation by professionals. Quality indicators are set out for schools by which they are judged in inspections. Self- evaluation is strongly promoted by HGIOS and it is made very clear that it is only worthwhile in as much as it leads to changes and improvements in practice and outcomes for pupils. Teachers are expected to work in teams, peer review and engage in professional dialogue. In the NIM (new inspection model) inspections are proportional, in that those who perform well, are given less frequent and less intrusive inspections. Also, inspectors can disengage from their evaluative role and engage in development after a day and a half (of a five day inspection), the latter involves discussion of issues of professional development and in support of areas already identified by staff in their self-evaluation. A key part of the inspectors' judgement is 'how well they know themselves', so the quality of their self-evaluation determines this outcome. The authors argue that the NIM shifts to one in which knowledge production is seen as crucial and supporting schools to be learning organisations is a key aspect.

Gustafsson, Jan-Eric; Lander,Rolf & Myrberg, Eva (2014) Inspections of Swedish schools: A critical reflection on intended effects, causal mechanisms and methods. Education Inquiry 5:4, 461-479.

Abstract: The article outlines a programme theory for the Swedish school inspection as part of the broader project about the impact of inspection in six European countries (Ehren 2011). 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 term simply 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.

Data used was taken from inspectorate documentation (Skolinspektionen 2008, 2009a, 2009b, 2010a, 2010b, 2011, 2012b) which explain the procedures and expected effects and a sample of reports of regular supervision and thematic audits. Regular supervision is the inspection of individual schools, normally every five years with a differentiated procedure depending on judgements based on schools results. Thematic audits use visits to a sample of schools to collect evidence on themes such as 'leadership' or a school subject with the objective of building up a knowledge base that can be used to contribute to the improvement of schools across the country. The Swedish National Financial Management Authority (ESV 2006) outlined a programme theory for the NAE school inspection. We also met with some Swedish inspectorate (SSI) officers to obtain their views on possible elements of the theory. 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.

Based on analysis of documentation, the authors provide 14 'if…then…because' statements of the mechanisms of inspection:

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.

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.

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. (pp 468-9)

Taken together, the authors say that, if valid, these mechanisms would lead to school improvement. However there is a lack of research in Sweden to test this. The authors point to such research as is available, in Sweden and elsewhere to point to evidence either in favour or as barriers to the effects of their suggested mechanisms.

Hargreaves, D. H. 1995. Inspection and School Improvement. Cambridge Journal of Education, 25, 117-125.

This is a discussion piece that examines the proposition that Ofsted inspections have the ability to improve schools, particularly in terms of cost-effectiveness. Hargreaves cites the cost (in 1995) of £100m per year for inspections. Three logical ways that inspections could lead to improvement: 1) governors and teachers will take up improvement initiatives in the school as a result of the inspection process 2) schools 'wither away' as parents take pupils away from the school 3)external bodies take over the school to transform it. He suggests little evidence (at the time) for 3) and that not enough parents overcome obstacles in terms of travel, relocation, or lack of available alternative, for 2. Therefore, efforts depend on the first proposition.

In this regard, Hargreaves suggests that, while inspections serve as a very useful audit and inspectors have great expertise, there are also problems: inspections lead to 'gaming' in the way that internal audits do not; weaknesses are not hidden in internal audits; less likely to lead to defensiveness. Overall, inspections give the impression that teachers and local education authorities cannot be trusted. Also, audits affect the thing they are measuring, people set up the system such that they can become auditable. Finally, Hargreaves make the point that, as schools internalised Ofsted criteria, they tend towards conformity, thus decreasing diversity in the education system. Hargreaves points to the evidence that joint work (and trust) is needed to make the difficult changes required to improved failing schools and this needs external support. The aim should not be quality control (failure has already happened) but quality assurance. The correct system is not good value for money. Recommends: identify good practice and offer advice, promote and lead discussion on the nature of quality and how to promote it; quality assurance role and checking of school's own mechanisms of continuous improvement; disseminate inspection skills; all inspectors should spend time in school (a third at least); make the identification of failing schools clear so that these can be intensively supported to improve.

Klein, G. 2000. Improving Inspection for Equality. Improving Schools, 3, 38-43.

This is a commentary piece about the Government's response to the MacPherson report (in the wake of the Stephen Lawrence case on institutional racism) to give Ofsted a key role. The author

refs to research by Osler and Morrisson (already summarised) that suggest that race issues were a very low priority for inspectors. They refer to prior evidence suggesting that race issues are brought up in only 2% of reports although bullying comes up in 17%. Inspectors felt more training was required and Headteachers also wanted some guidance on how to monitor race issues in schools. The article mentions that Ofsted has long had the powers to look at the access to curriculum and learning opportunities for all students. However, few inspectors comment on, for example, how setting and streaming my limit such chances for some students and reinforce stereotypes of what some students can achieve. Klein suggests that Ofsted should engaged in dialogue about the self-evaluation tools that schools use and learn from existing examples of good practice that support the idea that measures to improve equality can also improve overall quality.

Learmonth, J. (2000). Inspection: What's in it for schools? London: Routledge Falmer.

Ch. 4 Does inspection help schools in difficulty?

Author mentions that Ofsted report in 2000 gives the impression that schools can get out of difficulty entirely under their own steam and that any mention of external factors that may contribute to their difficulties would be seen as 'making excuses'. Referring to the number of SM issued to schools, author cites a study by Levacic and Glover (1994) concluding that schools in disadvantaged contexts were more likely to have an adverse Ofsted report – around two thirds of schools in SM coming from areas of social disadvantage (Ofsted own figures). He suggests that school failure is an interaction between school, LEA and national policies, therefore unfair to blame just on individual school and its staff.

Learmonth cites Myers and Goldstein (1998) categories for failing schools: 'striving', 'swaying', and 'sliding' but suggests that even schools not in SM may fall into these categories. Furthermore, each school is individual and needs to be treated as a special case with unique circumstances.

Attributing blame does not always work and can compound the problem through the damage in morale. He suggests that Ofsted is a political organisation designed to 'inspect' rather than improve and that the quasi-market that emphasises school choice is making it increasingly difficult for some schools to succeed despite their efforts.

Citing Thrupp (1999) he makes the point that schools have less impact on students' achievement than is often claimed. Therefore emphasising school autonomy and 'self-improving' systems' tacitly supports a political system that removes administrative and funding support. Citing also James Coleman (1966), he points to the importance of social capital in determining outcomes for children and suggests that one approach may be for schools to try to influence this (referring to work by Gene Maeroff, 1998).

Learmonth charts the rise of school effectiveness and improvement research in emphasising how it is possible for schools to get the same level of achievement regardless of social disadvantage but suggests that to do so for the same costs is unreasonable. He also shows how Ofsted have been ambivalent about this issue, in that they write about the effects of housing, health, economy in the local area and accept that this means some schools do not have the internal capacity to improve but their framework does not allow enough flexibility to take this into account. Learmonth suggests that school should take responsibility, but that the report and considerations about SM need to more adequately reflect local challenges.

Learmonth suggests that Free School Meals (FSM) is a good indicator of social disadvantage as it correlates highly with several indicators (p.69 for list, e.g. children from large families, no or shared

WC; no one employed in family). Learmonth cites an Ofsted report (see p70 of book) that outlines the challenges for schools with a high proportion of FSM pupils:

Chasing absenteeism
Highly mobile population
Low achievement on entry
Previous school exclusion
Poor English
Lack of experience of schooling
Trauma from refugee background

The effect of naming and shaming, despite the above factors, leads to a bereavement process and other types of emotional damage which may be disruptive. Ofsted names and shames, then leaves the improvement to other parties, such as the LEA and the school stakeholders.

Isolation: Learmonth cites work by Kathryn Riley (1998) that suggests teachers in SM schools are quite frequently isolated from their colleagues, possibly reflecting the overall climate in which schools tend to isolate themselves from each other in a competitive market created by successive government policies. Citing Rosenholtz (1989) and Reynolds (1995), 'Stuck' or failing schools may have the characteristics of blaming their pupils and a lack of basic competencies to do what is needed, failure to take responsibility to change things (someone else's job), lack of knowledge of alternatives, fear of being exposed to the public, dysfunctional interpersonal relationships. Reynolds suggest more informed interventions, with greater awareness of such issues:

Rebuilding relationships
Information provided on improvement
Building evidence from various stakeholders
Easily achievable goals
External agents come and deliver the harsh truth

Ofsted themselves (2000 report cited on p.79) state that schools need external support for training of staff and organising this training; help to priorities areas, swift action on leadership and support, not just constant monitoring.

Ch.5 Does inspection help schools improve?

Learmonth presents the idea that the New Right agenda was reinforced to an extent by school effectiveness and improvement research; it put the onus on the teacher and the school to make improvements and made a case for socio economic status having no impact on student achievements. The converse (in the 60s) was possibly the case, in that crude expectations of students based on SES led to lazy generalisations and disempowered the school from its role in making a difference. Citing Reynolds and Stoll (1996), Learmonth says that school effectiveness research is not always compatible with school improvement research. The latter sometimes fails to learn from the former and teachers are not sufficiently in tune with research to make sense of it. Citing several thinkers (Schon, Stenhouse, Hargreaves, Fullan) he thinks that teachers should have a more active research role.

Summarising work by Gray (1998) and Sammons et al (1997), he says that there is more variation within a school than between, schools make up about 10-15% of variation in pupils' performance,

the majority of schools (around two thirds) are doing as well as might be expected given the starting point of their pupils and subject departments vary considerably within schools, only a minority have strong depts across the board.

Author cites work by Stoll and Fink (1996) to outline what might be considered as 'school improvement' (p87), including enhancing pupil outcomes, a school taking control of its direction, changing the culture or building capacity to change. There are also 'doors' to getting started on improvement, some of which are external (including inspection or national strategies) or internal ones, including self-evaluation or building collegiality (Joyce 1991, Stoll and Mortimore (1995) on p 88). Barth (1990) also compares externally mandated change to internally derived and suggests the former is indicative of Ofsted's approach, ie. Change is directed at improving pupil outcomes on standardised tasks, teachers can be trained to do what is best practice and school improvement is about getting teachers to do what has been identified as needed. He contrasts this with the notion of where schools use own capacity to change, creating conditions for this to occur and improving the culture and interpersonal qualities. Author suggests that the right type of support should be identified based on where the school is placed at the moment or depending or using different approaches from time to time in the same school. Overall four key areas emerge from SE and SI literature:

Quality of Leadership
Managing changes to the culture of the school
Focus on teaching and learning
Systematic professional development of staff

Regarding leadership. Author says that Ofsted/DfEE (from Stark, 1998) evidence did not point to charisma as being important rather, strategic skills, monitoring, collegiality, staff management; staff development; resource management ambassadorial skills

Managing changes to the culture of the school. Variety of different classificaitons of culture are identified. Author cites Stoll and Fink's (1996) features of a culture that lead to SI: shared goals, responsibility for success; collegiality; a belief in continuous improvement; lifelong learning for children and adults; risk-taking as essential part of growth; mutual support; mutual respect; openness; celebration and humour (p. 94).

Learmonth also cites Hargreaves, D (1999) three core capabilities in the culture of an effective school: Monitoring (scanning and auditing the environment for pressures and potential etc.); Proactive: 'can-do philosophy, long as well as short term thinking; resource deployment: directing towards key purposes of schooling.

Focus on teaching and learning: some barriers to improvement here are identified by Learmonth, such as overly prescriptive systems that dictate one method for all teachers and classes; over reliance on assessing learning and analysing what this means, the tendency of teachers to be territorial about own classes and lack of time to support improvements to teaching and learning

Systematic professional development of staff: i) Need to look at the impact of PD on pupils' learning and achievement. Cites 200) DfEE paper on Professional development: support forTeaching and Learning' that emphasises the school and the classroom as the basis for PD and quote by Joyce et al (1997) that suggest the need to radically reformulate teachers' time to allow for collegial activity to continuously improve and study learning and pedagogy (see p. 99). Ii) monitoring and coaching relationships: needed in order to incorporate and master new teaching ideas and techniques into

practice. Iii) appraisal: Learmonth suggests that performance appraisal and improvement has not yet been adequately established but that some form of appraisal of professional learning is needed iv) Developing and understanding skills of the change process:

Fullan (1992): initiation, implementation and institutionalisation. Author suggests adding 'evaluation' to the list. Fullan and Hargreaves' (1992) guidelines to help develop a school learning community with the Head as 'lead learner': including: develop a risk-taking mentality; trust process as well as people commit to working with colleagues; encourage interaction between departments and colleagues. (p101).

School inspection and improvement:

Evidence is mixed and restricted as to whether Ofsted leads to improvement, some such as Cullingford and Daniels (1999) show dips in GCSES in year of inspection of schools. Lonsdale and Parsons (1998) showed that the process was disruptive and demeaning and questions its worth. Ouston and Davies (1999) question whether improvements where shown, could be achieved in a more cost effective manner. Citing work by Scanlon (1999) surveys of Heads and teachers suggested support for idea of Ofsted but that it should provide more than a snapshot, be more supportive, SE play a greater part and more LEA support.

Ch 6 What kind of inspection would help to raise standards?

Author suggests a framework for inspection that is based on the above chapter standards and questions whether recent (including 2000 change) to framework are based on SE and SI research. Learmonth paints a picture of LEAs often not performing adequately to improve teaching and learning standards in failing schools and the trend to outsource SI support but suggests that the jury is still out in this area.

Overall, Learmonth suggests a number of features of a framework for inspection that is needed, long list, including: (p114): consistent framework that allows for sufficient flexibility to take into account individual contexts; one that satisfies external accountability; complementary roles of inspectors, LA, government etc.; detailed collection of evidence to paint a picture of the school as it 'typically is'; complementary evidence from a range of stakeholders that know the school well; internal and external mechanisms to support schools before they get into a crisis; system that makes it clear to public the degree to which the school is responsible for its strengths and weaknesses; proportional to the situation of the school; continual training in evaluation, since doing this reliably is very difficult; recognition that internal self-evaluation can be as or more rigorous than external evaluation; complementary internal and external evaluation processes.

David Hargreaves (1995) suggests that internal SSE is not rigorous enough, too parochial and staff lack the training to do this well, therefore external needed too.

Accountability:

Professional accountability (Eraut, 1992) is a laudable aim but current trends towards external accountability have weakened teachers feeling that they belong to a profession that holds itself accountable. Initiatives to promote internal SSE, Peer review and teacher research (e.g the Best Practice Research Scholarships) are seen as complementary to such efforts. Collaboration and openness is essential, otherwise the way that professional accountability is defined can become too narrow. Using serving Heads and senior staff as AI's is important but training needs to be more extensive.

Supported self-evaluation:

Trond Alvik (1999) three categories of internal and external SE:

Parallel: both are conducted and findings shared

Sequential: SSE first and then external uses this as a basis to conduct review. Or external review first, and this provides information to the school to do its SSE

Cooperative: Discussion and negotiation of the process

Different patterns internationally:

Scotland: Moving from parallel to sequential. Tools provided to school in 'how good is our school? Self-evaluation using performance indicators. This includes student views on the quality of educational provision. Only a sample of schools is inspected:

The education authority, governing body, school board and parents give view of school prior to inspection; draft report of the inspection is discussed with the head teacher, director of education or chair of governors or school board; recommendations form agreed agenda for a follow up visit in 12-18 months time. Follow up report is a letter to the director of education or chair of governors; range of inspections designed to evaluate the authority and its quality assurance; headteachers and staff complete survey on inspection at the end of the process.

USA: Example given of the David Green's Chicago School Change through Inquiry Project (SCIP) and Tom Wilson et al in Rhode Island.

SCIP: Both internal and external reviewers are termed 'enquirers', rather than 'experts'. Quality review training includes various parties: state education dept, local principals of schools and representatives of business and industry. Trained in Teaching and learning; student learning, progress and achievement and schools as learning communities.

External quality review is a five year cycle, followed by annual review conducted by the school.

All parties encouraged to be honest and to see the school over time rather than just a snapshot. Evidence is interrogated and discussed and results published and shared with a wider community.

Rhode Island follows SALT (School Accountability for Teaching and Learning). These follow several activities:

Self-study (school based inquiry activities)

SIP, which combines evidence and professional judgement

School visits: four days, teachers, a parent, school administrator, a school ctte member, HE member, a member of Rhode Island Ed. Dept. This occurs every five years. Team writes a report based on three SALT areas and recommends areas fro improvement.

A compact for learning is drawn up by school, district and State Dept of Education. This ensures the school has capacity to make improvements and specifies what district and SDof Ed will do to support the school.

School report night: schools report to parents and community once a year one plans and how they've improved and what community can do to support it.

Lesson observations take place in self-study and school visits. A variety of other evidence is used, including surveys, test scores, programme evaluation. Review team is trained in inspection and audit methodology, report is designed to persuade rather than prescribe, school visits include shadowing students, carrying out surveys and interviews. External reviewers are seen as 'learners', they have discussion with teachers about teaching and learning, rather than ER's prescribing against a checklist.

Channel Island: Validated school self-evaluation system. Features:

Overall framework for development and review of schools (Ofsted style model, agreed by representative parties and externally evaluation of whole scheme.

Funding made available to schools to support self-review and implementation of action plan

Training for staff on internal SSE and validation process

Register of approved validators, including officers nad advisers in the Islands, heads and deputies, heads of dept, subject coordinators and Ofsted trained inspectors.

Teachers can develop evaluative skills in other schools and share effective practice

The Netherlands:

HE influences SSE: Uni of Twente's work on school effectiveness contributes to sequential system: first school does SSE, then external reviewers validate the school's procedures. Inspectors have to coordinate advice networks from a range of providers, including LAs as well as offer advice. Extensive three month training with one year top ups. Inspectors in charge of 100 primary schools or 30 secondary. Schools inspected annually, varying from half to one day.

Victoria, Australia:

Works on three principles:

External evaluation best when combined with well-developed school SE

SSE without external Evaluation lacks rigour to effect lasting improvement

Evaluation should look at management practices, not just performance. Emphasis on analysis of data and MIS

3 key elements are: School charter, school annual report and triennial school review. Planning, monitoring and performance-review framework over three year period.

School charter: 3 yr negotiated agreement with Victoria Dof E and school. School annual report shows performance and school review reports full three year cycle. School self-assessment and external verifier provided by Dept of Ed. Reviewers are trained by Office of Review. All schools carry out an evaluation prior to the review, rate performance of reviewer and suggestions for improvement in the process and comments indicated on the day and in the report. Office of review quality assure the process and the reviewers. The independent reviewer's role is crucial in setting challenging goals and improvement priorities for the next three years.

London Borough of Wandsworth:

Many LEAs have Education Development Plans that include supporting SSE. Some have own frameworks, developed with schools. These range in level of formality (e.g informal Derbyshire's Quality Development Dialogue).

London Borough of Wandsworth (LBW) has a well-established Annual School Review (ASR). Schools can apply for LEA accreditation of their SSE process. The criteria for accredited self-review (ASSRE) agreed by working part y of heads and inspectors, with Ofsted framework in mind.

Criteria:

Whole school improvement of educational standards, e..g pupil attainment, attendance and behaviour

Whole school improvement of educational provision, e.g. quality of teaching, effectiveness of the curriculum and assessment

Management and leadership, ie. In term of the above

School has to provide evidence to LEA that appropriate procedures are in place

LEA role has changed to one of raising standards rather than one of control. LEAs should not just look at schools in high risk as this may distort picture of local area and capacity.

A future inspection system

Learmonth discusses the work of OFSTIN, a group of educationists who, since 1966, meet to review the work of OFSTED.

 Six basic changes suggested by them (published in 1999):

Dispense with the contracting system and use professional, qualified and trained inspectors

Schools should be responsible for own performance, and this should be monitored by external inspectors

Inspectors' judgements should be discussed with teachers and offer advice and assistance

Less money on inspecting and more on helping schools to improve themselves

National inspectorate should be accountable to an independent body

Learmonth sets out roles of national inspectorate and LEA in future system:

National inspectorate:

advising Government

monitoring quality of school reviews and collecting national data

Monitoring and supporting LEA's priorities in their EDPs

Inspection of sections of provision not amenable to single school inspection, e.g. continuity between Key stages 2 and 3

Coordinating training programme for school evaluation nationally and regionally

"Development of criteria for judging school performance sensitive to the benefits and challenges which different socio-economic communities provide" (p. 132)

Contribute a national inspector for a sample of schools to allow central-regional flow of information

LEA or other regional agency:

Coordinate and validate SSE frameworks

Provide, with HE, training in evaluation skills and promoting action research

Monitor student achievement and intervene early if problems emerge

Coordinate and assemble local teams of inspectors and training, including to members of LEA

Contribute to national database from visits

Support headteacher and teacher appraisal and the monitor the distinction between SSE and performance management.

The school:

Develop and implement SSE in cooperation with LEA, all stakeholders should be able to contribute

Foster inter and intra school networks on professional practice

Regular reports to local community about progress and problems. These can be validated by LEA.

Conclusions:

Learmonth suggest that the infrastructure needs to be aligned and complementary, each level above supporting the one below – school-LEA- national Government. He suggest much can be learned from the Scottish system.

Meuret, D. & Duru-Bellat, M. 2003. English and French Modes of Regulation of the Education System: a comparison. Comparative Education, 39, 463-477.

Mostly a general comparison between the two systems of education so not too relevant and non-empirical. However, does make the point about the unintended consequences of inspection –i.e. that the focus on school inspection (only teachers are inspected in France, not schools) puts more pressure on teachers in general (in France it is apparently very hard for the Principal to fire poor teachers and parents reflect this concern). Authors suggest that the recruitment problems in England are less the case in France. In general the difference in performance on international comparison does not suggest England is working better in this respect (France ahead in most comparisons but not by much).

Mok, M. M. C. 2007. Quality assurance and school monitoring in Hong Kong. Educational Research for Policy and Practice, 6, 187-204.

Non-empirical. Looks at various changes over the last few decades in HK. Makes the point that 'quality' is defined in a very wide range of ways. This is quite good as it covers the needs of a wide range of issues/stakeholders but at the same time aspects can be in tension. E.g selective intake schools are not favoured by value added measures of student attainment as there is less 'wiggle room' for improvement. It also makes the point that, with a shrinking birth rate, the competition between primary schools increased and the publication of external reviews (inspections) popularity and intake could be make or break (school numbers have declined. "The number of primary schools in Hong Kong dropped from 815 schools (enrolment 15,013 students) in 2001 to 720 schools (13,353 students) in 2005 (Education and Manpower Bureau 2007)." (p.191).

Ozga, J. & Lawn, M. 2014. Frameworks of Regulation: Evidence, Knowledge and Judgement in Inspection. Sisyphus-Journal of Education, 2, 7-14.

This is an introductory article based on a special issue of the journal about governing through inspection, drawing on research from the wider project: Governing By inspection: school inspection and education governance in England, Scotland and Sweden'. The authors suggest that:

"Inspectorates are often translators of data-based system knowledge into actionable or practical knowledge for their national governments, as well as-in varying degrees-for schools, teachers and pupils."(p.9). They look at the system steering role of inspectorates. Inspectorates have a unique role in combining data and expert judgement into use by policy makers. As well as developing knowledge about improvement, inspectorates hold governments to account, staying independent from them, evaluating their policies.

ParsonS, C. 1998. The shame of OFSTED: not improving but policing. Improving Schools, 1, 38-42.

Reflections on the role of Ofsted. Puts in context of NPM and also compares with other professions that are inspected. In terms of having low esoteric knowledge, low on collective forms of working and high on public concern over competence, teaching comes out worse than doctors and the police. (Hughes, G., Mears, R. and Winch, C. (1997) 'An Inspector Calls? Regulation and accountability in the public services'. Policy and Politics 25.3 pp. 299-314). It suggests that individual and institutional inspection is not proven to drive up standards. Ofsted described as 'unethical' – lack of partnership with teachers, inspection punitive, inspectors 'mercenaries'; lack of trust of teachers, encouraging a system of surveillance.

Richards, C. 2014. Ofsted's Annual Report and its aftermath: are lessons being learned for school inspection in 2014? Education Journal, 187, 11-12

Refers to assertions by MW from the 2012/13 annual report that standards in leadership are improving, 8 out of 10 schools in England are good or better and the new MATs are using their freedoms well. He questions the ability of Ofsted to make judgements due to the quality of the evidence gained: "Unfortunately that evidence is based on a flawed inspection framework which is unduly reliant on questionable performance data, which ignores process and outcomes in large parts of the curriculum, which unrealistically expects its inspectors to be able to judge progress in short observations and which results in schools not so much teaching-to-the-test as teaching-to-the-inspection." (p. 11). Schools are increasingly gaming the system and the evidence on which to base policy decisions is not good, including the variable quality of inspection teams. He also questions the assertion that schools have improved due to Ofsted 'raising the bar'.

Richards, C. 2012. Ofsted inspection inspected: an examination of the 2012 framework for school inspection and its accompanying evaluation schedule. United Kingdom.

This is a detailed critique of key documents in relation to a new Ofsted framework for inspection and its possible consequences for schools. The authors makes a number of critical points about the new framework: the lack of reference to commenting on the quality of the curriculum. He suggests that the curriculum is a key aspect of school provision that has a big impact on pupil achievement and therefore should be commented in more detail than whether it is 'broad and balanced'. He also makes the point that if the school's aims are not clear then Ofsted will not be able, with clarity to comment on its effectiveness in meeting these. Issues such as 'quality and 'improvement' are also not admitted to be contentious and are not explained in any of the Ofsted documentation. The latter, he suggests is also indicative of the relative lack of stated aims and values of the English

education system. Ofsted gets round this by instructing inspectors to assess the effectiveness of the school in relation to its own aims and values but this raises the question about the extent to which a school has these, how clear they are (or how appropriate). The requirement of reporting inspectors to appeal to a variety (almost any) of audiences means that they are supposed to use 'everyday' language. He suggests that many terms, even in the quality rubric are imprecise and therefore more guidance is needed, e.g. what is meant by a 'challenging' task in a class or 'specific learning needs'. In terms of language, he also notes that previous references to 'standards' are largely absent and much more is now made of 'achievements' or attainment'. Achievement is generally equated to academic achievement and therefore devalues other forms of achievement. Achievement is also equated to attainment. Richards also surfaces 2 distinct aspects of 'teaching' and 'learning' that inspectors are required to make judgements on: i) task: i.e. learning as a process of engaging in something etc. or 'doing a teaching activity/strategy' ii) achievement: students have 'learned' something, i.e. they have had a shift in conceptual understanding or gained new knowledge or are able to apply new knowledge; in this sense teaching is the success in having achieved the latter among students. The achievement sense of these terms is much more problematic and difficult to ascertain through obsservation. Thus, the prime determinant of the grade is 'achievement' (i.e. attainment in examination results). Inspection of the 'behaviour and safety of pupils' is one area that Richards suggest is something that good inspectors should be able to make reasonable judgements about and that this need not require a recourse to achievement data. Richards also looks at the requirement to assess leadership and management and identifies many criteria that relate directly, again, to achievement – which he suggests is a contentious term. Thus, inspectors may start with an a priori position about the grade of leadership due to achievement and look for evidence and reasons to support this judgement. Even in the grading, i.e, a four-point grading scale: 1 (outstanding), 2 (good), 3 (satisfactory) [7] and 4 (inadequate), the first is norm referenced and the last is criterion referenced. He suggests changing to excellent. He suggests changing 3 to not consistently good and 4 to unsatisfactory due to their connotations. The lack of clarity and guidance and ambiguity in terms or use of contentious terms like 'achievement' lead to problems in the framework that mean the validity and reliability of judgements is in question. Many aspects are practically and conceptually impossible to test (e.g. the spiritual and moral development of all children).

Richards, C., Winch, C., Smith, R. & Philosophy Of Education Society Of Great Britain 2001. School inspection in England : a re-appraisal, [S.l.], Philosophy of Education Society of Great Britain.

Authors suggest that Ofsted 'promises more than it can deliver'(p2). One of the key deficits are stated aims and purposes of education which means that they cannot/should not make judgements about 'effectiveness' (in meeting these aims). Authors point out a number of limitations: The requirement to use 'everyday' language ignores the value-ladenness of many words which means that reports should not be seen as being so precise.

Standards: these are often judged mostly in relation to performance data, whose interpretation is much more controversial than is acknowledged in the reports. Judgements by inspectors also need to be seen as 'contestable' by staff at the school.

Teaching and Learning: very hard for inspectors to comment on attainment in lessons; they can only report observable expressions of engagement in activities etc..

Progress: Authors argue that there is not enough evidence for inspectors to comment on the school's contribution to improvements over time (for school or pupils).

He (principally views of C. Richards) concludes by suggesting where Ofsted should/can have a useful role:

(1) check on whether and how far schools are complying with relevant statutory requirements and ascertain and report back to central government any problems or issues arising from attempts to comply (or to avoid compliance);

(2) evaluate and report on observable features of school provision such as the state of repair of the fabric of buildings or the quality and quantity of resources;

(3) evaluate and report on the progress made, and the problems encountered, in introducing particular initiatives;

(4) evaluate and report on the effects of central or local government policies on policy and practice in schools;

(5) collect, report and evaluate the perceptions of interested parties (pupils, teachers, parents, governors) in relation to identified issues;

(6) offer possible explanations of how particular outcomes have been achieved in particular schools and disseminate that information to other schools and interested parties;

(7) offer tentative, broad-brush judgements as to how far individual schools appear to be meeting their own aims and values or the aims and values of school education in England (if these were to be agreed);

(8) offer tentative judgements as to how well lessons are conducted and on pupils' observable responses to teaching

(9) offer broad, tentative judgements about the quality of pupils' performance in particular subjects compared with those in schools in roughly comparable contexts;

(10) offer inspectors' interpretations of activities they see as a basis for dialogue with those who have been observed and who may have differing interpretations;

(1 1) validate schools' processes of self-review.

(p. 42/43).

Robarts, L. & Leslie, R. 2010. A day in the life of two school inspectors. United Kingdom. Education Today Vol. 60, Issue: 1.

An interesting article that compares the work of a school inspector in 1851 (Matthew Arnold) with one in 1997. The former had no teaching experience (he was a professor of poetry). He checked standards in 'elementary' schools (up to 13/14 yrs old) of middle and working class children at 104 schools in England and Wales. He commented on physical conditions of the school (ventilation, drainage, heating and other equipment) and also on methods of instruction, class discipline and attainments in arithmetic, reading and writing. The report was six pages long for each school and he prepared an annual report which summarised these, for the government. He also had the power to examine and appoint 'pupil-teachers. Arnold was apparently very sympathetic to the conditions and workload of teachers in large classes and would frequently make his visit as short as possible, for which he was criticised for being too soft. A revised code awarded funding only to schools for each student that was examined, and passed arithmetic, reading and writing. This set the precedent that only aspects that were examined/measured could deserved attention (and funding).

1997: Inspectors almost all have teaching experience. The workload involved in the preparation, visit and post visit were similar to that experienced in 1851, however now with a team of inspectors, working into the evening in their hotel. The visit involved lots of meetings and observations and is now much more systematic and thorough.

Similarities: nervous, terrified teachers; the lack of understanding of the broader aspects of education by inspectors; no right of reply. The language of reports in both centuries highlighted failings by schools and teachers. Both systems disrupted children's learning and encouraged narrow/shallow forms of learning, such as by rote. Both inhibited inspirational/innovative approaches to teaching. "It is a fundamental truth that education does not thrive on cruelty" (p26).

Vanhoof, J. & Petegem, P. V. 2007. Matching Internal And External Evaluation In An Era Of Accountability And School Development: Lessons From A Flemish Perspective. Studies in Educational Evaluation, 33, 101-119.

Very interesting non-empirical article which sets out to describe, define and suggest the appropriate relationships between summative and formative evaluation, quality assurance, internal and external evaluation. Quality = "meeting the expectations shared by the stakeholders in an appropriate manner" (p.103). Expectations can be internal and external, legally binding or not. Quality assurance = "an umbrella concept which covers all activities undertaken to investigate, monitor, improve – and perhaps also even to make public – the quality of schools" (p.104). Evaluations are a part of QA. QA can be for accountability or improvement purposes. Evaluations can be summative or formative. The authors make the point that by combining an accountability role and development role this can have a negative effect on intrinsic motivation (e.g Deci and Ryan, 1985). (also relates to achievement vs performance orientation in individual psychology). If the self-evaluation has an extrinsic role, this can remove the intrinsic motivation needed for school development and can lead to 'gaming', e.g. providing a falsely positive self-evaluation or hiding details in the SSE that could be 'used against' the school. The authors therefore argue for a strict separation between accountability and school development in policy; see evaluation as a process not a one-off activity (by contrast with an 'audit'); select themes carefully for self-evaluation (must be important to school and go beyond strict requirements/expectations and may involve using own framework; recognize professionalism of each stakeholder in the process; guarantee the quality of self-evaluations (both the conduct of and providing support, resources and tools); monitor the quality of self-evaluation (tricky to determine criteria but evaluation of tools is useful here). Authors set out six domains from their taxonomy:

Table 1: The Six Domains of Quality Assurance

	Legally-anchored expectations	Non-legally anchored external expectations	School-internal expectations
Accountability	Domain 1	Domain 3	Domain 5
School development	Domain 2	Domain 4	Domain 6

(p113).

Relevant domains: 2, 4 and 6 (i.e. these are to do with development):

Domain 2: authors suggest that the SSE be focused on external expectations and that inspectorate audits this but that another body follows this up (LEA or in Dutch context 'support services'), the latter 'guarantees' meeting the targets.

Domain 4 and 6: External evaluations can have a 'scope broadening', encouraging and legitimizing impact on self-evaluation. The external inspectorate could be asked to add a 'supplementary' section on aspects of their internal development, e.g. 'to shine light on part of the route travelled in the direction which the school wishes to go" (p. 115). Other bodies (support services in Holland) can help by providing training, making tools available, setting up links between schools and other parties and arranging exchange visits around a joint framework and also providing frameworks.

Waterman, C (2013) Ofsted's latest plans add to the de-construction of the local authority. Education Journal, (163), 8-9.

This is a commentary piece by Chris Waterman about Ofsted Feb 2013 consultation on inspecting LA school improvement services. The recommendations were that: the inspection would be "focused only on where it is needed most; not use a four-point scale to judge effectiveness; report against a published evaluation schedule, which will specify quality criteria; report and publish the findings in a letter rather than in a lengthy report; for those local authorities deemed not to be exercising their duties effectively, inspectors would consider whether the local authority will require re-inspection after a suitable interval." (p. 8).LAs have no powers with regard to intervening in academy schools. However, inspectors could explore whether the LA had discussed any concerns that might have with individual academies with them. The article reports that LAs are unhappy with the proposal and that the academies programme has vastly reduced funding available for supporting schools. The author suggests that this is a political attack on local authorities.

References:

Anonymous journalist (2014). Academy chains under attack. Education Journal, (193), 25

Anonymous journalist (2014). Listening to Andreas. Education Journal, (193), 3

Anonymous journalist (2014). Damning Ofsted reports on E-ACT academies. (2014). Education Journal, (196), 5.

Anonymous journalist (2013) Ofqual warns Gove on accountability. Education Journal, (162), 1.

Anonymous journalist (2013) New approach to accountability and inspections suggested at NUT national conference. Education Journal, (172), 10

Anonymous journalist (2014) New blueprint for inspection. Education Journal, (195), 1.

Anonymous journalist (2014). Ofsted to bring inspections back in-house. Education Journal, (201), 4.

Anonymous journalist (2014) School partnerships. Education Journal, (187), 20

Barber, M. (2004) 'The virtue of accountability: system redesign, inspection and incentives in the era of informed professionalism'. Journal of education, 185: 1-38.

Barton, V. 2005. Inspection and school self-evaluation : a practical guide for school governors, Ely, Adamson.

Brauckmann, S. & Pashiardis, P. 2010. The clash of evaluations. International Journal of Educational Management, 24, 330-350.

Chapman, C. 2000. Improvement, Inspection and Self-Review. Improving Schools, 3, 57-63.

Coryton, T. (2014) HMCI launches Ofsted's Annual Report. Education Journal, (220), p 14-18)

Cuttance, P. 1995. An Evaluation of Quality Management and Quality Assurance Systems for Schools. Cambridge Journal of Education, 25, 97-108.

Cuttance, P. (2005). Quality Assurance Reviews as a Catalyst for School Improvement in Australia. The Practice and Theory of School Improvement. D. Hopkins, Springer Netherlands: 101-128.

Ferguson, N. 2000. Improving schools and inspection the self-inspecting school [Online]. London: Paul Chapman.

Ferguson, N. & Earley, P. 1999. Improvement through inspection: a better system? Management in Education, 13, 22-28.

Gray, J. 2000. Causing concern but improving: a review of schools' experiences, Nottingham, DfEE.

Gray, J. 2004. Frames Of Reference And Traditions Of Interpretation: Some Issues In The Identification Of 'Under-Achieving' Schools. British Journal Of Educational Studies, 52, 293-309.

Grek S and Ozga J (2012) Governing through Learning: School Self-Evaluation as a Knowledge-based Regulatory Tool. Recherches sociologiques et anthropologiqes 2012/2 83-103

Gustafsson, Jan-Eric; Lander,Rolf & Myrberg, Eva (2014) Inspections of Swedish schools: A critical reflection on intended effects, causal mechanisms and methods. Education Inquiry 5:4, 461-479.

Hargreaves, D. H. 1995. Inspection and School Improvement. Cambridge Journal of Education, 25, 117-125.

Klein, G. 2000. Improving Inspection for Equality. Improving Schools, 3, 38-43.

Learmonth, J. (2000). Inspection: What's in it for schools? London: Routledge Falmer.

Meuret, D. & Duru-Bellat, M. 2003. English and French Modes of Regulation of the Education System: a comparison. Comparative Education, 39, 463-477.

Mok, M. M. C. 2007. Quality assurance and school monitoring in Hong Kong. Educational Research for Policy and Practice, 6, 187-204.

Ozga, J. & Lawn, M. 2014. Frameworks of Regulation: Evidence, Knowledge and Judgement in Inspection. Sisyphus-Journal of Education, 2, 7-14.

ParsonS, C. 1998. The shame of OFSTED: not improving but policing. Improving Schools, 1, 38-42.

Richards, C. 2014. Ofsted's Annual Report and its aftermath: are lessons being learned for school inspection in 2014? Education Journal, 187, 11-12

Richards, C. 2012. Ofsted inspection inspected: an examination of the 2012 framework for school inspection and its accompanying evaluation schedule. United Kingdom.

Richards, C., Winch, C., Smith, R. & Philosophy Of Education Society Of Great Britain 2001. School inspection in England : a re-appraisal, [S.l.], Philosophy of Education Society of Great Britain.

Robarts, L. & Leslie, R. 2010. A day in the life of two school inspectors. United Kingdom. Education Today Vol. 60, Issue: 1.

Vanhoof, J. & Petegem, P. V. 2007. Matching Internal And External Evaluation In An Era Of Accountability And School Development: Lessons From A Flemish Perspective. Studies in Educational Evaluation, 33, 101-119.

Waterman, C (2013) Ofsted's latest plans add to the de-construction of the local authority. Education Journal, (163), 8-9.

Annotated bibliography of literature reviews.

This annotated bibliography summarises major literature reviews on the impact of inspections.

Chapman, C. 2001. Unlocking the potential: inspection as a mechanism for school improvement. Improving Schools, 4, 41-50.

Chapman cites Fitz-Gibbon 1998 and Wilcox and Gray 1996, who conducted research into the reliability of OFSTED judgements and concluded that Government may be making decisions based on inaccurate data. Wilcox and Gray also noted that the short term effects gained in the preparation to the inspection lead to improvements that are short lived. Citing various authors, Chapman suggests that an Ofsted generated action plan has the potential for school improvement but that this depends on a number of factors, such as the schools' internal capacity for change, the judgements being accurate to begin with and valid judgements in the context of the school about effectiveness being made (depends on skills and expertise of the inspectors). If there is some overlap with school's own development plan this also helped. However, research also showed that development slowed down in most schools after inspection (Ouston, Fidler and Earley 1996). Chapman mentions feedback from lesson observations as a potential mechanism for improving teaching quality but says the success of this would be dependent on:
The ability of the inspector to identify areas for development
The interaction and communication between inspector and teacher
The willingness of the teacher to listen and implement suggestions

Citing research by Brimblecombe et al (1990) and Brunel University and Helix group (1990), Chapman shows evidence that teachers changed after inspection. In the former, 38% teachers 'intended' to change teaching style or method as a result of feedback by the inspector having been observed (the intention came before the publication of the report so it must have been the feedback). The Brunel study surveyed Headteachers and 58% of them stated that they had changed teaching styles and curricular organisation. In the latter study, this was asked after the report had been published. He also cites Lowe (1998), who stated that only 1 in 7 case study schools had 'substantially implemented inspection recommendations related to teaching and learning' (p.45); 3 had made some changes, and 3 none or very little change. He also quotes Brighouse and Moon (1995) in stating how some teachers may be sticking to their own idea of best practice and reacting against the orthodoxy of Ofsted. Chapman looks at the build up to inspection; the week itself and the impact of the inspection. He notes that previous research shows that there is generally a positive relationship with inspectors but that anxiety is greatest in the build-up and worse than the experience itself. The way that senior management prepare staff for the inspection event is also said to have a big effect (citing Shaw, Briddlecombe and Ormiston).
During the inspection: Lessons are more highly prepared than normal. About a quarter of teachers deliver more didactic lessons than normal. A fifth of teachers noted a change in their own behaviour and 50% a change in pupil behaviour when an inspector was in the class. Citing an ofsted review in 1994, chapman notes that staff were disappointed not to have the chance for discussion with the inspectors. However, the potential for increasing potential and for this to have a positive effect, is noted by Chapman as being limited. Feedback to teachers did reduce their sense of isolation and a little feedback was better than none it was felt. However, more time would be needed in the inspection process to feed back to teachers for this to be more effective in raising standards. Chapman cites previous work (Chapman and Harris, 2001) and says that schools with more positive cultures had a better response to the inspection process and that that this likely led to more improvements. Overall, he concludes that the Ofsted core purpose of 'improvement through inspection' was "a major weakness in the system" (p. 47)

Churches, R. and C. McBride (2013). "Making external school review effective."

This is a conference report for the CFBT the Windsor International conference on school improvement through inspection and external review. 2 day conference The conference presenters were chief inspectors and academics from around the world. Case studies from Bahrain, Dubai, India, Jamaica, South Africa, the UAE Federal Ministry of Education and the United Kingdom. How can inspection and review improve learning outcomes? Was the principle question of the conference. Pre-conference interviews were conducted with key members of the CfBT Education Trust; a literature review was also conducted on 635 peer review journal articles and paper between 2002-2012 with school inspection or school quality review in the title or abstract, 156 of these scrutinised in detail. This literature was related to the conference outputs in the form of a matrix. This was synthesised into 5 characteristics that underpin effective external school review:

Characteristic 1 Use of a robust review framework underpinned by research evidence
Pam Sammons suggests 10 processes that are important for school improvement:
Clear leadership
Developing a shared vision and goals
Staff development and teacher learning
Involving pupils, parents and the community
Using an evolutionary development planning process
Redefining structures, frameworks, roles and responsibilities
Emphasis on teaching and learning
Monitoring, problem-solving and evaluation
Celebration of success
External support, networking and partnership
Characteristic 2 Inclusion of parent and student views during review
Characteristic 3 Deployment of a skilled reviewer workforce
Characteristic 4 Objective evaluation and transparency
Chris Taylor from CfBT reported on work with the Government and stressed importance of some separation at government level, transparency, making student academic and personal development the key indicators; the importance of good leadership in creating the right conditions for students to thrive.
Characteristic 5 Alignment of evaluations with school internal review processes and development planning (does not explain how but just adds the internal review works over time)

and five principles:

Principle 1 Judgements and reporting are explanatory (the relationship between 'inputs such as teaching and output of learning')
Principle 2 Evidence is used to feed forward into future system reform as well as to feed back on the current state of the system, with the best practice also using the evidence to 'feed sideways' to provide school-to-school post-inspection support
Principle 3 There is a matching of method and framework content to the local educational ecology and system maturity (some mature systems may not require inspectorates, esp with high teacher effectiveness). Levels of pressure and support need to be judged accordingly too.
Principle 4 There is a sense that inspection is 'done with schools, not to them'
Principle 5 There is rigorous quality assurance and consistency
(p.11).

Pages 19-22 give the matrix – a checklist that describes the literature and how it fits in the five characteristics and five principles.

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.

de Wolf and Janssens' review paper provided an overview of studies into effects and side effects of inspection and publication of pupil performance data. They conclude that there is insufficient evidence to determine if inspection has a positive effect and, although schools believe that performance indicators are important, these are little used by parents or students when choosing schools. They note that several of the studies reviewed referred to unintended consequences of inspection and publication of data, such as such as putting on a show for inspectors. They cite early research in England by Brimblecombe et al. (1996) who found 38% of teachers intended to make changes shortly after an inspection visit, particularly in relation to the way they teach and organise classes. The tendency to change increased the higher up the teacher is within the organisation. Also in England, Learmonth (2000) and Ouston et al. (1997) found that inspection visits led to changes in behaviour among a large majority of school principals. In general, younger, less experienced school principals were more likely to implement changes as a consequence of inspection visits than older, experienced school principals.

Klerk, M. (submitted). The effect of school inspections: a systematic review. Oxford Review of Education.

Klerk's 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 were identified as relevant. Effects of school inspections are discussed in relation to: 1.school improvement; 2. behavioural change of teachers; or 3. student achievement results. Klerk found that there was no evidence that school inspections automatically lead to the improvement of the educational quality with complex interactions between different characteristics of school inspections and the inspector and with individual school factors. She reports on the case studies by Ehren and Visscher (2008) of 10 Dutch primary schools with data collected before and after an inspection. Whether school improvement resulted or not depended on: the quality of feedback about weaknesses; the assessment of weak points as unsatisfactory; the agreement between an inspector and the school regarding improvement activities.

Klerk reviewed research which considered the impact of inspection on poorly performing schools, noting also some unintended consequences. Thus Perryman's (2005, 2006, 2007, 2009, 2010) single case study of a school as it moved out of 'special measures' showed that long term improvement process developed, however with negative attitude to intensive monitoring as a failing school and teachers feeling that they needed to 'perform' against Ofsted criteria.   Willis (2010) (cited in Klerk (submitted) conducted a multiple case study of her own school and two other primaries which had all been placed in 'special measures'. Willis noted that despite a lack of confidence in the Ofsted judgements, it had focused the attention of the schools on student achievement results. Matthews and Sammons (2005) (cited in Klerk (submitted) used changes in Ofsted judgements on the quality of teaching from those originally given to a sample of secondary schools placed into special measures and those given in subsequent inspections two years later. They found substantial improvement in the quality of teaching.

Klerk refers to Chapman (2001) whose case study of five English schools just after an Ofsted inspection found that high quality feedback may be the key to teachers' intentions to change practice. Approximately 20% of teachers studied felt that inspectors' feedback had prompted changes in teaching practice. Klerk also presents evidence of negative effects of inspection in England for example, that of Case et al. (2000) who investigated well- being of teachers both before

and after Ofsted inspections in ten primary schools. Negative effects on well- being and three year disruption to teaching was found.

With regard to student achievement, Klerk cites Luginbuhl et al. (2009) who used a quasi-experimental design to investigate the effect of inspection on the test scores of pupils in primary education in the Netherlands between 1999 and 2003. They found a small positive, or no effect, concluding that inspections do not appear to have negative effects on student achievement but cannot be said to have a positive effect. In England she reviews research from Shaw (2003), Rosenthal (2004) and Matthews and Sammons (2005) noting that each study finds different, slight effects, indicating that the evidence for the impact of inspection on student achievement results is inconclusive.

Organisation for Economic Co-operation and Development. (2013). Synergies for better learning: an international perspective on evaluation and assessment. [Paris]: OECD.

 OECD's (2013) report includes a chapter which draws together previous research findings on external evaluation (inspection) and SSE. They note that research suggests that external school evaluation has differing impact on schools and that certain conditions are associated with schools accepting and acting on feedback from external school evaluation. OECD report that evidence that external evaluation reports are used to drive improvement in school policy and management comes from the Netherlands (Bekkers et al., 2012; Janssens, 2011), New Zealand (Nees, 2006), Korea (Kim et al., 2009) and Sweden, (Ekonomistyrningsverket, 2006). OECD also noted research which suggests that nature of feedback had a greater impact on school improvement than the amount provided (Matthews and Sammons, 2004; Ehren and Visscher, 2008). Follow up measures, including intervention support, when external inspection reveals weaknesses has been found to support school improvement in the Flemish Community of Belgium (Flemish Ministry of Education and Training, 2010), Korea (Kim et al., 2010; Jung et al., 2008) and the Netherlands (Inspectie van het Onderwijs,2010a, 2010b). However, research has shown that there are mixed findings about the impact of incentives from the Flemish Community in Belgium (Flemish Ministry of Education and Training, 2010) and Korea (Jung et al., 2008). There is also evidence that acceptance of findings is necessary to drive improvement from the French Community in Belgium (Blondin and Giot, 2011) but not sufficient (Ehren et al., 2013). Ehren et al.'s findings based on 2200 survey responses from six European nations (Styria in Austria, the Czech Republic, Ireland, England, the Netherlands, Sweden) suggest that 'where external school evaluation sets clear expectations, norms and standards and stakeholders are engaged with and knowledgeable about the external evaluation process, this has significant impact on schools … Expectations set in external school evaluation and stakeholder sensitivity to the results of external school evaluation are also significantly related to schools improving their self-evaluation processes. (Ehren et al., 2013).' (cited in OECD, 2013, p 391). These positive findings about the impact of external review on school self- evaluation were confirmed in research from New Zealand (Schagen and Wylie, 2009), Korea (Jung et al., 2008) and Portugal (Portuguese Ministry of Education and Science, forthcoming). Negative effects of external review prompting compliance rather than improvement are noted (Faubert, 2009) and more recently in England (West, Mattei and Roberts, 2011) who suggest that hierarchical and market accountabilities, through publication of league tables and school competition and the risk of sanctions for failure in inspections are likely to result in compliance to inspection criteria rather than improvement. OECD conclude that more research on impact of different approaches to external school evaluation are needed.

Penzer, G. & Cfbt Education Trust. 2011. School inspections what happens next? [Online]. Reading: CfBT Education Trust.

This examines differences across inspectorates in 17 countries: Czech Republic • Northern Ireland • Denmark • Portugal • England • Scotland • Flanders • Singapore • Germany • Slovakia • Hong Kong • Spain • Ireland • Sweden • Netherlands • Wales • New Zealand. This looks at how each inspection proposes to make improvements. Variations are: some countries publish all reports, some publish some and not others and some publish none. In HK, schools can decided (but cannot reverse their decision). Reports are written to the Govt, the school, or wide audiences; follow up can be enforced or voluntary; tools for improvement can be carrot or stick. Ideological/cultural factors: the importance of information being provided to enable 'parent choice' (e.g. England); in others, it is more about ensuring equality of provision (e.g. Spain); the extent to which the school vs the inspectorate have responsibility (e.g. in Germany it is entirely the school); emotional, ethical and professional aspects all have a part to play in the inspection arrangements, communication etc. In most cases, there is a situation in which the school can 'fail'. This leads to more frequent and sooner follow up, increased support, sometimes leads to tendency on school focusing on 'passing'. Most inspections provide areas for improvement, generally 4 steps are needed to achieve improvement: 1) school staff, governors need to be convinced of the validity of the recommendations; 2) school needs to obtain the resources it needs to implement the improvements; 3) staff need to be motivated and/or sufficiently confident to implement required changes; 4) carrot and stick measures for when school achieve/do not achieve the improvements. Schools are generally disinclined to accept unfavourable inspection conclusions; the credence, communication skills and demeanour of the inspectors has an effect on this aspect. Other ways to help schools accept (and take on) the recommendations include: encouraging self –evaluation and making this part of the inspection; publishing the report to put the pressure on; requiring that criticisms in a report meet a higher evidential standard. Making resources available: this can include funding but also availability of SSE tools (such as HK). In the latter case the school is encouraged to reflect on the findings of the external review and respond. Motivation: the way the inspection is conducted and the way staff perceive it has a direct influence on the response of the school. Rewards/sanctions: increased freedoms (less frequent inspections or more autonomy), good publicity for the school and enhanced professional standing (which may in turn lead to some financial benefit), in particular for headteachers. Singapore has awards for successful schools. Coasting schools that become good or good ones that become better are not usually rewarded. Overall conclusions: accountability is in tension with improvement; more is needed to translate inspection outcomes into improvements; contributing thematic, overall reports to the evidence base is a useful aspect; taking lessons back to schools (e.g. by inspectors who are serving heads (England)) is a useful aspect; clear, explicit statements are needed on what is wrong and what needs to be done; decisions about which schools to inspect should be 'strategic', i.e. choose according to the extent to which 'lessons can be learned' from particular cases.

References:

Chapman, C. 2001. Unlocking the potential: inspection as a mechanism for school improvement. Improving Schools, 4, 41-50.

Churches, R. and C. McBride (2013). "Making external school review effective."

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.

Klerk, M. (submitted). The effect of school inspections: a systematic review. Oxford Review of Education.

Organisation for Economic Co-operation and Development. (2013). Synergies for better learning: an international perspective on evaluation and assessment. [Paris]: OECD.

Penzer, G. & Cfbt Education Trust. 2011. School inspections what happens next? [Online]. Reading: CfBT Education Trust.