FORUM
Volume 54, Number 2, 2012
www.wwwords.co.uk/FORUM

Ofsted Inspection Inspected: an examination of the 2012 framework for school inspection and its accompanying evaluation schedule

COLIN RICHARDS

ABSTRACT Ofsted has always courted controversy. With the appointment of a strident new chief inspector its operations are likely to remain, or become increasingly, controversial. This article provides a detailed critique of key documents which describe the new inspection regime that for good or ill will have major consequences in schools. Although in certain limited aspects they represent an improvement on the previous inspection regime, the new requirements have many highly problematic elements which undermine the integrity and validity of inspection judgements. The article argues that schools in disadvantaged areas are likely to suffer most from some of the deficiencies highlighted.

Introduction

The Office for Standards in Education (Ofsted) was created in 1992. Since then the purposes of school inspection which it regulates have been variously defined. According to the inspection framework introduced from January 2012 consequent on the 2011 Education Act, school inspections:

> - provide parents with an expert and independent assessment of how well a school is performing and help inform those who are choosing a school for their child to attend in the future
> - keep the Secretary of State for Education (and Parliament) informed about the work of schools; this provides assurance that minimum standards are being met, provides confidence in the use of public money and assists accountability
> - promote the improvement of individual schools and the education system as a whole. (2012a, p. 4)

Ofsted was one of a number of instruments of school accountability set up by central government in the last decade or so of the twentieth century. Others included the enlargement of the powers of school governing bodies, the requirement laid on local authorities to monitor school performance, the introduction of national testing of selected outcomes of schooling and the publication of performance or 'league' tables. Unlike other instruments of accountability, Ofsted evaluates, and reports on, not just the outcomes of education in individual schools but also the quality of the education provided. It also claims to seek out and report on interrelationships or associations between processes and outcomes based on a range of evidence including first-hand observation of work by inspectors.

Ofsted inspection of schools has been the subject of controversy since 1992 – in terms of its purposes, its independence, its methodology, the expertise of its inspectors, its contracting procedures, the use of its results by successive chief inspectors, secretaries of state and others and particularly in terms of its effects on schools and individual teachers. This article is not primarily concerned with inspection procedures or with the effects of Ofsted inspection on schools, teachers and the education service. Instead it attempts to identify some of the limitations, as well as some of the possibilities, of the 2012 Ofsted inspection framework and its accompanying evaluation schedule for the inspection of maintained primary and secondary schools and academies.[1][2]

It needs to be acknowledged at the outset that taken together these documents constitute an improvement on the previous inspection regime as evidenced by a sharper focus, the abandonment of limiting judgements and much more reliance on classroom observation. However, the documents still contain highly problematic elements which undermine the integrity and validity of inspection judgements.

Inspection, Aims and Values

Some of those reading the framework might assume that inspection is a clearly understood, unproblematic activity whose nature is uncontroversial. But is it? Perhaps referring to the legal framework under which Ofsted operates might help in elucidating what inspection is and what it should evaluate.

At the beginning of the inspection framework it is clearly stated that following the Education Act 2011 inspectors are required to judge and report on

> the quality of education provided in the school, its overall effectiveness, and in particular cover:
> - the achievement of pupils at the school
> - the quality of teaching in the school
> - the quality of leadership and management of the school
> - the behaviour and safety of pupils at the school.

In reporting, inspectors must also consider:

> - the pupils' spiritual, moral, social and cultural development at the school;
> - the extent to which the education provided by the school meets the needs of the range of pupils at the school, and in particular the needs of disabled pupils and pupils who have special educational needs. (2012a, p. 4)

However, within the framework and the evaluation schedule there is no explicit requirement to inspect in any detail the quality of the school's curriculum despite the fact that it is the major vehicle for transmitting knowledge, understanding, skills and attitudes and so makes a very important and direct impact on achievement. Nowhere in the framework or the evaluation schedule is guidance given on how to judge the quality of the curriculum beyond token references to provision that is 'broad and balanced' - themselves contentious, value-laden and undefined adjectives. In a different document Ofsted does provide half a page of subsidiary guidance (Ofsted, 2012c, p. 19) but this is nowhere as detailed or as focused as the guidance provided on the inspection of the four main foci identified in the paragraph above. Ofsted claims that under the new arrangements 'inspectors are focusing more sharply on those aspects of schools' work that have the greatest impact on raising achievement' (2012a, p. 5) but presumably does not believe the curriculum to be one of those key aspects.

Inspection involves far more than observing, gathering evidence and reporting it orally or in writing. Inspectors are not simply a human form of camera neutrally capturing what goes on in schools or a human form of computer registering performance and other data. They have to make complex *judgements*, not precise measurements; they have to interpret, not just record; they have to make judgements about whether what they are observing or scrutinising is *worthwhile*; they can only make those judgements validly against a background of aims and values. But what are those aims and values? Ofsted is silent; the new framework and the evaluation schedule contain no discussion of, or reference to, these. Ofsted requires its inspectors to make a summary judgement of a school's overall effectiveness but in the absence of explicit aims and values underlying the process of inspection itself, how can that overall judgement be justified? (Richards, 2001).

Many of the aspects of schools on which Ofsted inspectors report are value-laden and contentious: 'quality', 'effectiveness', 'improvement', 'outstanding', 'inadequate' are obvious examples. Nowhere does Ofsted acknowledge, let alone discuss, the 'value-ladenness' of many of the key terms it uses.

The absence of any explicit references to educational aims and values in both the framework and the evaluation schedule is not surprising since the English educational system has long lacked a sufficiently detailed statement of the aims and values which should underpin and inform its activities, including those of school inspection. One major attempt was published in 1999, although

this dealt with only one, though central, facet. This was the statement of the values, purposes and aims underlying the school curriculum published in the National Curriculum handbooks for primary and for secondary teachers in England (DfEE/QCA, 1999a, b). This was incomplete, however, focusing as it did on the school *curriculum* rather than on the *education* provided in schools. Some of its phrases were loosely worded and subject to diverse interpretation. But even so, as Bramall & White pointed out, it presented 'a coherent, humane vision by which schools can be guided' (2000, p. 14).

After 1999 and before 2012 there was only one official attempt at spelling out aims – in this case for the National Curriculum in Key Stages three and four. These involved enabling young people to become:

> - successful learners, who enjoy learning, make progress and achieve
> - confident individuals, who are able to live safe, healthy and fulfilling lives
> - responsible citizens, who make a positive contribution to society.
> (QCA, 2007)

Most recently, the non-governmental Cambridge Primary Review proposed a set of more detailed 'core educational aims' for primary education that are capable of adoption/modification for secondary and special education. These were detailed under 12 headings:

> well-being
> engagement
> empowerment
> autonomy
> encouraging respect and reciprocity
> promoting interdependence and sustainability
> empowering local, national and global citizenship
> celebrating culture and community
> exploring, knowing, understanding and making sense
> fostering skill
> exciting the imagination
> enacting dialogue.
> (Alexander et al, 2010)

The Education Act of 2011 makes no reference to explicit aims or values but the authors of *The Framework for the National Curriculum: a report by the expert panel for the National Curriculum Review*, published later in 2011, suggest five – clearly influenced by what they perceive to be the priorities of the current Secretary of State. These are to

> 1. satisfy future economic needs for individuals and for the workforce as a whole …
> 2. appreciate the national cultures, traditions and values of England and the other nations within the UK, whilst recognising diversity and encouraging responsible citizenship;

> 3. provide opportunities for participation in a broad range of educational experiences and the acquisition of knowledge and appreciation in the arts, sciences and humanities, and of high quality academic and vocational qualifications at the end of compulsory schooling;
> 4. support personal development and empowerment so that each pupil is able to develop as a healthy, balanced and self-confident individual and fulfil their educational potential;
> 5. promote understanding of sustainability in the stewardship of resources locally, nationally and globally. (DfE, 2011b, p. 16)

Yet in the absence of a clear, officially endorsed statement of the aims and values which the English education system deems important and which primary and secondary schools are expected to pursue, how can an Ofsted inspection validly

> - gather evidence about a school's activities;
> - interpret that evidence in the light of aims and values;
> - evaluate that evidence through judging the worthwhileness of the activities in fostering commonly agreed aims and in embodying those values;
> - and report those judgements?

It cannot.

Currently then, the Ofsted inspection process involves collecting, evaluating and reporting evidence but without any explicit reference to aims or values. So, for example, unless they are tied into explicit values and aims, judgements that x and y are 'good' or 'outstanding' mean nothing apart from conveying a general sense of approval. Judgements of the 'effectiveness' of schools are empty unless it is made clear what aims are being effectively achieved and what values successfully embodied. Judgements of 'strengths' and weaknesses' in schools can carry no weight with the parties to the inspection enterprise unless they are in broad agreement on the aims of the activities being inspected.

HMI inspections pre-Ofsted avoided *in part* the problem of how to make qualitative judgements in the absence of explicit aims and values by evaluating how well individual schools were pursuing *their own* aims and values. This idea might usefully be reconsidered in any future revision of Ofsted inspection methodology. However, it has to be acknowledged that such school-derived aims were often unclear, indeterminate and rhetorical, rather than informative of practice. Also, in asking themselves 'Was what was intended good enough?' HMI were able to smuggle their own implicit values and aims into the inspection process and into the judgements they made about practices in individual schools. Neither old-style nor new-style inspection has yet successfully addressed the issue of underlying aims and values.

As a national inspection regime Ofsted needs to clarify how its framework and evaluation schedule relate to explicit aims and values – whether those it generates for itself, those embodied in the National Curriculum handbooks, those proposed by the Cambridge Review, those suggested by the National Curriculum expert review panel or those eventually (if ever) produced as the officially endorsed aims of English schooling. Or alternatively schools might be required to generate statements of their own aims and values and inspectors be required to evaluate how successful they are in meeting them. If, as is currently the case, the inspection process is not informed by these kinds of explicit aims and values, then Ofsted itself or its individual inspectors will inevitably invest their judgements with implicit aims and values of their own, which could in some respects run counter to those of the educational system itself or to those of the schools they are inspecting.

The Language of Inspection

While introducing some semi-technical phrases of its own, e.g. schools requiring 'special measures' or schools given ' a notice to improve', Ofsted uses everyday English in which to express its guidance to inspectors and its findings and judgements to those in schools.

In a publication issued 13 years ago Ofsted expected its reports to

> be clear to all its readers, governors, parents, professionals and the public at large; …
> use everyday language, not educational jargon, and be grammatically correct; …
> use telling examples drawn from the evidence base (of the inspection) to make generalisations understandable and to illustrate what is meant by 'good' or 'poor'; …
> employ words and phrases that enliven the report and convey the individual character of the school.
> (Ofsted, 1999b, p. 145)

Neither the 2012 framework for inspection nor its accompanying evaluation schedule refers to the qualities of the language required in Ofsted communication, but presumably at least the first two criteria still apply.

The sentiments underlying the 1999 guidance are admirable. Putting aside the impossibility of ensuring clarity to *all* readers, how possible is it for both inspection guidance and reports written in 'everyday' English to be clear, unambiguous and free from the possibility of subtly diverse interpretation ?

That 'everyday' language is potentially flexible, rich and subtle, though in practice the constraints placed by Ofsted's detailed grade descriptors are likely to lead to rigid, formulaic, impoverished writing. But the language of the framework, of the evaluation schedule and of inspection reports is *inevitably* shot through with ambiguities, imprecision and the possibility of being differently

interpreted, not just because of the nature of educational phenomena but because of the nature of 'everyday' language itself.

Take just one example, drawn from Ofsted's criteria for judging the quality of teaching:

> inspectors will evaluate
> - the extent to which teachers secure high quality learning by setting challenging tasks that are matched to pupils' specific learning needs.
> (2012b, p. 11)

Teachers and inspectors will have somewhat different views as to what the terms used in this criterion mean based on their previous experience, knowledge, understanding, values and beliefs. For example, what one inspector may judge as an example of 'challenging tasks', another may judge as an example of 'mismatch' if some of the children fail to engage in the activity.

What, then, is meant by 'pupils' specific learning needs'? Is there agreement on what the term means? More specifically, in relation to any particular lesson, what 'specific learning needs' are being referred to ? How do inspectors come to a common understanding of what these needs are? What is meant by 'matching' tasks to 'need'?

Is 'high quality learning' defined in the framework or schedule ? If not, as is the case, does it mean subtly different things to different people, including the various members of an inspection team?

Similar points could be made in relation to all of Ofsted's inspection criteria. The attempt made in the evaluation schedule to prescribe closely how these criteria are to be used fails to do justice to the nature of the language in which the criteria have to be expressed.

The point of this is *not* to argue that Ofsted inspectors should use a technical language, as free as possible from diverse meanings or interpretations. They could not, even if they were required to. The nature of educational phenomena precludes such a language. Ofsted does *indeed* need to issue guidance and to report in 'everyday' terms if it is to communicate reasonably effectively with readers, but inspectors and teachers need to appreciate the imprecise nature of the language used in documentation, and the inevitable variety of interpretations that will be placed on what is written. They need to recognise that the language used 'has countless hair-triggers inside it … . The 'drunkedness of things being various' may be unleashed by any plain, simple and sober-seeming word (Bowie, 1998, p. 320).

The Inspection of Achievement

'Achievement', the first focus of the 2012 framework, has been a major concern of Ofsted and before that of Her Majesty's Inspectorate of Schools. No one seriously disputes that inspection should involve judgements of whether children are achieving well or otherwise. There is, however, considerable

controversy over what 'achievement' means, how it relates to 'attainment' and 'progress', and how judgements related to these are determined.

The Downgrading of the Concept of 'Standards'

Under previous education acts such as the Schools Inspection Act of 1996, the inspection of educational standards was seen as central to inspection. For example, Section 10 of that act required inspectors to report on

> - the educational standards achieved in the school;
> - the quality of education provided by the school;
> - whether the financial resources available to the school are managed efficiently;
> - the spiritual, moral, social and cultural development of pupils at the school.
> (Ofsted, 1999a, p. 4)

Similarly, the 1999 schedule to which all Ofsted inspections had to adhere highlighted as one of its first major sections:

> How high are standards?
> 2.1: The school's results and achievements
> 6. Pupils' attitudes, values and personal development.
> (p. 34)

This centrality of 'standards' to the inspection process has now been very largely superseded in the 2012 framework. The framework makes only a small number of cursory references to 'standards' and uses the term 'achievement' instead. 'Standards' have lost their pride of place and where they do appear their inclusion appears primarily rhetorical.

It is particularly noteworthy that clause 40 of the 2011 Education Act quoted earlier replaces 'standards' by 'achievement' when it requires inspectors to judge and report on the quality of education provided in the school.

Elsewhere in the framework there are only four brief references to 'standards':

> (a) the use of inspection to provide the Secretary of State and Parliament with 'assurance that minimum standards are being met' (p. 4);
> (b) the use of inspection criteria and grade descriptors to 'illustrate the standards of performance and effectiveness expected of the school' (p. 4);
> (c) the reference to 'an acceptable standard of education' in defining schools judged to be inadequate (p. 13); and
> (d) the consideration to be given by inspectors to 'the standards attained by pupils by the time they leave the school, including their standards in reading, writing and mathematics' (p. 14).

In none of these cases are 'standards' defined.

An earlier inspection handbook, however, was clear about their nature:

> In this guidance we use the term standards to denote the educational attainment of pupils in relation to some clear benchmark, such as National Curriculum levels, or descriptions, at the end of a Key Stage. (Ofsted, 1999b, p. 23)

The 2012 framework is not clear. It does not spell out what it means by its few references to 'standards'. However, its interpretation becomes clearer when the fine detail of the accompanying evaluation schedule is examined.

As will be argued later in the article, the underlying conception of 'standards' is a much impoverished version which relates not to levels or descriptors across the full range of subjects in the curriculum but only to a sub-set and, within that sub-set, only to those amenable to testing/examination. In the evaluation schedule test/examination results are equated to 'standards' and 'standards' to 'achievement' – two very problematic contentions.

The Huge Significance Placed on Performance Data

According to the evaluation schedule, judging the achievement of pupils at the school requires inspectors to evaluate:

> - the standards attained by pupils by the time they leave the school, including their standards in reading, writing and mathematics and, in primary schools, pupils' attainment in reading by the end of Key Stage 1 and by the time they leave the school
> - how well pupils learn, the quality of their work in a range of subjects and the progress they have made since joining the school
> - how well pupils develop a range of skills, including reading, writing, communication and mathematical skills, and how well they apply these across the curriculum
> - how well disabled pupils and those who have special educational needs have achieved since joining the school
> - how well gaps are narrowing between the performance of different groups of pupils in the school and compared to all pupils nationally
> - how well pupils make progress relative to their starting points. (2012b, p. 6)

In both the framework and the evaluation schedule 'achievement' is equated quite explicitly with *academic* achievement – a very significant emphasis. This devalues other aspects of a school's work and of individuals' achievement in non-academic, but arguably equally important, areas. As used by Ofsted, 'achievement' is a hybrid amalgam of academic attainment, progress and the quality of learning.

In its evaluation schedule Ofsted provides detailed grade descriptors for the judgements to be made of the achievement of all pupils. The descriptor for

'good' achievement is used here as an example to try to tease out Ofsted's assumptions and to suggest problematic elements in that characterisation:

> *Good*
> Pupils are making better progress than all pupils nationally given their starting points. Groups of pupils, including disabled pupils and those with special educational needs, are also making better progress than similar groups of pupils nationally. Performance is likely to exceed floor targets. Pupils acquire knowledge quickly and are secure in their understanding in different subjects. They develop and apply a range of skills well, including reading, writing, communication and mathematical skills across the curriculum that will ensure they are well prepared for the next stage in their education, training or employment. The standards of attainment of the large majority of pupils are likely to be at least in line with national averages for all pupils. Where standards of any group of pupils are below those of all pupils nationally, the gaps are closing. Where attainment, including attainment in reading in primary schools, is low overall, it is improving at a faster rate than nationally over a sustained period and the gap is closing. (2012b, pp. 9-10)

Nowhere in this paragraph or elsewhere is the distinction between 'achievement' and 'attainment' made clear. Nor is the concept of 'progress' spelled out explicitly.

However, the underlying assumption is that all three can, indeed should, be defined in terms of performance data from test/examination results. Hence the references to 'floor targets', 'national averages', 'standards of attainment' and 'low attainment'. There is a further fundamental basic assumption, i.e. that complex phenomena such as learning and understanding 'must be reduced to simple elements accessible to quantitative measurement, without undue worry whether the specific characteristics of a complex phenomenon … may be lost in the process' (Koestler, 1989, p. 3).

Here and elsewhere, 'achievement' expressed in terms of performance data is treated as unproblematic and uncontentious. There is no hint that performance data are other than definitive, objective *measures* of performance made in relation to commonly agreed and interpreted criteria. Ofsted's uncritical use of performance data to characterise 'achievement', 'attainment' and 'progress' belies the fact that there are any number of well-documented controversies over whether, for example:

1. it is possible to *measure* (as opposed to *appraise*) performance in relation to whatever is meant by 'standards' or 'achievement';
2. it is possible through testing to assess 'connected' knowledge and understanding (as opposed to the 'thin' knowledge required to answer test questions)[3];

3. it is possible through successive testing to assess progress over time, especially in relation to conceptual understanding;
4. all subjects or just a 'core' should be tested or examined;
5. the most important aspects of particular subjects are, or can be, subject to testing or examining;
6. the particular tests or examining methods used can reliably and validly assess performance in relation to particular level descriptions or grade descriptors;
7. the particular tests and examinations can be marked, and the results reported, fairly and consistently;
8. the results of different tests and examinations can be compared over time;

and many other issues, some of which are alluded to, but dismissed, in the *Report on Key Stage 2 testing, assessment and accountability* (DfE, 2011a).

Ofsted's use of performance data, gathered directly from the reading checks for six-year-olds, from the tests used to inform teacher assessment at the end of Key Stage 1, from the tests at the end of Key Stage 2 and from external examinations at the end of Key Stages 4 and 5 means that 'achievement', 'attainment' and 'progress' are being characterised entirely in terms of those areas which the tests/examinations purport to measure.

Admittedly in the grade descriptor above there is one reference to 'pupils being secure in their understanding in different subjects' but what these subjects are and how that understanding is to be demonstrated as 'secure' are not indicated here or elsewhere in the three documents. In the outline guidance to inspectors related to the 'Achievement of pupils at the school' (Ofsted, 2012b, pp. 7-10) there is no explicit reference to specific non-tested subjects. In none of the documents does Ofsted offer its inspectors any specific guidance on the procedures and evidence needed to assess 'achievement' in non-tested/non-examined subjects. But why would it if the paramount source of evidence for 'achievement' is test/examination data? Ofsted thus seems to have very largely, perhaps entirely, abdicated responsibility for evaluating and reporting on 'achievement' or 'standards' in areas of the primary curriculum other than the tested elements of mathematics and English and in areas of the secondary curriculum which are not examined at the end of Key Stages 4 and 5.

Though test results are used as the predominant source of evidence for evaluating 'achievement', the evaluation schedule does acknowledge (perhaps a little grudgingly?) that 'inspectors should take account' [4] of

> evidence gathered during the course of the inspection on the learning and progress of different groups of pupils … drawn from
> - observation of lessons and other learning activities and discussions with staff and senior leaders
> - scrutiny of pupils' work to assess standards, progress and the quality of learning of pupils currently in the school
> - discussions with pupils about their work
> - parent, pupil and staff questionnaires

- case studies of individual pupils.
(2012b, p. 7)

However, like performance data, each of these has its limitations as a source of evidence for 'achievement'.

Lesson observation *can* provide evidence of pupils' current achievements but only when inspectors have experience of inspecting the same subjects in a wide range of other schools (including so-called 'outstanding' ones),when the level descriptors (or whatever constitute 'standards') are in the forefront of their minds in making their judgements and when through moderation/discussion with other colleagues their interpretations of these descriptors are broadly consistent. Given that current inspections exclude 'outstanding' schools, are of very short duration and rarely provide opportunities for moderation of judgements of observed achievement, these conditions rarely, if ever, obtain. In addition, as will be argued later , there are serious limits on the possibility of assessing aspects of learning and progress through classroom observation.

Of the other sources of evidence on which inspectors 'should draw', it is difficult to see how questionnaires could yield direct evidence of achievement, learning and progress. Case studies of individual pupils could be a useful, though very partial and time-consuming, source of evidence but from these it would not be possible to draw more than highly tentative generalisations. Discussion with pupils could well be a major source of data not only about their perceptions of their progress and learning but also about the genuine extent of their understanding, provided in-depth discussion with a representative range of different groups and ages of pupils is built into inspection schedules – but such circumstances will rarely, if ever, obtain within the limitations of the short inspections governed by the 2012 framework.

Equally problematic are judgements of achievement, progress and the quality of learning made over a number of years based on scrutiny of work from children in different year groups. When required to make judgements of the progress children make over the course of a key stage or from 'their starting points' to the time of the inspection, inspectors usually have to scrutinise samples of the work of different year groups of pupils collected in the same school year – not samples of work of the same year group of children collected over the whole period of a key stage or over the period since their 'starting points'. In such scrutiny; the different populations whose work is being sampled and examined over time are bound to vary from one another in a variety of ways, thus precluding direct judgements of progress. Even when it is possible to compare the work of the same year group over time, changes in the quality of work scrutinised might be due to factors which are no longer operative. Scrutiny of pupils' work within the constraints of an Ofsted inspection is usually conducted hurriedly and inevitably superficially due to time constraints. Such scrutiny is a very inadequate basis for judging the progress children have made, though it can yield useful evidence in relation to other issues such as the quality and consistency of marking.

Given the limitations of test data, the limitations of other sources of evidence and the lack of guidance to inspectors in assessing non-tested/non-examined subjects, it is inevitable that the inspection of the 'achievement of pupils at the school', itself seen only in restricted academic terms, is bound to be very partial, limited and subject to all kinds of reservations. Yet the pronouncements made about 'achievement' are likely to be reported with a degree of authority, certitude and precision that the inspection process does not permit. In particular it needs to be stressed that in primary schools that apparent degree of certainty comes from an unjustified reliance on performance data from only two subjects of the curriculum.

The Inspection of the Quality of Teaching

In the 2012 framework the quality of teaching is one of the four main areas on which inspection judgements focus. This emphasis is appropriate and uncontroversial in principle. There are, however, contentious issues related to the basis on which judgements of teaching quality are made, and especially the relationship between the quality of teaching and children's learning and progress.

The evaluation schedule accompanying the framework stresses that 'The judgement on the quality of teaching must take account of evidence of pupils' learning and progress' (2012b, p. 11) .Ofsted is in no doubt as to the tightness of the link between the two, but just how tight is it and how easy is it to make judgements of quality and progress? To answer it is necessary to distinguish between the *task* and *achievement* senses of both learning and teaching.

Learning

It is possible to distinguish between two senses of a concept like *learning*. The first is the *task* sense in which, for example, children in year 3 can be said to be learning aspects of place value but haven't yet understood these. If, however, they have understood them, then it can be said that they have learnt them in the *achievement* sense of learning.

The 2012 evaluation schedule includes criteria related to learning (and teaching) in both *task* and *achievement* senses. For example, inspectors are required to evaluate

- the extent to which well-judged teaching strategies, including setting challenging tasks matched to pupils' learning needs, successfully engage all pupils in their learning
- how well pupils understand how to improve their learning as a result of frequent, detailed and accurate feedback from teachers following assessment of their learning
- the extent to which the pace and depth of learning are maximised as a result of teachers' monitoring of learning during lessons and any consequent actions in response to pupils' feedback

> - the extent to which teachers enthuse, engage and motivate pupils to learn and foster their curiosity and enthusiasm for learning. (2011b, p. 11)

The *task* aspect relates to pupils' observable responses to the act of teaching. The evaluation schedule usefully draws attention to such questions as 'Are children engaged in the work?' 'Are they enthused?' 'How are they responding to the teacher's questions?' 'How are they contributing to discussion?' However, none of these – enthusiasm, engagement, concentration, active contribution – is *necessarily* indicative of learning taking place. For example, a child might exhibit all those qualities but be simply engaged in revising what he or she already knows. Nevertheless, it seems likely that when all or most of these qualities are in evidence, learning of some kind or other is going on. Secondly, there is bound to be an element of uncertainty when children's observable responses are being evaluated. For example, they may be feigning these qualities in order to impress the teacher or the inspector. Thirdly, inspectors are likely to vary somewhat in their interpretation of what constitutes appropriate 'engagement', 'response', 'contribution', etc. But with appropriate training, experience and discussion among inspectors it should be possible for their judgements on these aspects to be broadly harmonised and a reasonable degree of certainty of judgement achieved. Here the 2012 inspection framework is on reasonably secure ground; generally valid judgements *can* be made about children's observable responses to the teaching they receive.

The framework also requires inspectors to evaluate learning in the *achievement* sense by answering questions such as 'Do children understand the challenging tasks presented?' 'Do they understand how to improve their learning?' 'Has their learning been improved by frequent, detailed and accurate feedback?'

But how feasible is it to expect inspectors to be able to make such judgements? Here, the framework is on much less secure ground. Presumably, the judgements are to be based very largely on lesson observations. Certainly the 2012 framework stresses Ofsted's intention to 'increase the proportion of inspectors' time in school that is spent observing teaching and gathering evidence of *learning, progress* and behaviour' (2012a, p. 5, my italics).

Evaluating how well pupils have learned in a lesson involves at least two sets of judgements: (a) judgements of children's understanding and/or skills which they 'bring' to the lesson; and (b) judgements of the knowledge, understanding and skills they 'take away' at the conclusion of a lesson. Except in lessons involving, for example, the learning of specific physical competences or drama skills which children cannot perform at the beginning of a session but can demonstrate at the end, or those involving the learning of factual information which pupils do not know when tested at the start of the lesson but can demonstrate at its end, inspectors do not usually have the *detailed* knowledge of either (a) or (b) and therefore cannot gauge the degree of change in pupils' understanding as a result of an act of teaching. Of course, it is likely that most

pupils do learn something new in most lessons, though some sessions might quite appropriately be concerned with practising or applying previous learning. However, changes in conceptual understanding are not detectable to any significant degree through observation or brief discussion with pupils by inspectors. Conceptual learning in its 'achievement' sense might be able to be evaluated to a certain extent if inspectors had sufficient time to question closely an appropriate sample of pupils before and after a lesson and to take account of their oral, written and other responses to the teaching received. The Ofsted inspection framework and evaluation schedule do not refer to such close, in-depth questioning. How could they when most lesson observations are for no more than fifty minutes and often less? Davis (1999) concurs with the impossibility of assessing conceptual understanding through such short observations:

> If the lesson aspires to develop children's 'connected' understanding, this cannot be assessed by means of the limited evidence available to the most perceptive of inspectors. It would be difficult even to check whether just one pupil had learned in this rich fashion as a result of the teaching observed. (p. 36)

But if assessing learning in lessons is problematic, can scrutiny of children's work provide good evidence of progress in learning? Work scrutiny *can* yield valuable information about the kinds of tasks set, about the nature of teachers' written feedback and about pupils' response to that feedback but for reasons discussed earlier it is a very problematic source of evidence for progress in learning in its *achievement* sense.

Teaching

As with learning, it is important to distinguish between two senses of teaching. To say 'A teacher taught aspects of place value to children in year 3' could mean she was *attempting* to teach place value (the *task sense*) without implying that she was successful, or it could mean that she *was* successful in teaching those aspects to that particular group of children (the *achievement* sense). Ofsted's criteria for judging how well teachers teach, quoted at the beginning of the previous section, involve the use of *teaching* in both senses.

The *task* sense of teaching refers to the way in which lessons are conducted. The criteria quoted above refer, for example, to whether teachers display high expectations; set challenging tasks; monitor ongoing learning; give appropriate feedback; enthuse and motivate children, etc. It seems reasonable to argue that the presence of such features is likely to enhance the possibility of the teaching resulting in the intended learning. Through classroom observation inspectors can gain sufficient evidence to be able to make justifiable judgements of teaching quality in this *task* sense. Of course, even here there is likely to be an element of unreliability of judgements between inspectors but such unreliability can be kept within acceptable limits by discussion and moderation by inspectors

themselves. Evaluating the quality of the teaching in the *task* sense (and making suggestions as to how that teaching might be improved) is, or perhaps should be, at the very centre of inspectors' expertise.

The judgements to be made also involve criteria related to the *effectiveness* of the teaching in bringing about learning, i.e. judgements about teaching in the *achievement* sense. In order to make such a judgement inspectors have to have gathered reliable evidence of the learning that has been fostered by the teaching observed in those lessons.[5] If the arguments advanced earlier about the evaluation of learning in the *achievement* sense are accepted, then judgements about conceptual learning and understanding are simply not possible, without extensive discussion with children. This is not a problem which Ofsted could easily solve by modifying its criteria or making minor changes to its procedures; it would require a very different, time-consuming, research-oriented methodology based on in-depth interviewing of children before and after lessons and observing them during lessons. As with the *achievement* sense of learning the Ofsted framework is on insecure ground when requiring inspectors to make judgements of the quality of teaching in its *achievement* sense, except in relation to observable skills and the recall of factual information.

Teaching, Learning and Performance Data

What are inspectors most likely to use as the basis of judgments about the quality of teaching (in its *achievement* sense)? The evaluation schedule provides the answer in the first sentences of the grade descriptors for 'outstanding', 'good' and 'satisfactory' teaching:

> *Outstanding*[6]
> Much of the teaching in all key stages and most subjects is outstanding and never less than consistently good. As a result, almost all pupils are making rapid and sustained progress. (2012b, p. 12)
> *Good*
> As a result of teaching that is mainly good, with examples of outstanding teaching, most pupils and groups of pupils, including disabled pupils and those who have special educational needs, are achieving well over time. (2012b, p. 13)
> *Satisfactory*
> Teaching results in most pupils, and groups of pupils, currently in the school making progress that is broadly in line with that made by pupils nationally with similar starting points. (2012b, p. 13)

The prime determinant for those judgements is 'achievement' (yet again!) – 'achievement' once more expressed in terms of performance data over time set against national benchmarks as indicated in RAISE online. This raises the important issue as to whether and to what extent judgements about the quality of learning and teaching can, or ought to, be made independently of judgements of pupils' 'achievement' as measured by test/examination data.

There is a very strong likelihood that instead of classroom observation forming the basis for judging teaching quality, inspectors' prior knowledge of performance data will directly influence their judgements both of teaching in individual lessons and in the school as a whole. For example, how likely is it that inspectors will judge the quality of a school's teaching (in the *achievement* sense) to be inadequate if performance data are well above average? How likely is it that they will judge teaching quality to be good if the data are well below average? There is a grave danger that the judgement of teaching quality will, consciously or unconsciously, be unduly reliant on performance data rather than on classroom observation. If, as a result of the 2012 framework and evaluation schedule, too many judgements of the quality of teaching (in the *achievement* sense) are not made independently of judgements of children's 'achievement' expressed in terms of test scores, then justice is unlikely to be done to the quality of teaching of those children whose progress is not deemed 'broadly in line with that made by pupils nationally and with similar starting points' (2012b, p. 13).

The Inspection of the Behaviour and Safety of Pupils

The 2012 Framework places the behaviour and safety of pupils as the third of the areas on which inspection should focus. These are important to parents, children and teachers and need to be inspected closely if there are concerns over any shortcomings. However, it is debatable whether these issues merit the high status they are being accorded in the 2012 documentation, given that under previous frameworks the large majority of schools have been given positive, or very positive, evaluations on behaviour and safety. There is *no* inspection evidence to suggest that in these areas standards have declined nationally or are so problematic as to require this renewed, major focus under the new inspection arrangements. There *is,* however, a likelihood that in giving them the priority it does, Ofsted may be reacting to political and tabloid pressure.

The accompanying evaluation schedule requires inspectors to evaluate:

- pupils' attitudes to learning and conduct in lessons in and around the school
- pupils' behaviour towards, and respect for, other young people and adults …
- how well teachers manage the behaviour and expectations of pupils to ensure that all pupils have an equal and fair chance to thrive and learn in an atmosphere of respect and dignity
- pupils' ability to assess and manage risk appropriately and keep themselves safe
- pupils' attendance and punctuality at school and in lessons
- how well the school ensures the systematic and consistent management of behaviour.

(2011b, p. 15)

Here the framework and evaluation schedule are on much firmer, far more defensible ground. It *is* possible for experienced inspectors to make justifiable judgements on children's behaviour, on their attitudes to learning, on the management of behaviour, on attendance rates and on the incidence of lateness – gathered from their own observations, school documentation and the views of children, parents, carers and staff. It is less easy, though still possible, through discussion and observation to come to a tentative assessment of children's ability to manage risk and keep safe. Perhaps the only element of unreality in the criteria is the requirement on inspectors to judge whether teachers' management of behaviour *ensures* that *all* pupils have an equal and fair chance to thrive and learn. Inspectors cannot possibly judge whether any action can *ensure* future fairness nor how any element of behaviour management can possibly impact on each and every pupil.

However, in contrast to the other three foci of the framework, this one does not require inspectors to make evaluative judgements with performance data directly or indirectly in mind.

The Inspection of the Quality of Leadership and Management

The quality of leadership and management is the last of the four key areas. That school inspection should judge such quality is uncontroversial, but as with teaching and learning there are contentious issues around the criteria inspectors have to apply. These require inspectors to evaluate whether the school's leadership:

> - demonstrate an ambitious vision for the school and high expectations for what every pupil and teacher can achieve, and set high standards for quality and performance
> - improve teaching and learning including the management of pupils' behaviour
> - provide a broad and balanced curriculum: that meets the needs of all pupils, enables all pupils to achieve their full educational potential and make progress in their learning; and promotes their good behaviour and safety and their spiritual, moral, social and cultural development
> - evaluate the school's strengths and weaknesses and use their findings to promote improvement
> - improve the school and develop its capacity for sustaining improvement by developing leadership capacity and high professional standards among all staff;
> - engage with parents and carers in supporting pupils' achievement, behaviour and safety and their spiritual, moral, social and cultural development
> - ensure that all pupils are safe.
> (2012b, p. 18)

Some of these criteria (the second, fourth and fifth) are appropriate and uncontentious in principle, though they are inevitably susceptible to subtly different interpretations (for example, what constitutes 'improvement'?) and not easy (though *not* impossible) to apply to management and leadership practice.

Some of the others (the first, third and sixth) either feature 'achievement' and 'performance' explicitly or smuggle in these notions through phrases such as 'make progress in their learning'. This would be uncontentious if Ofsted's perspective on 'achievement ' was uncontentious. But it is not. As argued above, the inspection of the 'achievement of pupils at the school', viewed almost entirely in academic terms and in primary schools focused on the two subjects of mathematics and English, is partial, limited and subject to all kinds of reservations.

There is a very real danger that in applying the first, third and sixth of these criteria inspectors will begin with examining performance data and infer from that data the effectiveness or otherwise of management and leadership practice rather than forming independent judgements of that practice and then exploring the relationship between that practice and the achievement of children broadly conceived. That exploration could, in some cases, reveal interesting, perhaps uncomfortable, relationships. Good management does not necessarily lead to good or high achievement on the part of children, especially if achievement is narrowly construed. Likewise, high test performance does not necessarily correlate with good management practice. Lying behind the development of these particular inspection criteria is the assumption that managment and leadership can be, indeed should be, held *directly* and *almost totally* accountable for pupil performance however defined. This is to ignore the many non-school factors affecting children's learning and progress, over which the school and its leadership have little or no control or influence. Unless judgements of management practice are made separately from those about children's performance there is a real danger of injustice, especially to the leadership of schools in areas of disadvantage where 'achievement' is likely to be below average.

The third and seventh of the criteria are impossible for inspectors to judge and equally impossible for schools to achieve. To take the third criterion, it is impossible to identify every one of the needs of each and every pupil, even if agreement could be found on what those value-laden 'needs' are. It is impossible for a curriculum, however 'broad and balanced' (and these too are value-laden terms) to enable each and every pupil to achieve 'their full educational potential' – a much-used and much-abused notion which assumes a fixed, determinate, yet distantly achievable, quantum of whatever makes up 'educational potential'. To take the seventh criterion, no manager can '*ensure* that all pupils are safe', though their actions to promote safety can be, and should be, judged. The sixth criterion makes reference to aspects of children's development whose problematic inspection is discussed briefly in a later section of this article. In relation to leadership and management, as with other aspects of the 2012

framework, Ofsted expects more of schools, more of its inspectors and more of itself than they can possible deliver.

Grading: a brief note

For each of the four main areas and for the judgement of a school's overall effectiveness inspectors are required use a four-point grading scale: 1 (outstanding), 2 (good), 3 (satisfactory)[7] and 4 (inadequate). For each of these grades in each of these areas grade descriptors are provided. Putting to one side problems with the language in which they have to be expressed and the very real difficulty of matching the complex provision of a school to generalised descriptors, the wording of the grades themselves presents difficulties. Three are criterion-referenced, qualitative terms but one ('outstanding') is norm-referenced. To be consistent 'outstanding' needs to be replaced by a criterion-referenced qualitative term such as 'excellent'. Over time the grade 'satisfactory' has been devalued; it no longer carries the connotation of 'good enough' (as it did in old-style HMI inspections) but instead is now taken to mean its opposite, i.e. not good enough! Getting rid of this third grade altogether seems unwise, as it roughly matches provision in some schools. Perhaps grade 3 might be described as 'Not consistently good', thus allowing for the likelihood that there are aspects of good practice within such schools. 'Inadequate' has condescending, highly derogatory connotations and might more appropriately be replaced by 'unsatisfactory' until that term in turn is devalued.

Other Aspects of the Inspection Framework

The 2012 framework makes it clear that in addition to the four main foci of the inspection process inspectors must also consider

> - the spiritual, moral, social and cultural development of pupils at the school
> - the extent to which the education provided by the school meets the needs of the range of pupils at the school and, in particular, the needs of disabled pupils and those who have special educational needs. (p. 5)

It is unclear whether the reference to 'also consider' reflects the secondary importance of these aspects compared to 'those aspects of schools' work that have the greatest impact on raising achievement' (2012a, p. 5). The fact that the evaluation schedule provides no specific guidance on how these judgements are to be arrived at gives credence to the assumption that they are of secondary significance – perhaps their inclusion as a form of rhetorical window-dressing? It is certainly possible to argue that, if inspectable, they are important aspects which need to be considered when judging the effectiveness of a school. However, there are severe, indeed insuperable, limitations on inspectors' ability

to make sound judgements about aspects of them, as they are currently expressed.

Reservations about identifying, let alone meeting, the 'specific needs' of 'all' or 'the range of pupils at the school' have already been briefly alluded to and need not be rehearsed again.

There is, however, an interesting ambiguity over the meaning to be accorded to the requirement for inspectors to 'consider the spiritual, moral, social and cultural development of pupils at the school'. This could be interpreted as asking inspectors to judge the adequacy or otherwise of the *provision* the school makes, both through its curriculum and its wider life, in promoting these fourfold aspects of pupils' personal development. Arguably, judgements could be made relatively straighforwardly about the quality of the experience the school offers in relation to children's social and moral development, but less obviously in respect of cultural development and even less clearly, and more contentiously, in respect of spiritual development (is the latter capable of conceptual clarification and amenable to value consensus by inspectors and schools?).

However, the documentation implies a much more contentious interpretation. The subsidiary guidance requires inspectors to 'gather evidence of the *impact* of the curriculum on developing aspects of the pupils' spiritual, moral, social and cultural (SMSC) development' (p. 19, my italics). This would require inspectors to come to a view of the state of children's development in the four aspects and of how far the school has promoted these. Such requirements would be empirically and conceptually impossible to meet. What evidence could count towards the judgements? How could it be collected and validated? How could inspectors *judge* the state of others' personal development? How in particular could fallible human beings possibly judge others' spiritual development? Not even the chief inspector can play God, never mind his subordinates!

Ofsted needs to take steps to clarify the meaning and the wording in relation to the inspection of pupils' personal development and of the school's contribution to meeting the 'needs of the range' of its pupils.

The Inspection Framework and Overall Effectiveness

The evaluation schedule requires a final, overall judgement of the school's effectiveness. Inspectors are required to evaluate 'the quality of the education provided in the school' by drawing on their judgements of the four areas (achievement, teaching, behaviour/safety and leadership/management) and taking into consideration the two broad aspects discussed in the previous section. In principle this seems a sensible, uncontentious requirement – but only if the framework and schedule used have no serious limitations and only if all key aspects of the school have been considered. This article argues that these conditions have not been met.

The framework and schedule employ an impoverished view of what constitutes 'achievement' focused almost entirely on performance on tested/examined subjects and based on quantitative data. They claim too much for inspectors' ability to judge teaching, learning and progress in the 'achievement sense', though they make appropriate claims for these aspects in the 'task sense'. Also, too tight a relationship is implied between teaching and learning and the documentation's impoverished view of 'achievement'. The framework and schedule are on much more defensible ground in the judgements they require in relation to children's behaviour and safety. In contrast, they tie judgements of leadership and management of both primary and secondary schools too tightly to pupils' achievement and require judgements in relation to some of their criteria that are impossible to meet. The severe limitations on inspectors' ability to make sound judgements about aspects of personal development and about the school's ability to meet the needs of the full range of pupils are not acknowledged. The school's curriculum, its main vehicle for the transmission of understanding and values, is not given explicit attention as one of the key foci. Finally, the framework and schedule require inspectors to make a summary judgement of a school's overall effectiveness in the absence of explicit aims and values. In summary, the overall judgement required of inspectors in terms of the effectiveness of a school is seriously compromised.

The same can be said of the overall effectiveness of the 2012 inspection framework itself and its accompanying evaluation schedule. The two documents along with the supplementary guidance often claim more than they can justifiably deliver. The judgements they require are more partial, more uncertain, more limited, and more dependent on problematic performance data than Ofsted publicly acknowledges.[8] A fundamental reappraisal of the premises on which both framework and schedule are based is needed if Ofsted inspection is to be 'fit for purpose'. This article is offered as the beginning of such a reappraisal.

Appendix: 'outstanding' nonsense?

The chief inspector believes all 'outstanding' schools should have 'outstanding' teaching – a challenging but certainly justifiable stance. Ofsted believes that 'outstanding' teaching should meet a range of demanding criteria – equally challenging and justifiable. What is unjustifiable is that inspectors should or could use those criteria to identify 'outstanding teaching' and so give a school that badge of honour.

Ofsted argues that 'the main evidence [for the quality of teaching] will come from inspectors' direct observations of teaching and learning and their discussions of what they have seen with teachers, other adults and pupils' (2012b, p. 12). The criteria used by inspectors to judge 'outstanding' teaching bear examination. Here three examples are highlighted, though others could be cited.

The first criterion is that 'all teachers have consistently high expectations of all pupils'. For that criterion to be met every single teacher, for example in a 1200-pupil school, would have to be observed a number of times in a variety of teaching situations, their expectations of each and every one of their pupils elicited and these then judged 'high' or otherwise by an omniscient inspector. It is simply impossible.

A second: 'teachers use well-judged and often imaginative teaching strategies that … match individual needs accurately'. How can inspectors as outsiders in a class for 30 minutes or in a school for two days in all possibly know what the individual needs are of every pupil to judge whether teachers are matching strategies to needs accurately? They cannot possibly do this.

A third: 'almost all pupils are making rapid and sustained progress'. How can progress, especially progress in understanding, be judged in an observation of 30 minutes or less? It cannot. How in a short inspection can inspectors possibly know what progress, if any, is being made by, say, 27 out of 30 pupils in a class or in the whole school context by, say, 975 out of 1000 pupils? They cannot.

Many, though not all, of Ofsted's criteria for 'outstanding teaching' are equally impossible to apply. The problem is not with the criteria *per se;* they do embody teaching excellence. After all, everyone involved in education wants teachers with high expectations and with imaginative strategies that as far as possible meet pupils' needs; all of us want pupils to make as rapid and sustained progress as possible. And so on for the other criteria.

There are, however, two problems. One is that many of the criteria cannot be applied in a literal sense in any inspection. The second is that if they could be applied fully no teacher or school could ever meet the impossibly high standards expected consistently day in, day out. Ofsted's criteria are outstanding nonsense.

Update (July 2012)

A revised framework for inspection and an accompanying evaluation schedule have recently been published as part of the arrangements for school inspection from September 2012 onwards. Some of the minor criticisms levelled in this article about the original 2012 framework and schedule have been addressed. However, that is not the case with the fundamental criticisms made here. Though there is a brief acknowledgement that 'achievement' includes broader aspects of the life of the school, the new documentation continues to stress an impoverished view of what constitutes 'achievement', still focusing almost entirely on tested subjects. It continues to ask too much of its inspectors in judging teaching, learning and progress within the confines of such a short space of time, whether that be a single lesson, a single half-lesson or the period of the inspection itself. It ignores the difficult issue of what values can or should inform the inspection process and the qualitative, value-laden judgements which ensue.

Notes

[1] This article's main points can also be applied, with modifications, to the inspection of nursery schools, classes and other early years settings.

[2] The article also makes some reference to a third document: *Subsidiary Guidance: supporting the inspection of maintained schools and academies from January 2012.*

[3] For these important distinctions see Davis's pamphlet (1999) *Educational Assessment: a critique of current policy.*

[4] Interestingly the draft evaluation schedule originally stated that 'inspectors *may* take account of such evidence', implying no obligation on inspectors so to do (my italics). That at least has been rectified.

[5] Of course some lessons are appropriately not intended to foster 'new learning' but to practise and/or apply previous learning. The evaluation schedule does not recognise this.

[6] See Appendix for further problematic elements involved in the inspection of 'outstanding teaching'.

[7] Having devalued the term 'satisfactory', Ofsted, through its new chief inspector, is replacing it by 'requires improvement'. This has three disadvantages. Firstly, it is true of all schools including so-called 'outstanding' as well as 'inadequate ones'; all can and should improve. Secondly, it fails to recognise that there are likely to be pockets of good practice in every satisfactory school .Thirdly, whatever the new designation, it will be devalued in turn, as will grade 2 ('good') and grade 1 ('outstanding') before too long.

[8] Similar criticisms can be offered of Ofsted's inspection of early years provision – perhaps even more so as too often data are required on the progress of 3- and 4-year-olds whose development is very complex and uneven and not amenable to quantitative assessment.

References

Alexander, R. et al (2010) *Children, their World, their Education: final report and recommendations of the Cambridge Primary Review.* London: Routledge.

Bowie, M. (1998) *Proust among the Stars.* New York: Columbia University Press.

Bramall, S. & White, J. (2000) *Will the New National Curriculum Live Up to its Aims?* Impact No. 6. London: Philosophy of Education Society of Great Britain.

Davis, A. (1999) *Educational Assessment: a critique of current policy.* Impact No. 1. London: Philosophy of Education Society of Great Britain.

Department for Education (2011a) *Report on Key Stage 2: testing, assessment and accountability.* London: The Stationery Office.

Department for Education (2011b) *The Framework for the National Curriculum: a report by the expert panel for the National Curriculum Review.* London: The Stationery Office.

Department for Education and Employment (DfEE)/Qualifications and Curriculum Authority (QCA) (1999a) *The National Curriculum Handbook for Primary Teachers in England.* London: DfEE/QCA.

Department for Education and Employment (DfEE)/Qualifications and Curriculum Authority (QCA) (1999b) *The National Curriculum Handbook for Secondary Teachers in England.* London: DfEE/QCA.

Koestler, A. (1989) *The Ghost in the Machine.* London: Arkana Books.

Office for Standards in Education (Ofsted) (1999a) *Inspecting Schools: the framework.* London: Ofsted.

Office for Standards in Education (Ofsted) (1999b) *Handbook for Inspecting Primary and Nursery Schools.* London: The Stationery Office.

Office for Standards in Education (2012a) *The Framework for School Inspection.* London: Ofsted.

Office for Standards in Education (2012b) *The Evaluation Schedule for the Inspection of Maintained Schools and Academies.* London: Ofsted.

Office for Standards in Education (2012c) *Subsidiary Guidance: supporting the inspection of maintained schools and academies from January 2012.* London: Ofsted.

Qualifications and Curriculum Authority (QCA) (2007) *The National Curriculum at Key Stages 3 and 4.* London: QCA.

Richards, C. (2001) *School Inspection in England: a re-appraisal.* Impact 9. London: Philosophy of Education Society of Great Britain.

COLIN RICHARDS is Emeritus Professor of Education at the University of Cumbria, and a sub-editor and contributory author to the *Cambridge Review of Primary Education.* He has been an honorary professor at the University of Warwick and a visiting professor at the universities of Leicester and Newcastle. He is chair of governors of a Cumbrian secondary school which has twice decided not to pursue academy status. As a critic of much (though not all) of government education policy (both New Labour and Coalition) he is a frequent contributor to the national press, particularly to the *Education Journal* and to *The Times Educational Supplement* where he has had two hundred articles and letters published. He treasures the epithet 'an old-fashioned HMI', bestowed by a former Chief Inspector of Schools.
Correspondence: colin@sparkbridge.freeserve.co.uk