

# A Better System of Inspection?

A Report by  
**Chris Boothroyd, Carol Fitz-Gibbon,  
John McNicholas,  
Meryl Thompson, Elliott Stern  
and  
Ted Wragg**

edited by  
**Michael Duffy**

illustrated by  
**Bill Stott**

# **A Better System of Inspection?**

---

A Report by  
Chris Boothroyd, Carol Fitz-Gibbon,  
John McNicholas, Elliott Stern,  
Meryl Thompson & Ted Wragg

edited by  
Michael Duffy

illustrated by  
Bill Stott

published by  
**The Office for Standards in Inspection**  
c/o 9 Quatre Bras, Hexham,  
Northumberland NE46 3JY

OFSTED OBSERVED

AN OFSTED COUPLE IN BED

HOW WAS IT  
FOR YOU DARLING?

SATISFACTORY, AND  
IN PARTS, VERY GOOD.



# CONTENTS

---

Foreword		iv
Chapter 1	The Oral Evidence Meryl Thompson	1
Chapter 2	The Written Evidence Chris Boothroyd & John McNicholas	7
Chapter 3	Ofsted's Methodology Carol Fitz-Gibbon	16
Chapter 4	Inspection and School Self-Evaluation Ted Wragg	21
Chapter 5	The Evidence Assessed Elliott Stern	26
Afterword		31



# FOREWORD

---

In June 1996 more than a hundred people attended a conference at New College, Oxford to share experiences and ideas about Ofsted inspections of schools. The conference focused on four issues: the soundness of Ofsted's methods; its impact on schools, teachers and the education service; the extent to which it represented value for money; and the way forward for inspection as a means of accountability and a tool for school improvement.

A brief report of that conference, entitled *Improving School Inspection*, was published in November 1996. It contained a summary of the key issues raised in debate, and of the recommendations that emerged.

This second publication puts flesh on that skeleton. It begins with an impression of the 'oral evidence': the contributions of teachers and headteachers with experience of inspection who responded to the invitation to give brief accounts of how it seemed to them.

Many people who could not travel to Oxford to contribute in person sent written submissions for consideration. They cover many themes; but the dominant concern, documented in Chapter 2, was the lack of conviction carried by inspectors' judgements.

To supplement the direct experience of those who spoke and wrote, the conference heard two academic presentations: one from Professor Carol Fitz-Gibbon on methodology, and one from Professor Ted Wragg analysing weaknesses in the current system of inspection and proposing better alternatives. These talks are the basis of Chapters 3 and 4.

Concerned that the debate about inspection has been too narrowly conducted, essentially within the world of education, the conference organisers invited a panel of people representing a wider constituency to assess the evidence and arguments. Chapter 5 is a summary of the responses of members of this panel to what they heard.

Central to their conclusions was the suggestion that it was time to expect government to establish an independent review of the effectiveness and efficiency of the Ofsted inspection system, supplemented by appropriate research.

*The Conference was organised by, and this Report is published by, Ofstin ('The Office for Standards in Inspection') – a voluntary, informal and independent group concerned to promote debate on the role and future of inspection.*

# CHAPTER ONE

---

## The Oral Evidence

Meryl Thompson, Head of Policy Unit, Association of Teachers and Lecturers

‘Anyone can become angry,’ Aristotle wrote. ‘That is easy. But to be angry with the right person, to the right degree, at the right time, for the right purpose, and in the right way – this is not easy.’ Those giving evidence to the conference had not chosen the easy way. They had marshalled their evidence; acknowledged the quality of the *Framework for Inspection* and its power to concentrate attention on effective teaching; and accepted the need for open, professional accountability. They generously recognised the professionalism of many inspectors and appreciated the positive effect of the preparation and planning, the talking and team-work, and the imposed discipline of improving paperwork. They readily described the hard work and loyalty of their colleagues and the support of parents. They spoke of their determination to go forward positively. But they had not decided not to be angry. The evidence we heard in Oxford gave an indication why.

As evidence it is open to criticism – as indeed is this impressionistic, and not quantitative, summary of it. Over six hours thirty-five contributors – governors, heads, teachers and parents – were questioned by two panels. I attended one panel as a ‘collator’, briefed to take an overview and to report to the assessors. Professor John Gray, of Homerton College, attended the other in the same role and provided his report for me. The readers of this account are being asked, therefore, to trust my perception and my judgement. I have to establish credibility just as do inspectors. You are being asked to trust the integrity of governors, teachers, headteachers and others in collecting data, in assessing the impact of inspection on their colleagues and on educational processes, and in ‘telling it like it was’. For those inclined to doubt, I suggest comparison of this oral evidence with the findings of other research into the experience of Ofsted. My impression is that the main doubts and criticisms raised here are consistent with it.

### 1.1. OFSTED’S METHODS

In general, inspection conclusions were regarded as neither unreasonable nor unexpected. Overall, inspectors were seen as fair and professional, although they were likely to be seen less favourably by teachers than by headteachers. Most frequently, the evidence for methodological inadequacies related to doubts about the sufficiency and typicality of evidence. For example, is a sample of 150 hours of observation, out of a possible 1,600 teaching hours available to the inspectors, *reliable*? When teachers believed that students behaved better than normal and students, from the evidence we were given on pupils’ perceptions of inspections, thought that teachers taught them differently, did inspectors see a ‘normal’ school? John Gray caught the underlying concern when he asked whether alternative ways of collecting and analysing inspection evidence might not lead to different judgements and different agendas. Indeed, should we, he queried, in an age of sophisticated methodologies for school evaluation, be using this primitive methodology of the ‘snapshot’ to judge schools?



Most of the criticism we heard was about the reporting process and particularly the written reports. But the process of reporting is not only a crucial part of the methodology of inspection, it is also the vital link with other spurs to improvement, the mechanism of accountability to parents and the foundation of the action plan. If the report can be called into question so can the efficacy of the entire process. The evidence was that report writing was a particularly flawed aspect of inspection and was itself likely to raise or crystallise doubts about the quality of the inspection evidence and the objectivity of its judgements. It was not unusual for inspection to be considered fair and feedback useful, or at least satisfactory – until the report.

Reports comparing a school's performance against 'national expectations' led schools to question how inspectors can make valid comparisons and, if they can, what is it that is 'broadly in line' with them. Reports referring to a 'high' proportion of lessons falling into a given category – as, for example, 'always sound or better' – often carried little or no meaning to teachers. They were seen as 'woolly and unhelpful'. One school, which exceptionally found that the inspection process was enjoyable and positive and that its oral feedback was good, described its report as 'bland'. It was so bland, we heard, that the Registered Inspector even apologised for it. If reports are so opaque that they have to be decoded and carry little or no meaning to teachers and governors they are unlikely to convey an accurate message to parents or contribute meaningfully to school improvement. Poor reporting reduces the impact of the inspection process.

Some of the evidence offered as dealing with Ofsted methodology was more concerned with the inspection process and its impact upon the school. We heard, for example, about staff becoming 'saturated by visits' and so experiencing a 'climate of intimidation'. If a distanced, cold objectivity on the inspectors' part is integral to the methodology it certainly has its impact on teachers. One set of inspectors was described as 'frosty' and 'clinical', so much so that they walked out without saying goodbye. Not surprisingly, the subsequent report was perceived as reflecting that negativity. However, other teams were seen as friendly and supportive. Evidence of similar contrasts, which affected the inspection process, was given in examples of inspection teams who were or were not prepared to talk to teachers before the inspection or to provide departmental and individual feedback.

Some evidence questioned the limitations of Ofsted's methodology. It was argued, for instance, that an assumption that ethnic diversity is invisible was built into Ofsted's method. Attention was drawn to the undisputed male bias in inspection teams. One headteacher felt that the inspectors had made an instant reaction to the school: 'They didn't like the look of the kids.' Others gave evidence of inspectors making disturbing statements and expressing their personal opinions. Can we be sure that inspection does not have a cultural or sub-cultural bias and that the process is free of prejudice and stereotyping? Have we allowed Ofsted to impose pre-determined orthodoxies upon the profession?

We heard emphatically that schools lost confidence in the inspection process when, for example, inspectors showed factual ignorance, as of what were

and were not statutory duties; when they made unnecessary demands for duplicate information; when the originally-notified inspection team went through so many changes that it was impossible to believe that it was a team or when the Registered Inspector did not know the team; when the report went into several drafts or was so ill-written and inaccurate that it went into three negotiated drafts taking 50-60 hours of school time. Lack of confidence was particularly acute in primary schools, who often felt that they were assessed against secondary criteria, in special schools, and in relation to special educational needs (SEN), when inspection teams had not undertaken the enhanced training for SEN, did not have knowledge of special units in mainstream education, and had no recent direct experience of SEN. The accumulation of this evidence led John Gray to question whether enough attention has been given to the personal qualities of inspectors as revealed in their own professionalism, their knowledge of how to apply the criteria and their understanding of the most effective ways of securing institutional change and professional development.

Implicitly some evidence raised questions not of the method's objectivity but of its fairness. If a school had had two headteachers and three acting headteachers in the last six years, wasn't that a factor in its performance? Or if supply staff had replaced conscientious senior teachers absent with stress-related illness? Were factors such as budget deficits, social conditions, deprivation, environmental depression and school fires irrelevant? Why was value-added data ignored? If the *Framework* is changed, how can reporting really be consistent over time? Particularly, teachers in schools in difficult conditions, often in inner cities, felt that Ofsted was anything but a level playing field. They and others resented, too, the absence of a formal complaints procedure. Was it right that where there was evidence of poor and inaccurate report writing Ofsted could not act to correct this publicly and the school had no public redress?

## 1.2. THE IMPACT OF INSPECTION

Throughout the sessions we were presented with unarguable evidence that both the inspection methodology and the process of inspection has had a major impact on teachers, both personally and professionally. We heard repeated references to the stress caused to teachers, headteachers and governors and, one must surely infer, to children. There was evidence that the overwhelming majority of teachers experienced anxiety *before* the inspection and that frequently this was compounded by the stress of an increased workload. Long lead-in times meant that inspection dominated lives for a protracted period. In evidence collected from four schools, all with good inspection reports, anxiety which affected teachers' marriages, health and personal lives and even led to cancelled holidays was cited as the worst part of inspection.

Teachers said that they couldn't teach as well as usual because of the anxiety and stress. Half the pupils surveyed in a research project reported to us felt that some of their teachers had been nervous and all of them were aware of the stress on teachers, including the impact on their mood and temper. Half the pupils reported that inspectors created an artificial atmosphere. In one school the chair of governors had resigned with stress two weeks before the



inspection and others reported incidences of nervous breakdowns, both of teachers and headteachers. Most schools reported 'post-Ofsted malaise' – increased absenteeism and tired and exhausted teachers. One headteacher considered the experience as 'inhumane'. It took months, she said, to restore morale. On the positive side, we learnt that Ofsted had made teachers pull together and had often strengthened management teams. A minority of teachers did find inspection a positive experience.

The stated intention of Ofsted is to bring improvement through inspection. However, the evidence we heard suggested that, even where the actual inspection process was seen as fair, a lack of confidence in any part of the process or team was likely to lead to cynicism. The effect was to undermine senior management in getting a positive response to the action plan. The evidence indicated that very frequently inspection had impaired the process of self-evaluation and improvement. It was often the opinion that 'Ofsted did nothing for me or my school'. Such negative reactions were offered as evidence that there is a fundamental fallacy in the assumption underlying Ofsted, which is that if you are criticised you will improve. When inspection is seen as a 'witch-hunt', we were told, it will not achieve even the minimum of improvement.

In fact inspection could lead not only to a state of sheer fatigue and post-Ofsted malaise, but also to a very real loss of purpose, even to a loss of identity and to anger and low morale. One headteacher, following a good Ofsted report, found her view of herself as a thinking, forward-looking person undermined by 'the feeling Ofsted creates that you are running around, closed in and dominated by the process – in the middle of it all but not there'. There was very little evidence of an experience that was invigorating, challenging or supportive in the drive for improvement.

Above all, we had evidence of the impact on schools deemed failing and in need of special measures. 'Nothing prepares you for the feeling of physical and mental assault, the distress and the effect on your quality of life,' one head said. This was from a person who had joined the school after a period of great discontinuity, knew there were serious weaknesses and had tried to prepare the staff for them without undermining their confidence. In this case teachers were in tears after the first day – the negative feedback had been so distressing. Their imagery was of 'vultures circling round the school'. Although parents were supportive, staff were 'caned' on radio and television. There was no support during the critical action-planning period – and then the DfEE tore the action plan apart. Inspection had made improvement much more difficult. The school was 'damaged by a gravely flawed process'.

Little work has been done on the impact of inspection on the learner. We were fortunate to have evidence of this presented to us. It shows that secondary students are very aware of the preparation and arrangements for inspection. They feel ignored and perceive inspection as an interruption to their work, especially in examinations. During inspection they are generally sympathetic to their teachers and good students tend to try harder. They are aware, too, that during inspection teachers act differently. Staff were more punctual, explained things better and were more helpful and

encouraging. Lessons were more planned, quieter and more work was done. For pupils, in other words, the period of inspection is not typical of their experience of school life. So there are clearly some positive effects. Are they worth the cost?

### **1.3. IS INSPECTION VALUE FOR MONEY?**

Not surprisingly, teachers' and governors' observations and doubts on methodology and process and their experiences of the impact of Ofsted influenced the evidence they offered of whether it was value for money. Sometimes it has to be inferred. For example, in the small survey of four schools all the respondents said that their Ofsted report told them nothing they did not know. Their view was that Ofsted was spending some £27,000 on each of them for nothing. Other evidence indicated that the Ofsted report could give managers an agenda for action, for example, by publicly identifying a failing department, but questioned whether this was an appropriate outcome for such a cost. Some governors said that inspection had motivated governors and increased their participation; some said that they now regarded their experience of inspection as something to be erased. The evidence from school managers was generally that they had to put inspection behind them in order to proceed positively. In fact it had retarded, rather than opened up, opportunities, and, measured by the way it absorbed time, it had been a financial burden. The issue of inspection as value for money was most often presented in the context of whether inspection was a good model for improvement. Put simply, evidence was offered that it was not. No evidence was offered that it was.

### **1.4. THE WAY FORWARD**

It was not generally proposed that inspection should be abandoned, and certainly not that improvement was not required. The way forward was seen as finding ways for inspection to strengthen its contribution to school development by building in strategies for self-review and self-improvement. Governors, for example, wanted better links between the inspection report and its action points because it is difficult to get agreement on what these links are. Many proposals were about what was described as the 'anti-psychological' elements of Ofsted. These included proposals for more constructive criticism, more feedback, more value-added data, greater professional involvement, a more balanced emphasis on a school's positive features and, by inference, for a better trained inspectorate of higher quality people. It was perhaps implicit that the way forward, too, included rectifying the flaws of the present system, especially report writing, as well as incorporating changes such as involving students more and placing a greater emphasis on the performance of the school in context and over time, rather than only classroom observation, when evaluating the school's outcomes. Dissenting voices claimed that advice and support were more important for school improvement and said that higher quality management training was a prerequisite. One headteacher suggested that if the purpose of inspection was to create a two-tier system of schools, inspection was clearly working.

## 1.5. CONCLUSION

I share with John Gray a conviction, on this evidence, that inspection causes too much stress. This stress is dysfunctional, since it appears to have a disproportionate effect in demoralising and demotivating good teachers, and sometimes causes them to leave a profession they would otherwise have stayed in. I share with him, too, the conviction that the way forward for school improvement is for inspection to be linked with school self-review, and to be informed by the data we need to make realistic, accurate and contextualised judgements. Particularly, we need better data and understanding on the performance of inner city schools.

However, this summary would not be complete without a comment on the nature of the intensive oral evidence. It was serious, focused, factual, unhistrionic. There were people in the room who had to take both the odium and the responsibility for 'failing schools', yet to which they had only recently been appointed. There were people who were disoriented and disillusioned by the experience of inspection but who carried on professionally to support and sustain others, to pick up the pieces, and to counteract cynicism and restore faith in other means to improvement. There were many who raised pointed questions about consistency, veracity, the use of evidence, the reliability of the process and, above all, its effectiveness. There were even more who had witnessed the impact of anxiety and stress on colleagues and yet sought to balance this against the longer-term impact of openness and accountability. Pre-eminently, this was evidence of the profound impact of Ofsted and inspection on the teachers, and on teachers' self-esteem.

It is difficult to know which effects of such an innovation were intended and which unintended, but we have to assume that much of the impact of inspection had not been fully predicted. Retrospectively, this can only be seen as a culpable lack of imagination and empathy or a careless disregard for the literature on the implementation of change. The evidence we heard told a great deal about the will, character and self-restraint of the teaching profession. It is time now for the educational improvement strategists to devise a model of inspection for school improvement that builds on and does not undermine these conditions.

# CHAPTER TWO

---

## The Written Evidence

Chris Boothroyd & John McNicholas

People throughout the land took up Ofsted's offer to hear of their inspection experiences. From single sheets to dossiers of correspondence with Ofsted, nearly a hundred stories and observations came: from teachers, headteachers, governors, professional bodies, parents, inspectors and researchers, representing all sectors of schooling and all types of Ofsted verdict. They are coming still.

Each contribution is a particular viewpoint with a particular tone – cool analysis, passionate outrage, wry bewilderment, constructive advice. Through all of them come the voices of hard-working professionals, wrestling with serious issues in challenging circumstances. Looking for constructive criticism and practical help wherever it may be, they judge Ofsted inspections in that light. They are generally disappointed. In some cases they and their schools are seriously damaged.

Taken together, submissions deal with the judgements reached by inspectors, with their impact, and with possible improvements. The first issue stands out as troubling to most of those who wrote to us. Because the other factors receive attention elsewhere in this report, this chapter reflects particularly our correspondents' thoughts about inspectors' judgements. Quotations are from teachers, including senior staff, except where indicated.

The grand edifice of inspection, casting its huge shadow over the educational landscape, is founded on one simple assumption: that the judgements reached by thousands of inspectors across the country are reliable. They must be soundly-based, consistent, and capable of being used constructively by those whom they affect. If any one of those conditions does not apply, the exercise is futile.

From the accounts we received, eight basic qualities emerge as essential if inspection judgements are to inspire confidence. Ofsted appears to agree: for each requirement, we quote endorsement from its own guidelines.

### 2.1. 'FITNESS FOR PURPOSE'

'Inspectors must be appropriately qualified for the particular tasks they will carry out. Teams must include expertise in inspecting the subjects of the curriculum and all other areas covered by the inspection.'

*Making the Most of Inspection, p.10*

For many of our correspondents the Registered Inspector, or the composition of the inspection team, or both, proved unsuited to the school. The most common weakness was a mismatch of expertise and experience with the sector or subject being inspected. Poor knowledge of the National Curriculum, of the SEN Code of Practice, even of the *Framework for Inspection*, were regularly cited, as was lack of senior management experience. LEA teams inspecting in their own patch were often not seen as objective. High among the reasons for doubting the credibility of the whole

team were instances of what can only be described as unprofessional behaviour by individual inspectors.

Worthwhile judgements require a suitable team. Attention to those judgements demands that inspectors conduct themselves professionally.



*'How can I accept the criticism of an inspector who on her own admission had never taught primary age children, and of a team who criticised us for directed teaching, when our SATs results are in the top 25 per cent, our parents very happy with our methods, and Chris Woodhead himself is now advocating methods such as we employ?'*

*'The RegI had no school experience for twelve years. Prior to that he was a non-teaching head for many years, therefore his classroom experience was certainly not up to the standards required of the teachers in this school.'*

*'The lead inspector did not present as someone who would make valuable judgements. He slouched about the school, lying on tables during observation sessions. He left the building many times every day to smoke but generally did this in view of the children.'*

*'I was asked by an inspector why I was teaching young children when I had a Master's degree.'*

*'I was told, "If a pupil does much worse in one subject than in another, that's under-achievement".'* (Governor)

*'On the [Ofsted] course I followed, one participant had only taught at two boys' public schools. This person has since followed the primary conversion course and recently inspected PE in a primary school. I commented that he didn't know anything about PE and he observed that he could tell whether they were enjoying themselves.'* (Inspector)

*'We are still unsure of the ethics which allow a registered inspector to employ his wife as his admin assistant and thereby have her along at all his meetings and during inspection.'*

*'Prior to the inspection I raised the issue that there was nobody with primary teaching experience on the team, and a lack of any musical expertise when we pride ourselves on our music. No reply was received to my letter.'*

*'... especially since this same man – wearing his other hat – would be quite happy for us to have him back to put right all the things that he was highlighting as lacking in the school.'*

*'[The RegI] complained because some Year 2 children wouldn't leave him alone and he couldn't cope with them. A small group of children were trying to talk to him. He tried to send them away without success. He then told the teacher he would have to leave.'*

*'"I have no opinions, only criteria": these were the words of the RegI. Clearly this was not true. He had very strong opinions.'*



‘Our inspection team did not know each other, and there were obvious tensions within it. There was a public row between two inspectors in the hall. This did not inspire respect for judgements.’ (Governor)

*‘[The RegI] totally forgot three appointments and kept staff waiting.’*

‘I found [the RegI] pompous, single-minded, high-handed, insensitive, a poor listener, a man with his own agenda.’

## 2.2. INSPECTORS ARE THOROUGHLY PREPARED

‘Inspectors...will read any other documentation the school chooses to provide.’

*Making the Most of Inspection, p.13*

Schools spend much time, energy and money on preparing advance material. Distressing to our correspondents was the frequent evidence that inspectors had not spent enough time reading the information provided at such cost, often failed to assimilate or understand it, and did not always coordinate their ways of picking up issues arising from it. Lack of grip on the background to the school, sometimes even perverse readings of its context, were regularly mentioned as contributing to dubious judgements.

*‘A lengthy document which outlined the system for monitoring initiatives was ignored by the inspector, and he asked the school to cancel a meeting of the group which handles monitoring so that departmental feedback could be delivered. The report criticised management for having no monitoring procedures.’*

‘An inspector criticised that too much of the total resources was spent on staffing, yet local authority figures for comparable schools show spending is at the lower end of staffing budgets.’

*‘The Inspectors made assumptions about our children that weren’t true. They said that a “significant number” of children came into school knowing their alphabet. The inspector could not find his evidence when I challenged what was meant by “a significant number”.’*

‘Teachers’ invitations to refer to lesson plans were ignored.’

*‘The PICSI’ referred to 15 unqualified teachers. We have no unqualified teachers.’*

## 2.3. PARENT INFORMATION IS CORRECTLY INTERPRETED

‘Points raised by parents will be reflected in the report if supporting evidence is found.’

*Making the Most of Inspection, p.15*

A relatively minor theme overall, this was a vital issue for those who did identify it. These were cases where the inspectors’ judgements appeared to rely unduly on inferences drawn from the parents’ questionnaires and meeting. Doubts about the validity of such evidence rest on the small and self-selecting nature of the sample, on the naivety with which information was sometimes interpreted, on the lack of rigour in testing this evidence, and on accounts of Registered Inspectors ‘leading’ the meeting, particularly in a negative direction.

---

<sup>1</sup> The Pre Inspection Context and School Information data supplied by Ofsted to the Inspectors.

*'I was extremely concerned about the role assigned to parents in the inspection. The questionnaire was of an embarrassingly low quality, paying little regard to validity or reliability. I assumed that it was trying to goad parents into attacking the staff.'* (Governor)

'Parents complained that the meeting was being led in a way they did not like, that they were being discouraged from positive comment and encouraged to be critical, eg: "there must be something you don't like?">'

*'The credibility gap was there for all to see. How are the views of 60-odd self-selecting parents to be fed into the report, and how much weight will be given to them? Were boxes ticked, marks out of 5 awarded, or short-hand notes taken? [The RegI] was at some pains to reassure us that generalised murmurs and other indicators of support were being taken on board, though he did not specify how. If the inspectors' account of the parents' meeting was to be fair, it would have to be impressionistic, yet Ofsted claims credibility through its supposed objectivity and its scientific methodology.'* (Parent)

'The Ofsted attempt to manipulate parents is not effective but is unpleasantly negative.' (Governor)

#### 2.4. INSPECTORS SEE THE SCHOOL AS IT IS

'Registered inspectors must ensure that their judgements are valid, in that they accurately reflect what is actually achieved and provided by the school.'

*Framework for the Inspection of Schools, p.7*

Inspectors' judgements determine the recommendations for action, and thus affect the daily life of the school for several years. Those judgements must therefore be based on evidence that reflects the school as it normally is: a particularly important requirement when classroom observation is so significant. Most of our correspondents suggest that the inspection week is far from typical either for pupils or staff. While all sorts of minor influences are at work, people identify excessive anxiety and stress as the main distorting factor. The long period of anticipation, Ofsted's perceived high and judgemental profile, public and press expectation, and the grading of teachers are all mentioned as contributing to this.

*'The school was perceived by employees other than teaching staff as having a surreal quality during the days of inspection. We are genuinely concerned that such stress and the resultant unusual atmosphere serves only to undermine true teacher performance.'*

'We were disappointed that the inspectors made very little personal contact with the children. This left them without one of the essential elements required to judge the success of the teaching.' (Governor)

*'Primary teachers cannot reasonably be expected to be equally good at delivering every aspect of the National Curriculum. People inspected teaching what they know is their weakest subject are particularly aggrieved.'* (Professional Association)

'Inspectors seemed to expect deferential behaviour – not the best way of seeing what typically happens in a school.' (Governor)

*'The pressure of limited time can result in unreasonable emotional demands on staff and in more than half the inspections staff were reduced to tears. The intense nature of inspections and the need to gather information does mean that demanding responses sometimes undermines the confidence of staff. The 'interrogation' of a member of staff who may have several responsibilities by several inspectors in a short time is sometimes unavoidable in small schools. The opportunities to corroborate in such schools are extremely limited and judgements can be made on evidence that is so limited as to be at best questionable, and judgements in these circumstances lack credibility.'* (Inspector)

'25 per cent of observations at Key Stage 2 were of a teacher who was under disciplinary proceedings for competency.'

## 2.5. JUDGEMENTS ARE BASED ON VALID EVIDENCE

'Registered inspectors must ensure that their judgements are secure, in that they are rooted in a substantial evidence base and informed by specified quantitative indicators.'

*Framework for the Inspection of Schools, p.7*

Again and again, accounts suggested that inspectors' claims to make valid and measured judgements are spurious. In reality, it seems, samples of evidence were too small, too inconsistent and uneven and too little tested for the weight of judgement they had to bear. Inspectors were often casual in pursuit of facts and sometimes lacked objectivity.

Particular exception was taken to the often fragmentary observation of lessons and to its uneven distribution across teachers and classes – criticism made more strongly by those inspected under the new framework with its grading of teachers. Secondary schools with value-added data were often frustrated by some inspectors' apparent lack of knowledge or interest: where the information was a crucial counter-balance to 'raw' data, this was felt to be irresponsible. Many people noted that inspectors seemed reluctant to talk with pupils about their work, and commented on a poor level of communication, during inspection, between team and teachers.

Some sensed a pre-determined agenda, a set of previously-agreed expectations and issues, which unduly influenced interpretation and judgement. Others were surprised by how quickly, and how early in the process, judgements appeared to form, and how frequently these were influenced by prior subjective expectations. Many felt that inspectors were more interested in finding weaknesses than in recognising achievements, and that the school's context was largely ignored.

Two features in particular damage people's confidence in the process. Significant factual errors sometimes persisted into the final report. And a frequent observation was of substantial differences in judgement between the inspection team and the LEA (Ofsted-trained) advisers when working, in theory, to the same *Framework*.

*'It was established at the end of Day One that more than 90 per cent of pupils were making some progress. When pressed for evidence of this we were assured the correct procedures in the form of criteria from the Framework had been applied – an answer which became frequently used during inspection week.'*

'The collective judgement was positive to the school but individual judgements varied in consistency in that some were analytical, others were simply descriptive.'

*'There are unacceptable inconsistencies both between the individual members of inspectorial teams and, more seriously, between one inspection team and another. The mechanisms for ensuring consistency are not yet rigorous enough. The data are certainly not robust enough to bear the monumental weight of opinion currently laid on them [in HMCI's] public statements about the performance of the education system as a whole.'*  
(Professional Association)

'Rigour in the process is not yet matched by rigour in providing evidence.'

*'Positives were consistently under-stated; negatives consistently over-stated.'*

'Fifteen-minute lesson observation forms (lofs) which skew pilot (a computer program used by Ofsted and HMCI) judgements further undermine the reliability of the evidence base.' (Inspector)

*'The reinspection was carried out by two HMIs. A key factor in the success of the revisit only six months after the action plan was that the HMI were primary specialists. The findings of 50 per cent of teaching satisfactory or better in the final report improved to 84 per cent of teaching satisfactory or better in the report at the re-visit.'*

## 2.6. JUDGEMENTS ARE TESTED FOR ACCURACY

'It is good practice to test hypotheses with staff before judgements are finalised.'

*Framework for the Inspection of Schools, p.7*

Concern about superficiality and subjectivity was increased by a perception that inspectors frequently did not seek or accept explanations for matters on which they were reaching negative judgements. Gullible acceptance of vocal parent minorities was identified in several cases as an example of this failure. But it applied equally to situations where the context of lessons or the reasons for policies and procedures were not tested against evidence from other relevant sources. Similar lack of rigour is evident at the end of the process, with many people disturbed by the difficulty in challenging findings and by the lack of an appeal procedure.

*'The reliability of judgements was low. At no stage did any checking of facts take place. The final judgement was given to the senior management team by the lead inspector only after the inspection team had left. Factual errors which then emerged could not be challenged as the inspectors who made them were no longer available.'*

'When we received the final report we faxed an immediate response pointing out forty-eight serious factual errors. We were told it was too late to do anything.'

*'Judgements presented a picture of an unknown school, praising under-performing departments and slating good ones.'*

'Some inspectors are expected to compile their evidence on two or three aspects within two days. In small schools (two inspectors over four days)

an individual inspector is expected to report on a minimum of fourteen subjects and aspects. First drafts of Reports are produced within twenty-four hours. My first AI-led inspection finished on Thursday at 3.00 pm. I submitted my paragraphs for the first draft for Quality of learning, Quality of teaching, Mathematics, Geography, PE, Assessment, recording and reporting, Quality and range of the curriculum, Equality of opportunity and Resources for learning at 3.45 pm the next day and was handed the first draft of the Report at 5.00 pm of the same day. Where is the rigour or fairness in this?' (Inspector)

*'Two quotations from our report: "Systems are in place to monitor absence. Pastoral staff work hard to improve attendance, targeting many of the pupils, involving their parents and have achieved good results. Lesson registration and absence slips for monitoring truancy are efficiently completed and are having a positive effect in reducing unauthorised absence." With which compare: "The general organisation is of a high standard with secure and reliable systems in place for all aspects except attendance."!*

'A nursery nurse was sitting with the group during registration so that she could follow up work presented by the teacher. The RegI took this up with me and although it was pointed out that she does not, as a rule, sit with the group until registration is complete, we were penalised and our report said that we were "not quite giving value for money" because we were not utilising our "resources" efficiently. I pointed out that this nursery nurse arrives in school at 8.00 am and does not leave till 6.00 pm (she helps to run an after school club which generates income which has helped us to pay off our overspend). She also accompanies our Year 3 classes on a 5 day school journey, every year, and does not receive any additional payment for being on call for 24 hours a day.'

*'The financial advice of the RegI conflicted with that of the LEA auditors.'*

## 2.7. JUDGEMENTS ARE CLEARLY EXPRESSED

'The summary report must be written in a clear and comprehensible style.'

*Framework for the Inspection of Schools, p.13*

Many of those who wrote to us were highly critical of the language in reports and, to a lesser extent, of the style of oral feedback. Reports were criticised as stilted, often badly-written, sometimes internally inconsistent, and – in a few minds – meaningless. Reports were commonly described as unhelpful, in their structure and expression, to parents and governors.

Conclusions quite often appeared to change at various stages between initial oral feedback and draft report, between draft report and feedback to governors, between that and final report. The reasons for these changes were often obscure, while the malleable nature of judgements did not increase confidence in them.

*'The report language was simplistic and infantile.'*

'All the [parent] comments received by the school were critical of the Report. They found it confusing, contradictory and badly presented.'





'The quality of learning is sound in 3/4 of lessons...The quality of teaching is sound in 4/5 of lessons.' Is creating a learning milieu independent of teaching?' (Governor)

'The report that we received was badly written. Effective communication requires skill in interpretation of evidence. Our report was bland, repetitive to a point of incoherence and demoralising to read for our whole team.'

'The final report we eventually received was for another school in a different part of the country.'

'The final report was weakly constructed, almost the result of cloze procedure.'

'Distinctions are drawn between differences which are indeterminate or speciously precise.' (Governor)

'It is disappointing that a report which more than once mentions low standards of literacy is written in poor English, with little regard for spelling, punctuation, grammar, and the conventions of paragraphing.'

## 2.8. JUDGEMENTS CAN BE ACTED ON EFFECTIVELY

'The key issues for action...should provide a clear and practicable basis on which the appropriate authority and the school can act.'

*Framework for the Inspection of Schools, p.13*

Just 3 per cent of our correspondents considered their inspection judgements fair and helpful to development. 40 per cent considered them fair but marginally useful, in that the school was already working on the issues identified. 57 per cent regarded the overall judgement as unsound and in many cases damaging to development. The majority view was that at best there was a tenuous link between the report and school improvement, while a smaller group felt that the required Action Plan was mostly unrelated to the school's previously agreed priorities.

Many identified a long-term post-inspection malaise. Frequently mentioned were: collapse of energy, confidence and morale; loss of the impulse to develop; a new disaffection and division among the staff; and a sense of being deskilled by the process. Sadly, stress-related illness recurred as a theme of these comments.

*'The Ofsted system has, in this case, constructed a situation where the end result is possible school closure, destruction of professional lives and the negation of five years' school improvement work. Can this really be "Improvement through Inspection"??'*

'This has been a soul-destroying experience. We are hard-working, caring, professional people who no longer trust our own judgements and question our own competence.'

*'After inspection, panic writing and re-writing of development plans threw all medium and long term planning into chaos and confusion.'*

‘Confident, reflective practice was dealt a body blow.’

*‘The failure to identify fundamental and pervasive weaknesses has left the staff disillusioned.’*

‘Our recent experience of the inspection process has been damaging – to our morale, to our health, to our attitude to the current educational climate. We suffered a model of assessment which was, in our opinions, unfair. It was critical, not supportive or developmental.’

*‘It would be extremely difficult to be sure that any improvement in students’ learning resulted from the Inspection. Some of the developments that have subsequently taken place could have been implemented earlier if we had not had to break off our own programme to ensure that we were well placed for the Inspection.’*

‘We learnt nothing that we did not know and felt that our achievements were ignored. Years of encouragement to analyse and question what we do, which is based on confidence, were negated.’

*‘Inspection seems to have induced changes in policy statements, curriculum documentation and administrative procedures and practice.’ (Researcher)*

‘The post-Ofsted collapse of energy, a cumulative and collective experience, is likely to burn out much of the best teaching that could have been expected for the next half term.’

Our last words come from the staff of a school who, observing ‘perhaps we should say here that the end result was, in Ofsted speak, a very good report,’ went on to say: *‘What follows are the words, thoughts and feelings of a group of people who have been through the Ofsted experience three months ago and still feel incensed by it. The feeling of impotency, rejection, depression and demoralisation it has left behind has no precedent in any other experience we have had.’*

# CHAPTER THREE

---

## Ofsted's Methodology

Professor Carol Fitz-Gibbon, School of Education, University of Durham

### 3.1. THE ISSUE OF RELIABILITY AND VALIDITY

There is nothing inherently unscientific in using human judgements as measurements. In fact such procedures may often be essential. However, there are certain agreed and fundamental standards to be observed in the collection and analysis of such judgements. If these are not observed there can be little confidence in the evidence presented.

Ofsted's methods have failed to meet minimally acceptable standards. Essential studies that would be expected from a properly designed inspection system have not, apparently, been conducted. Further, it appears from a survey of 158 headteachers<sup>2</sup> that Ofsted's inadequacies are widely recognised. (Survey results are reported below in indented paragraphs.)

It might be argued that an inspection need not meet research standards. On the contrary: because of its potential to mislead and distress parents, pupils and teachers and because of the apparent faith placed in Ofsted by politicians, the public is entitled to expect the highest standards from the Office for Standards in Education. The fact that Ofsted has been allowed to operate without adequate validation is indicative of a serious weakness in UK education – the failure to understand and make use of this century's developments in the methodology of investigation. In this sense, certainly, Ofsted has produced evidence that we do have 'failing' schools in the UK. They are those that educate our politicians.

The methodological concepts considered below are those of sampling, consistency (reliability) and validity. The argument is *a priori* but the views of the headteachers quoted are of interest as it is they who are supposed to benefit from inspection.

#### 3.1.1 How Representative is the Sample?

It seems unlikely that the pre-announced visits by Ofsted inspectors provide them with a view of the school as it is normally functioning. Given the statement: '*An Ofsted inspector sees a school as it is normally*' – 81 per cent of headteachers disagreed.

Even if the lessons observed were representative, we would still need to ask whether the number of lessons constituted a sufficiently large sample on which to base important judgements – such as rating the teacher on a seven point scale. There have long been methods available for estimating adequate sample sizes. Why are there no Ofsted studies of this important issue?

---

<sup>2</sup> C. T. Fitz-Gibbon & N. J. Stephenson, *Evaluating Ofsted's Methodology* (draft report, University of Durham, 1996).

### 3.1.2 The Question of Consistency (Reliability)

If different inspectors would arrive at different judgements, then which judgement should be used? This problem can be addressed only by obtaining fair measures of 'inter-rater reliability': ie, measures of how *consistent* inspectors are. Do different inspectors arrive at the same conclusions when observing a lesson or when rating a school? There are numerous accounts of extremely inconsistent ratings when in Ofsted training sessions would-be inspectors are shown videos of lessons. There are no reports, it seems, to show that this fundamental problem is overcome.

How did headteachers view this issue? Two questions in the survey related to it:

Given the statement: *'Two Ofsted teams working without contact would come to the same conclusions about a school'* – only 15 per cent of headteachers agreed.

Given the statement: *'I believe the Ofsted teams have no difficulty in reconciling the judgements of each team member to provide a corporate view'* – only 37 per cent agreed.

Furthermore, in a progress report by HMI on the first year of secondary inspections, published by Ofsted<sup>3</sup>, we find:

*'The majority of RgIs were able to make appropriate decisions about conflicting evidence.'* (p.10) – a statement which clearly implies that the majority of Registered Inspectors *were* presented with conflicting evidence.

### 3.1.3 Validity

The question of validity is the question as to whether Ofsted's judgements are, despite the clear shortcomings already noted, correct.

Several types of validity should have been checked before Ofsted was allowed to operate.

**Construct validity.** Is it methodologically reasonable to apply a single label to a whole school given that there is almost certainly considerable variation within every school? Is a 'failing school', for example, adequately defined? In extensive value-added studies over many years we have seen that within-school variation is considerable and, furthermore, non-academic indicators can give a quite different impression from the academic indicators.

**Concurrent and predictive validity.** Since Ofsted defined effective schools as ones in which pupils make average or better progress, value-added measures are clearly the ideal concurrent measure which should agree with inspectors' judgements. Despite the widespread availability of value-added measures Ofsted has failed to publish any studies of this fundamentally important question relating to the validity of their judgements.

Ofsted appears to claim a knowledge of 'good practice' but, even if inspectors agreed among themselves as to what constitutes good practice, this could be a shared prejudice rather than accurate knowledge. Indeed

---

<sup>3</sup> *Independent Inspections of Secondary Schools 1993–1994* (Ofsted, 1994).

teachers are all too aware that views on what constitutes 'good practice' have changed over the years, as, for example, in the fashion for group work which is now replaced by a fashion in favour of more whole class teaching. But which fashion will the inspector of your class follow?

Nine questions in the form of statements to which respondents answered on a five-point scale from 'strongly disagree' to 'strongly agree' were used to assess validity as perceived by headteachers.

Fewer than one in three heads (29 per cent) agreed that Ofsted inspectors can '*correctly identify failing schools*'.

Twelve per cent of headteachers agreed that inspectors '*can assess pupils' abilities better than can the staff of the school*'.

Fewer than 7 per cent of respondents thought that inspectors could assess the '*spiritual, moral, social and cultural*' aspects of schools separately.

**Evidence of Discriminant Validity.** Are inspectors confused by extraneous factors? In the survey quoted, there was a strong tendency for inner city and disadvantaged schools (as shown by the proportion of pupils in receipt of free school meals) to receive poor ratings.

Is it possible that the strong relationship between a poor inspection rating and being an inner city school reflected the incapacity of the inspectors to make adjustments for the difficulties of working in these schools and for the handicaps that pupils continue to experience in the urban environment? However, the association could be an indication of genuine and alterable problems in some inner-city schools, in which case we need urgent research into what remediation is effective.

Whatever the reason for a strong association between receiving a poor rating and having a high poverty indicator, we can take the relationship into account statistically, and then ask whether there are yet other factors which relate to the ratings received from inspectors.

Even after taking account of free school meals there remained a relationship between the amount spent buying in pre-inspection help and the rating received by a school.

If inspectors judge what they see and not what they are told, and if they are evaluating a school as it actually is, rather than on the basis of a self-presentation exercise, why should the purchase of pre-inspection help apparently have such a noticeable impact? This link between money-expended and rating-received is a situation which needs monitoring and further investigation.

**Validity of different kinds of information.** Despite the limitations of numerical indicators, they appear to carry more weight than inspectors' judgements in measurable areas.

*'If the judgement of the Ofsted team differed from data available to me on Value Added I would tend to believe the Ofsted judgement'*. Only 8 per cent agreed with this statement (only 1 per cent 'strongly').

The supposition that inspectors can judge the progress of pupils by sitting in a few lessons and looking at a few work samples needs checking.



### 3.2. THE ISSUE OF OFSTED'S IMPACT

One impact of Ofsted that is immediately apparent is *stress* and *the diversion of effort into preparation for inspection*. One way to put a realistic figure on the impact of the stress is to look at absences which heads attributed to the inspection process.

Heads were asked *'In your view, were there any stress-related staff absences before, during and after the inspection?'* This part of the questionnaire was often written over with statements such as 'Too many' or 'Too difficult to estimate'. There was a wide range of response in terms of staff absence. Absences before inspection were substantial and reported by half the heads. The overall average was 15.1 staff days. Absences during inspection dropped to a level of 2.3 staff days on average, but then shot up again after inspection to 28 staff days on average.

If staff absence were a sign of nervousness because of incompetence it would be expected that the amount of absence would be related to the rating the inspectors attached to the school. No such relationship was found. Stress was apparently related to little else, certainly not to the Ofsted rating of the school.

As for costs in terms of preparation, median values were: 40 staff days preparing documents, 10 days of the head's time on documents, £250 on reprographics and photocopying, 5 staff days on extra meetings of staff, zero on extra meetings with parents although some schools reported much more. Two staff days were reported as the median for extra meetings with governors and zero with the press, but with some schools reporting very large amounts (90 staff days).

The actual impact of inspections will have to be monitored; but what were the opinions of head teachers? They were asked:

*How much information of use to you in improving schooling did you gain from the inspection?*

Given that Ofsted's slogan is 'Improvement through inspection', it is worth looking at the distribution of results on this item.

Eighty nine heads had had inspections and the distribution was as follows: 5 per cent of heads reported having learnt nothing; 16 per cent reported 'not much'; 40 per cent reported 'some' (the middle of the scale); 33 per cent reported 'quite a lot' and 6 per cent reported 'a large amount' – results not overwhelmingly positive nor overwhelmingly negative. However, considering the costs, having 61 per cent reporting less than 'quite a lot' calls into question the value for money.

### 3.3. OTHER FACTORS

#### 3.3.1 Classroom Observation

The aspect of inspection which is the most expensive in inspectors' time, the most costly to schools in staff stress, and the least validated, is the practice of having inspectors sit in classrooms making amateurish attempts at classroom observation and drawing unchallengeable conclusions about effectiveness. It is this aspect of inspection which should be immediately

suspended pending the development of proper standards. It is doubtful that business or industry would permit an inspection regime, centrally imposed, that was based on opinion about how the business or industry should be run, not on sound research. This is what is being imposed upon schools in the public sector, despite the intentions of the welcome Local Management of Schools legislation.

### **3.3.2 Inadequately qualified inspectors**

Ofstin was informed of complaints that inspectors were not well trained in Health and Safety and that some make poor judgements and recommendations. This important observation raised the issue of the need to consider the competences Ofsted inspectors should possess. If inspectors are interpreting a body of statistical data, then they should be examined in their understanding of such data. If they are interpreting the adequacy of account-keeping then they should be examined in their knowledge of accounts. If they are serving as Health and Safety Officers they should be qualified to the highest standards since nothing is of greater concern to parents than the health and safety of their children.

### **3.3.3 The Case for an Independent Review**

There can be no substitute for inspection, but in its present form it is an amateurish and anachronistic operation. It is also a source of serious stress – and often of serious distress – to the teachers on whom we rely for the care of our children and grandchildren. There should immediately be an independent review, involving representatives from business, industry, medicine and statistics as well as education, to consider the role and methodology of the inspection process.

# CHAPTER FOUR

---

## Inspection and School Self-Evaluation

Professor Ted Wragg, School of Education, Exeter University

### 4.1. THE WEAKNESSES OF THE CURRENT SYSTEM

It was with some regret that Tim Brighouse and I concluded, in our pamphlet *A New Model of School Inspection* (Wragg and Brighouse 1995)<sup>4</sup> that school inspection was not working properly. In the past a number of different frameworks for inspection<sup>5</sup> have been considered, such as: (a) privatising the whole business by putting inspections out to tender; (b) extending the number of HMI from the 450-500 it was then to over 2000, permitting more regular visits to schools; (c) linking HMI and local inspectors in some way; (d) establishing greater control over what local inspectorates did.

It was no surprise that the Conservative Government, in the 1992 Education (Schools) Act, opted for the first of these. The market philosophy espoused by successive education ministers, the beliefs expressed in some of the right wing 'think tanks' that HMI and LEA inspectors were part of a liberal consensus, and the reduction in the powers and budgets of local authorities, all combined to change school inspection significantly. When the Office for Standards in Education (OFSTED) was established, over half of HMI left or retired, most not being replaced. At local level, the moneys allowed for local inspection in the Revenue Support Grant were decreased, so LEA advisers' and inspectors' posts were terminated or substantially reduced in number.

Training by OFSTED consisted of a one week course, usually in a hotel. There is very little that one can train for in a week. It took me several months to get my collector's badge in the cubs. The system was only rescued because HMI and early retired LEA inspectors joined inspection teams.

Tim Brighouse and I have advocated<sup>6</sup> that for schools to be 'dynamic' – that is, to improve pupils' achievement by changing in a judicious way – they must secure the commitment of the whole community – head, teachers, pupils, parents, governors. Furthermore, raising standards of achievement should be seen as 'improving on previous best'. The present inspection arrangements are not effective. The following problems are but a sample.

- Inspection is divorced from advice, so that the whole process is

---

<sup>4</sup> E. C. Wragg and T. Brighouse, *A New Model of School Inspection* (Exeter University School of Education, Media and Resources Centre, 1995).

<sup>5</sup> E. Bolton, 'Alternative Education Policies: School Inspection' in S. Tomlinson (ed) *Educational Reform and its Consequences* (London, IPPR/Rivers Oram Press, 1994).

<sup>6</sup> E. C. Wragg, *An Introduction to Classroom Observation* (London, Routledge, 1994).

T. Brighouse, *What Makes a Good School?* (Stafford, Network Educational Press, 1991).

detached from the daily running of the school, thereby reducing the impact.

- ❑ Procedures are cumbersome and bureaucratic.
- ❑ There is little benefit during the week of the inspection itself, because of the tensions felt by the school staff, though there are some benefits during the period of preparation.
- ❑ Despite some of the benefits that accrue from preparation for inspection, schools are asked to provide too much information, a great deal of which is not necessary.
- ❑ Although schools provide information, there is no proper place for sustained self-evaluation in any substantial form over a period of time.
- ❑ High anxiety is generated; so much so that inspection is seen as a hoop to be jumped through, rather than as a central part of a sustained programme for improvement.
- ❑ Reports are written to a formula, with too much prominence given to comparison with national norms, and not enough to a thorough analysis of the school.
- ❑ The language of inspection reports, littered with phrases like 'generally satisfactory' or 'sound', bears little resemblance to the normal language of debate and discussion on educational matters and is too imprecise to be helpful.
- ❑ Formal recommendations concentrate too much on the structure and administration of the school and not enough on what happens in the classroom.
- ❑ Emphasis on inspection as a private profit-making business leads to an undesirable confusion over the role of inspectors, so that, following complaints from schools, OFSTED had to issue an official warning to inspectors of the dangers of 'misusing their position in order to seek work in a consultancy capacity'.
- ❑ The use of *ad hoc* teams of professional and lay inspectors, many of whom are retired or work on an occasional basis, makes it difficult to ensure consistency between teams in the application of criteria used for judgement.
- ❑ The shortage of primary inspectors has meant that the programme of primary school inspection has fallen behind schedule, and inadequately prepared secondary specialists are undertaking primary inspections.

There must be a better way of meeting the need for public accountability by ensuring that schools are operating effectively. Wholehearted commitment within the school to improving what has been achieved previously is the starting point. Local support is a very important feature, but external moderation is essential.

There have been two significant gains since the establishment of Ofsted. The first is that there can be a **regular cycle of inspection**. The second is the existence of a **published framework document** describing what

inspectors look for and record. The four year cycle is too frequent and does not discriminate between schools that are running well and those that are doing badly. The framework document has been too detailed and too closely linked to the crude and insensitive application of national norms. We should nonetheless like to retain these two gains in some form.

#### 4.2. FIVE PROPOSALS FOR IMPROVEMENT

We made five proposals to improve the quality of teaching and learning in schools.

##### 4.2.1 The Office for Standards in Education should be closed down in its present form and there should be a new national and local structure for school inspection. The advantages and disadvantages of local and national inspection are well known. The strength of local inspectors is that they

should know their local schools better than outsiders and can monitor a school over a period of years. The weaknesses are that local inspectors may take too parochial a view or be dominated by strong-minded individuals who can sponsor certain schools, teachers or teaching methods.

The number of inspectors in Her Majesty's Inspectorate should be a team of 450 fully trained members. Secondly, local authority inspectors should be available on secondment to HMI for 20 per cent of their time, one day a week on average, the other 80 per cent being for sustained and regular support, monitoring and advice for schools. Thirdly, a number of headteachers from primary, secondary and special schools should be released from their teaching for up to two years, to supplement the inspection teams, and to be available additionally for advisory follow-up work. The fourth point is that there should be trained lay inspectors in each team. Lay inspectors should not be asked to undertake the inspection of

classroom teaching: that is a professional matter. Lay inspectors should concentrate on the views of the parents and others, the use of premises, the appearance and maintenance of the school, the well-being of pupils.

##### 4.2.2 Although a school's structures, plans and administrative arrangements are of interest, the principal focus of school inspection should be on improving the process of teaching and learning in the classroom. Too much attention has been paid to structural and administrative matters and not enough to improving the nature and quality of classroom learning. When John Gray and Brian Wilcox<sup>7</sup> studied the recommendations made in Ofsted reports, they found that the great majority concerned curriculum documentation, management and administration, and curriculum delivery. The assumption

<sup>7</sup> J. Gray and B. Wilcox, 'In the aftermath of inspection: the nature and fate of inspection report recommendations', *Research Papers in Education*, vol. 10, 1, pp 1-18 (1995).





behind this approach is that teachers are propelled by organisation charts and policy statements. They are not. While these can be a useful framework for action, they are not a substitute for it.

One-off inspections are but a pin-prick. A study of teacher appraisal at Exeter University<sup>8</sup> found that fewer than half of teachers said they changed their classroom teaching as a result of appraisal. Teaching strategies are not modified dramatically after a single inspection. For teaching to improve, teachers must themselves be involved in making professional decisions. Real and lasting improvement is often spread over a period of time, not achieved instantaneously.

- 4.2.3** There should be a better framework for inspection, with certain core features that apply to all schools, as well as options that recognise that primary and secondary, urban and rural, big and small, rich and poor schools are different from each other. The present system assumes the false premise that all schools are exactly the same. It is as if the inspection of a hospital were to attach greatest importance to the discovery that death rates in the ward for the terminally ill are higher than those in the ward for patients with in-growing toenails. There is no recognition of different starting points, or of the need to compare like with like. When teachers read that their school is 'generally satisfactory' or 'sound', this imprecision does nothing to engage them in proper professional discourse about improvement. Schools should set and achieve their own targets, as happens currently in Birmingham and in some other areas. If every school improves, then the whole level of national achievement goes up.
- 4.2.4** In order to secure the full commitment of the head, teachers, governors and parents, there should be a proper place for a school's **self-evaluation**. This should be available to the inspecting team whenever the school has a formal inspection. There are certain requirements if self-evaluation is to work. A proper structure should be drawn up, which need not be a straitjacket, as it should encourage imagination and individuality as well as offer guidance. Secondly, the local inspectors, between inspections, ought to be identifying any schools that are not effectively evaluating what they do. Thirdly, any school unable to evaluate itself would not be licensed, as self-evaluation is a pre-condition for being awarded a licence (see 4.2.5 below).
- 4.2.5** Schools that are running well should be given a **five year licence**, allowing them to carry on under self-monitoring procedures. At the end of this five year period, or sooner if there appear to be problems, a school would receive a shorter inspection. Its five year licence could then be extended for another five years. After each ten year period, however, there would be another full inspection. Inspections could also be triggered by certain other events, such as a significant change in the nature of the school. The awarding or renewal of a five year licence following an inspection would not mean that a school had reached a state of perfection, but that it should now be allowed to get on with its business of improving teaching and learning under local supervision.

---

<sup>8</sup> E. C. Wragg, F. J. Wikeley, C. M. Wragg and G. S. Haynes, *Teacher Appraisal Observed* (London, Routledge, 1996).

### 4.3. AN INDEPENDENT NATIONAL COUNCIL

A National Council for the Inspection of Schools should be established. It should consist of people who are thoroughly familiar with the analysis of teaching and learning, school processes, and the improvement of teaching. Members should not be political nominees lacking expertise in the field. There should be people who are used to visiting classrooms to advise teachers, who have good experience of classroom observation, pupil learning and identifying effective and ineffective teaching. There should also be a place for some lay members with an interest as governors, parents or employers.

The National Council would be independent of government, the local education authorities and the inspectorate, though its members may well have experience working within these bodies. Any school failing to obtain a licence would be referred to the Council for an individual solution to its problems, a solution which, in an extreme case, should involve closure. The Council would draw up and review criteria for inspection and for self-evaluation.

We propose that a critical focus of both inspection and self-evaluation should be on the support and nurturing of members of the school staff. This would include professional development for all teachers, and looking at how the school dealt with teachers who were regarded as incompetent, including the use made of disciplinary and dismissal procedures where these were appropriate. Schools that did not handle this matter properly would not fulfil the conditions to obtain or keep their licence.

Inspection should be a public service, not a private profit-making business. It should concentrate on what happens in the classroom, not on the bureaucracy and paperwork of running a school. It should involve schools in self-evaluation as well as external inspection. It should combine the best of local and national inspection traditions. It should be fair but rigorous, with a core of activities undertaken in all schools and a set of individual activities carefully tailored to the school under scrutiny. It should challenge schools to improve what they do and then license them to do it.

# CHAPTER FIVE

---

## The Evidence Assessed

Elliott Stern, Director, Evaluation Development and Review Unit, The Tavistock Institute

### 5.1. THE PANEL OF ASSESSORS

An independent panel of assessors, drawn from a wide area of expertise in public administration, business, educational administration and research, was invited to attend the Oxford conference and review the evidence presented.

Members of the panel were:

Dr Tim Blackman	Oxford Brookes University (School of Social Sciences)
Andrew Collier	General Secretary, The Society of Education Officers
Michael Fischer	Chief Executive, Research Machines plc
Professor Anthony Flew	Reading University
Dr Janet Lewis	The Joseph Rowntree Foundation
Jane Seddon	The British Deming Association
Elliott Stern (Chair)	Director, EDRU, The Tavistock Institute, London
Professor Sally Tomlinson	Department of Education, Goldsmiths' College, London
Margaret Tulloch	The Campaign for State Education
Brian Whitworth	Independent Management Consultant.

The oral evidence accumulated during the first day of the conference was summarised for us by members of the conference organising team under four headings: the methodology of the Ofsted inspection process; the impact of Ofsted inspection on schools in both the short and longer terms; the cost effectiveness of the Ofsted inspection system; the views of the conference as to the best way to ensure effective inspection in the future. These summaries were presented in full conference, without schools or respondents being named, and opportunity was provided (and taken) for assessors to question and individual members of the conference to comment on the summaries as presented before the assessors withdrew to consider our response.

### 5.2. THE NATURE OF THE EVIDENCE

In some respects the format was difficult. For reasons which we fully understood, the evidence gathered by the conference tended to focus more on schools' short-term experiences of inspection than on its longer term effects. It was difficult for the conference rapporteurs, in the twenty minutes or so allotted to them, to prioritise and substantiate all the points that had arisen during the evidence-taking sessions. Some issues of interest to us, such as the precise basis of the methodology of inspection and the technical validity of the judgements based upon it, were only lightly

touched upon. We were conscious too that the evidence we were hearing might well be partial, in the sense that it came from a self-selected sample of schools whose experiences of inspection might not reflect the generality of opinion. In the event, though, both the composition of the conference – we heard the views of registered inspectors, governors and parents as well as those of heads and teachers – and the tone of the evidence presented to us reassured us on this point. Almost all the evidence we heard, for example, indicated acceptance of external inspection as a legitimate and necessary element in the proper accountability of schools to their parents and the public, and as a potentially powerful factor in their improvement. Our discussion, after the evidence-giving session, was largely about how inspection could be structured and managed so as to secure both of these expectations.

The evidence indicated that current practice falls short of this in some respects. The stress on the accountability function of inspection may be necessary and appropriate but it appeared to us that its effect is to create a climate in schools in which inspection is seen as a wholly external process – ‘something that is done *to* us’, as one respondent said, ‘but not *with* us.’ It was clear that in this context inspection can be seen as a threat to schools, and something to which they react negatively and defensively rather than positively and openly. There was clearly a view that both the tone of some of the HMCI’s public pronouncements about national data from inspections and the requirement that Ofsted should grade all teachers on the basis of the individual lessons seen were likely to contribute to this effect. Many teachers appear to believe that Ofsted is more concerned with identifying poor teachers than with identifying the other factors that contribute to poor school and pupil performance. True or false, this perception may weaken the potential of inspection as a tool for school improvement.

Some aspects of Ofsted methodology may have the same effect. It is based predominantly on lesson observation during the two or three pre-notified days when all members of the inspection team are in the school together. The argument is that inspectors can only make the judgements that the inspection contract requires on the basis of what they have seen in practice. But according to the submissions that we heard, schools and teachers often believe that the *context* of the lessons seen is thereby disregarded. Inspection, we were told, is essentially a snapshot; but education is a process, and the one can give a seriously misleading impression of the other. It may well be, the conference agreed, that too many teachers have low expectations of their pupils; it may equally well be that too many inspectors have an inadequate view of the circumstances in which schools are working and of the progress they have made to improve them. In this respect, we were concerned to note the requirement put upon inspectors to judge school performance against national norms, and to regard a marked deviation below those norms as an indication that the school is failing or at risk of failing. We thought that local circumstances need to be taken into account as well. The current procedures for identifying failure appear to be consistently singling out schools in inner city areas with a high incidence of economic and ethnic instability. This is likely, we felt, to be a self-fulfilling prophecy.

That said, it appeared to the panel that most schools' concerns about inspection were more about its effects than its methodology. In particular, schools were concerned about the levels of stress that inspection imposed. There was evidence to this effect in all the presentations made to us. Unacceptable teacher stress was reported in all types of school, and at all stages of the inspection process. In response assessors pointed out that stress was a factor in all types of employment, and was certainly not confined to schools. Nonetheless, it was clear to us that 'good' teachers in 'good' schools were just as likely to report it as were teachers in schools that had received a critical report, and that headteachers in particular were deeply concerned about its effects on staff energy and staff morale. Over-preparation is no doubt a factor in this: it is wholly understandable, given the anxiety of teachers and schools alike to get a good report. Nevertheless, there does seem to be an assumption in the current Ofsted process that pain is a necessary condition for improvement. Members of the panel wanted to challenge that perception, not least because teacher stress is very wasteful. In any school the time and energy of good and well-trained teachers are key resources for improvement, yet we heard disturbing evidence of increased ill-health, teacher absence and premature retirement. If this is directly attributable to inspection, it is clearly a factor which needs to be considered in any calculation of the cost-effectiveness of the process. How effective is inspection, pound for pound, when compared with other strategies for securing school improvement?

### **5.3. IS INSPECTION COST-EFFECTIVE?**

We thought this question both the most important and the most difficult to answer. It was our impression that, notwithstanding the stress involved, Ofsted inspection was seen by most schools (and by most governors and parents) as reasonable and professional. What was striking was that it was not generally seen as 'helpful'. Even after what were described as 'good' inspections, we were told, the inspectors' reports were often disappointing. They tended to be bland, poorly focused and prioritised, and more concerned with managerial changes – typically, policies and procedures – than with changing the way that teaching was practised and experienced. The fact that governors and parents shared these views led us to suspect that in this respect at least the inspection system might be falling short against both of its key objectives – accountability and improvement.

One difficulty is that it is hard to find evidence of school improvement that is directly related to inspection. This may be inevitable, given that improvement is necessarily a long-term process. Even so, it was disturbing that much of the anecdotal evidence identified a dissonance between the key issues for attention listed in the inspectors' reports and the school priorities as identified in the school's statutory development plan. Legislation requires the former to be addressed in a public action plan, approved by the governors within 40 days of the publication of the report and reviewed by them in twelve months' time. That creates the appearance of improvement but not necessarily the reality. What matters more, we suspect, is a whole school agreement on priorities and strategies, and that is likely to rest on changing attitudes and expectations rather than policies and procedures. How successful has Ofsted been in achieving this sort of change? That seemed to us to be a very open question.



There remained the question of the validity of inspection judgements. Registered inspectors are required to ensure that these are secure ('rooted in a substantial evidence base and informed by specified quantitative indicators'), first hand, reliable, valid ('in that they accurately reflect what is achieved and provided by the school') comprehensive and corporate. The evidence we heard suggested that most headteachers thought that their inspection had met these criteria. But this list begs as many questions as it answers. Inspectors make *qualitative* judgements, and until we know, say, that ten inspectors visiting a particular lesson will give it the same grade on a scale of 1 to 7, we are entitled to have doubts about the validity both of individual judgements and their aggregation into a whole school or even sometimes a whole system picture. It may well be that too much weight is being placed, at every level, on a process that is more impressionistic than precise.

That, in outline, is a reflection of the panel's thinking as members digested and discussed the afternoon's submissions. Over tea, we agreed the outline of our own, necessarily brief, report.

#### **5.4. THE ASSESSORS' VIEWS**

##### **5.4.1 Process – or Performance?**

The current inspection system is primarily about examining the process of school education. There is an important argument, not yet we think adequately addressed, as to whether the proper purpose of inspection should be to examine its results, leaving it to schools – that is, to governors, heads and teachers – to devise the appropriate methods.

##### **5.4.2 Evidence of Good Practice**

That said, few members doubted that the present inspection regime was having some beneficial effects. The evidence suggested that the prospect of inspection had helped schools to focus on improvement. What was less clear was whether inspection itself had brought about significant change; and whether significant change could be quantified into significant improvement. What was wanted, we suggested – and what was lacking in the present inspection system – was *evidence* about what constitutes good practice in schools, and about what could be shown to lead to improvement in school outputs. The collection of such evidence – a quite separate function from inspection – ought to be a national priority.

##### **5.4.3 The Measurement of Outcomes**

A third priority was the measurement of improvement. The Ofsted model of inspection measures the processes of teaching and school management against the model of education that is described in the Ofsted Handbook, but is unclear about the measurement of outcomes. Are schools being judged in terms of their own performance over time, or against the performance of other schools, or against an assumption of national targets and standards? There is a confusion here which needs to be resolved.

#### 5.4.4 Uncertainty over Purpose

Part of the confusion stems from an uncertainty over purpose. Beyond the two objectives of the system we have already identified, namely the provision of a mechanism for accountability and a means to school improvement, there are two others: the maintenance of minimum standards of quality and the collection of standardised national data about school performance. Current inspection aims to meet all four objectives in a single framework, applied to the letter and at laid down intervals in all schools, regardless of their circumstances, size and needs. We recognise that the *Framework* has been significantly changed (which casts into some doubt the validity and feasibility of the fourth objective) but it remains prescriptive, bureaucratic and managerial. Inspection under its provisions is also deliberately *external*, in the sense that nobody with an internal knowledge of the institution is admitted to the team. Conversely, no member of the team is allowed to work with the school and its governors on how best to put recommendations into practice. That may be good for accountability, but it is most unlikely to be good for school improvement.

#### 5.4.5 Are Ofsted's Methods Valid?

We do not know – nor, we suspect, does anyone – whether current inspection methodology is appropriate to the range of purposes it serves, and whether it meets the criteria that Ofsted itself lays down. We suspect, on the evidence made available to us, that it bears far too heavily on disadvantaged schools, and sometimes imposes additional disadvantage on them. We suspect, too, that in its complexity it puts primary schools at a disadvantage. In either case, its results may not accurately reflect the schools surveyed: a possibility emphasised by so many of the evidence-givers that we give some credence to it. We are surprised that Ofsted itself, which currently requires its inspectors to sign the Official Secrets Act, does not admit this possibility, and open its processes to public examination and if necessary to formal complaint. We therefore believe that research should be commissioned into the effectiveness and validity of the Ofsted methodology – and into the comparable methodologies of other inspection systems, in this country and abroad.

#### 5.4.6 The Cost of Inspection

Finally, we have serious concerns about the costs of the Ofsted inspection process, both directly, in terms of the resources switched to Ofsted from other educational budgets and indirectly, in terms of the cost of inspection to schools in time and stress. There is clearly a need for a cheaper and, we believe, a more flexible process. We doubt whether the lengthy period of notice of an impending inspection currently given is helpful to either the inspectors or the schools.

### 5.5. CONCLUSION

Our conclusion is that there is a real need for a thorough and independent review of the inspection process and methodology.

## AFTERWORD

---

In concluding that 'there is a real need for a thorough and independent review of the inspection process and methodology', the assessors echo many voices coming directly from schools, as well as much concern in academic analysis and among professional bodies.

Virtually all of these calls share three important features. They do not start from a complacent assumption that school improvement is unimportant or unnecessary: on the contrary, they regard it as a vital and abiding concern. They do not attack the principles of inspection or of accountability in the context of education: they take both as necessary and integral aspects of any major public service. They do have an awareness of methods of accountability, inspection and improvement in other spheres of public and private enterprise – methods beside which those of Ofsted frequently appear amateurish and archaic.

The growing evidence of the inadequacies of the Ofsted arrangements, in terms of processes, outcomes, and cost-effectiveness, is precisely why professional opinion calls for review and improvement.

The motto for the Oxford Conference was *Quis custodiet ipsos custodes?* That challenge remains unanswered. We hope that this report of the conference proceedings will re-focus debate on its importance, and on the urgent need to secure an inspection system that genuinely contributes to school improvement.



