



Viewpoint: Thoughts on whole school inspection post-Covid

Colin Richards

How to cite this article

Richards, C. (2020) 'Viewpoint: Thoughts on whole school inspection post-Covid'. *London Review of Education*, 18 (3), 511–18. <https://doi.org/10.14324/LRE.18.3.13>

Submission date: 14 July 2020

Acceptance date: 31 July 2020

Publication date: 13 November 2020

Peer review

This article has been through editorial review.

Copyright

© 2020 Richards. This is an open-access article distributed under the terms of the Creative Commons Attribution Licence (CC BY) 4.0 <https://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Open access

The *London Review of Education* is a peer-reviewed open-access journal.

Viewpoint: Thoughts on whole school inspection post-Covid

Colin Richards* – *University of Cumbria, UK*

Abstract

This article argues that the pause in whole school Ofsted inspections until the end of 2020, and possibly into 2021, offers the opportunity to re-examine the nature of school inspection as an appreciative process with inevitable limitations that Ofsted has so far failed to recognize. This raises the possibility of replacing the current adversarial system with a more consensual one grounded in basic principles similar to, but developed from, the process of school inspection pre-Ofsted by Her Majesty's Inspectorate of Schools.

Keywords: inspection, appreciative judgement, reliability, validity, quality of teaching, curriculum

The pause in whole school inspections consequent on the coronavirus emergency should be used to re-examine the principles underlying the process, and to institute reforms to policy and practice. Nothing less is required if the current confrontational stand-off between schools and the Office for Standards in Education, Children's Services and Skills (Ofsted) is to end and a consensual approach to inspection is to be reinstated.

What follows is a personal reinterpretation and updating of an approach to school inspection undertaken by Her Majesty's Inspectorate of Schools before its abolition in 1992 (Richards, 2016). It is offered as a set of principles and insights that could be used to underpin a redesigned inspection system post-Covid, which takes more cognizance of the subtleties and uncertainties involved in trying to assess the quality of education in a timely and sensitive manner. It offers no neat, detailed prescriptive blueprint to be readily adopted by Ofsted (or any successor body), but proposes a set of principles underlying an inspection system defensible in terms of the possibilities and limitations involved in the exercise of professional judgement in an educational context.

The nature of professional judgement

More so than in previous Ofsted formulations, any renewed approach to school inspection needs to recognize and make central the exercise of professional judgement. That is easy to say but very difficult to characterize; judgement is easier to appreciate than to define. Ofsted has made no attempt to characterize it, despite making sporadic references to its importance. Vickers (1983) gets close to it in his book *The Art of Judgment*. He usefully distinguishes two aspects of the kind of overall appreciative judgement that school inspectors have to make – reality judgements involving facts about the state of any system such as a school being inspected, and value judgements 'making judgments about the significance of these facts' (ibid.: 40).

Note that Vickers characterizes appreciative judgement as an 'art', not as a science, nor as an art informed by science – as in recent Ofsted publications from its 'research' department.

In an inspection, reality judgements are derived from observations and discussions in class, with leaders and around a school. Such judgements can seem to be firmly rooted in objective reality, but crucially they can only be mediated through inspectors' past experience, and they involve mental processes that are often complex and prolonged, resulting in inferences, forecasts and conclusions. Different inspectors may legitimately observe, report and assess facts differently. Reality judgements cannot be characterized as totally objective or be regarded as incontestable.

Similarly, the judgements that inspectors inevitably have to make about the value of what they observe or deduce 'cannot be *proved* correct or incorrect; they can only be *approved* as right or *condemned* as wrong by the exercise of another value judgment' (ibid.: 71, my emphasis) – itself inevitably contestable. In any reconceptualization of inspection post-Covid, the notion of objectivity needs to be replaced by that of 'value-informed judgement'. All this implies that inspection cannot and should not claim to be any more than the professional subjective judgement of a group of experienced, expert observers. As such, the findings of any inspection are contestable, open to interpretation and never definitive. Just as Covid has challenged many of our previous health-related certainties, so post-Covid inspections should respect the uncertainties of the judgement process.

To minimize those inevitable uncertainties, whole school inspections need to rely on the collective, not the individual, judgement and experience of the inspectors. As Vickers (1983) stressed, appreciative judgement and subsequent decision making, although they are mental activities of individuals, are part of a social process. They are taken within, and depend upon, a net of communication, which is meaningful only through a vast, partly organized accumulation of largely shared assumptions and expectations, a structure constantly being developed and changed by the activities it mediates. In the case of inspection, such collective judgement making needs to be based on participants' wide experience of a variety of institutions in different educational contexts nationwide. It is metaphorically 'forged' or 'hammered out' through lengthy discussion and deliberation with other similarly experienced colleagues. The result is a collective, unique but internally moderated set of judgements – an appreciation, not a set of off-the-shelf ones imported from elsewhere. The notion of collective 'hammered out' appreciative judgement is crucial. No published report, especially one determining or affecting a school's future, should be the work of one individual alone. Even in the smallest school, an individual's views need to be moderated with the perspectives of at least one other to arrive at a defensible, moderated judgement.

Practising professional judgement

At the risk of appearing pretentious, inspection is best characterized as a form of joint educational connoisseurship, not bound by clear-cut, straightforward, incontestable criteria. In considering how an 'expert connoisseur' makes aesthetic judgements, Wittgenstein gets close to helping us understand the nature of inspection judgements and how they are justified. He commented:

We learn certain things only through long experience and not from a course in school. How, for instance, does one develop the eye of a connoisseur? Someone says, for example, 'This picture was not painted by such-and-such a master'. He may not be able to give any good reasons

for his verdict. How did he learn it? Could someone have taught him? Yes – not in the same way as one learns to calculate. A great deal of experience was necessary. That is, the learner probably had to look at and compare a large number of pictures by various masters again and again. In doing this he could have been given hints. Well, that was the process of learning. But then he looked at a picture and made a judgment about it. In most cases he was able to list his reasons for his judgment, but generally it wasn't they that were convincing ... The value of the evidence varies with the experience and the knowledge of the person providing it, and this is more or less the *only* way of weighing such evidence since it cannot be evaluated by appeal to any system of general principles or universal law. (Quoted in Monk, 2005: 103–4)

Applying these insights to inspection implies that professional expertise cannot be acquired simply from 'a course' of professional development. It involves learning from a wide range of teaching and inspection experience in a variety of relevant contexts, national as well as regional or local. It involves looking at, and comparing, a large number of educational activities by 'various masters again and again'. It is not like learning from an inspection rule book or tick list. It involves learning from others more experienced in making judgements of teaching quality, who can 'hint' at what is required and who can discuss the complexities and intangibles of classroom observation. Like Wittgenstein's connoisseurs, inspectors should be able to 'list reasons' for their judgements, but these can never be absolutely 'convincing', given the difficulties involved in interpreting learning and teaching. The value of the judgements, and the evidence they use to back them up, depends upon the experience and knowledge of the person making them. To repeat Wittgenstein's comment, 'this is more or less the *only* way of weighing such evidence since it cannot be evaluated by appeal to any system of general principles or universal laws' (ibid.) – whether these are enshrined in an inspection handbook or in subsidiary guidance.

The reliability and validity of inspection judgements

One of the major criticisms levelled at inspection judgements is their lack of reliability and validity – concepts central to the literature of assessment/measurement. Teachers' professional organizations, professional interest groups such as the Headteachers' Roundtable (2019) and academics such as Coffield (2017) offer this critique. Ofsted has responded by claiming to use social science research methods to improve the validity and reliability of its judgements. But can notions of validity and reliability be straightforwardly applied to the making and justification of such judgements?

Examining the basis of theatre criticism, in many ways analogous to school inspection, can help answer that question. Theatre critics appraise a performance or run of performances, as school inspectors appraise schools, based on a series of observations. Critics judge the quality of the acting; likewise, inspectors judge the quality of teaching. Critics judge how far the performance reflects the content and intentions of the play text; similarly, under the current Ofsted framework, inspectors comment on the rationale and implementation of the 'text' of the curriculum. Critics assess the reactions of the audience; likewise, inspectors assess students' responses.

Critics judge the quality of what they see. So do inspectors. They do not measure what they see on any numerical scale, nor can inspectors. Critics make their judgements based largely on their experience of similar, although never identical, productions; likewise, their inspectorial counterparts. The criteria that theatre critics

use are largely intuitive and impressionistic, and they cannot be reduced to a checklist of clear, unambiguous components.

It does not make sense to ask of theatre criticism that it be reliable and valid in the ways that such terms are usually used in educational assessment. The same applies to school inspection. Both are value-laden enterprises with the concept of 'quality' at their heart, and thus are subject to a different kind of assessment logic than educational measurement. Academic and professional critics need to recognize this, as do those currently working within Ofsted's research division and trying (vainly) to bolster its 'scientific' credentials.

Some limitations of inspection judgements

Because of the complex mix of reality and value judgements involved in the act of educational connoisseurship, an inspection team can never claim that their interpretation of a school is the only correct one. Nor should inspectors ever claim a monopoly of objective, authoritative judgement. Equally importantly, that unique set of judgements cannot be directly or robustly compared with the equally unique set of judgements of a school in a different context, or even with the judgements of the same school (which never remains 'the same school') inspected at a different time. Each set of inspection judgements is in a sense *sui generis*. Direct comparison of inspection judgements over time, or from school to school, is at best highly problematic and at worst invalid.

With their focus on observation and discussion, inspectors can only validly report and interpret activities seen at a particular point in time – a 'snapshot'. They cannot comment with any plausibility on what has happened in the past or predict what will happen in the future. They cannot comment with any authority or conviction about progress over time, whether by groups of students or by the school as a whole, since inevitably they do not have first-hand access to the past. Admittedly, they may have past documentation or a past inspection report to refer to; but they do not have full access to their predecessors' assumptions, expectations or deliberations for comparison, nor can they know with any certainty what has transpired in the interval between inspections. Performance data from the past may be available, but such data are fallible, contestable, variously interpretable and very partial as indicators. They cannot be interpreted except in the light of close knowledge of the context in which they were generated, and this is denied to the inspectors visiting and reporting some time later. The judgements that inspectors make can only be as 'they seemed to them at the time'. Every inspection report is inevitably to some extent out-of-date immediately after the inspection, but that does not mean it is not useful in the short-to-medium term as a basis for professional reflection and development. It can, in fact, be very valuable as an 'outside' appreciation of a school's work, including aspects requiring consideration, as well as those considered praiseworthy by the inspection team. The time-specific 'instant' nature of inspection judgements, and their inability to comment meaningfully on progress, whether by the school or by its students, need to be more fully recognized in any re-valuation of inspection policy and practice.

Inspecting the quality of teaching

The heart of inspection is a professional judgement about the quality of teaching – as recognized in statute when Ofsted was established in 1992. Under the most

recent inspection framework, its centrality has been downgraded so that it is now seen essentially as a vehicle for delivering the content of the curriculum. In reality, it is far more than that; it is also very importantly the medium through which aspirations, interpersonal and cultural values, expectations and attitudes to learning are fostered.

In recent years, teachers' professional associations have expressed anxieties over inspectors' preferred teaching methods influencing their judgements. However, evaluating the quality of teaching need not – should not – involve looking for particular teaching methods and then gauging their effectiveness in terms of promoting learning; rather the reverse. Inspectors should look for evidence of students' learning in terms of their observable responses to teaching, and then work back to highlight those factors that may have promoted, or hindered, their learning. This involves close observation and discussion with both staff and students. The unanticipated success of the wrong method, as judged by students' responses to the teaching they receive, needs to be recognized and celebrated. Similarly, the unanticipated failure of the right method needs acknowledgement in inspection reports.

Judgements about the quality of teaching in lessons, and in the school as a whole, are properly tentative, and consequently they should be offered as such in any feedback to those whose work has been observed. There is inevitably a considerable degree of inference involved in feedback, especially about the extent to which learning has taken place. There is also inevitably an element of professional judgement as to which features of a lesson have contributed to, or inhibited, learning. That tentativeness is crucial to the context in which any feedback is being given. It offers the opportunity in dialogue for other tentative, evidence-based, interpretations to be offered by the teacher who has been observed. Downgrading the making of judgements about the quality of teaching in a school endangers the professional credibility of inspection, and threatens to leave performance data as the main, or even sole, source of evidence used in reporting on the quality of education.

Inspecting the curriculum

Until very recently, the curriculum has not been a major focus of attention in school inspections. In contrast, it is currently seen as the core, or heart, of what Ofsted (2019) terms 'the quality of education'. Its *Education Inspection Framework* (ibid.) focuses on three aspects: *intent*, *implementation* and *impact*. These are to be assessed as a result of discussion with school and subject leaders, scrutiny of documentation and observation of work in class. These are certainly important aspects of curriculum planning and management, and they can be inspected with varying degrees of plausibility and certainty. *Intent* is probably the easiest to characterize in general terms, although deciding on whether the curriculum is 'ambitious' enough, or whether it is 'coherently planned and sequenced towards cumulatively sufficient knowledge and skills for future learning' (ibid.: 90) is far from straightforward and inevitably shot through with value judgements that are far from uncontentious. *Implementation* involves appropriate sampling of lessons by inspectors, with activities observed being matched with written or oral expressions of intent – a tricky but not impossible undertaking, provided that inspectors have the necessary time and subject expertise. *Impact* is the most uncertain to judge, given the restricted timescale of a typical inspection. Talking with students about their progress could be a major source of evidence, provided that lengthy, in-depth discussions with a representative range of different groups and ages are built into inspection schedules, but these are

all but impossible within the time constraints of most whole school inspections. Likewise, in-depth examination of students' work, cross-referenced to evidence of *intent* and *implementation*, may be possible in theory, but it is very problematic and impressionistic in practice.

Unacknowledged by Ofsted, there is a major and fundamental lacuna in its approach to the school curriculum. It focuses entirely on how well the curriculum is planned and implemented, and on its apparent impact. It does not allow evaluation of the *worthwhileness* of what has been designed and implemented. It assumes that the current legally mandated national curriculum framework is both good and incontestable. It does not permit inspectors to comment on any inherent deficiencies in the content of the officially approved curriculum – only deficiencies in its planning and management by schools. This is deeply problematic. The rationale, aims, concepts and content all need to be considered *worthwhile* for a school's curriculum to be judged 'good', and inspection criteria should embody that value dimension. Ofsted's claim that it is able to report on the quality of the curriculum is thus a very partial and flawed one.

Inspection judgements and grading

The evaluation of teaching, the curriculum and other aspects of the school are inevitably qualitative. Nothing speaks for itself; everything needs interpreting, and that interpretation inevitably involves value judgements and the use of qualitative descriptors such as 'good', 'very good', 'excellent', 'satisfactory', 'reasonable', 'fair' and 'poor'. In advance of an inspection, there can be no stipulation as to which qualitative terms are to be used; the terms must 'fit' the perceptions of the activity or activities being evaluated. They cannot be reduced to just four numerical grades as under the current Ofsted framework; reality is much more complex than a fourfold categorization. The oversimplification fails to take into account the many varied facets of educational reality, which can only be captured (and then only in part) in well-crafted prose. Inspection teams need the freedom to dispense with artificial, misleading constructs such as overall inspection gradings, and to present schools in their idiosyncratic variety with idiosyncratic descriptors to match. Each inspection report has to be bespoke – not a formulaic account with minimal variation from school to school. Misleading oversimplistic grades should make way for prose that gives a vivid sense of what a particular school is really like – as seen by a group of experienced, expert observers. That qualitative richness needs to be built into a reinterpreted inspection system.

Inspection recommendations

No school, however notionally 'outstanding', is perfect. There is always more to learn from the experience of other schools, and inspectors can help bring that experience to bear when reporting their findings. Inspections should result in recommendations, not in diktats about 'what the school needs to do to improve'. Inspectors should raise issues that a school needs to consider, not necessarily to act on; that is a crucial distinction. However, there needs to be a legal and professional obligation on the part of schools to respond publicly about how they have considered and responded to those recommendations, even if it is to reject them in part. This would both reflect and reinforce a view of inspection as providing a set of provisional, tentative, time-specific judgements that inform, rather than necessarily override, the similarly provisional,

tentative and time-specific judgements of staff, governors and parents. Providing recommendations for schools to consider, rather than to comply with, would serve to respect rather than undermine the professional judgement of staff.

In contrast to the current legal position, inspection reports should never of themselves determine an institution's future, but they should instead inform local decision makers – a crucial distinction. Without being the sole determining factor, such reports can nevertheless be very powerful in their advocacy of the need to protect valued aspects of the work of a school or the need to consider changing policy and practice. That change in tone and substance from diktat to recommendation would need to be part of a reinterpreted inspection system.

Moving forward

Despite the introduction of a new inspection framework in 2019, the current inspection regime continued to be subject to considerable criticism up to the beginning of lockdown in 2020. Anxieties were, and are still being, raised over the practice of grading, over context-insensitive judgements, over continued over-reliance on test and examination data in determining standards, over the reliability and validity of inspectors' judgements, and over the high-stakes nature of inspections for schools, staff and students.

Before April 2020 there were calls by professional associations and other interest groups, such as the Headteachers' Roundtable, for a pause in school inspections to re-examine inspection principles and practice. In the event that pause did ensue – but for health-related, not educational, reasons. There have been virtually no whole school inspections since lockdown, and that will apply until at least the beginning of 2021. It is even possible that that pause will be extended until autumn 2021, as schools cope with the disruption to learning and examination preparation due to the pandemic. If whole school inspections are to commence early in 2021, there would need to be modifications to inspection criteria and grading to reflect the change in schools' circumstances, and to reflect the fact that performance data from 2020, and possibly from 2021, will not be available to inform judgements of quality and standards.

In any 'new normal' following the pandemic, there will be strong arguments for the replacement of what many see as an adversarial inspection system with a more consensual one focused on commonly agreed principles and based on greater awareness of the nature of inspection as an appreciative process – tentative and provisional, although still very valuable when conducted in a context-sensitive fashion.

This article is intended to contribute to that much-needed debate about the future of school inspection post-Covid. Teachers' unions, subject associations, parents' groups, the Chartered College of Teaching and personnel from universities, local authorities and multi-academy trusts should be involved, along with Ofsted itself and the Department for Education. Given its current leadership, Ofsted is unlikely to be willing, or to be able, to orchestrate a fundamental reappraisal that would retain the confidence of the various interests involved. The Education Select Committee might be a suitable lead body to set up such a fundamental, wide-ranging review, or an independent charitable foundation might be encouraged to do so. Even a public inquiry, or something similar, might be possible.

A post-Covid world will necessitate the re-examination of a wide range of previously held assumptions, policies and practices, not only in education but in other policy areas, such as health and social care. School inspection should not be an exception to that fundamental review.

Notes on the contributor

Colin Richards is Emeritus Professor of Education at the University of Cumbria and has been Visiting Professor at the Universities of Leicester, Warwick and Newcastle. A former primary school teacher, he was an H.M. Inspector from 1983 to 1996, specializing in primary education and in teacher education, and holding the post of staff inspector (Curriculum 5–16) followed by that of Ofsted's Specialist Adviser for Primary Education. As a critic of much (although not all) of past and present government education policy, he is a frequent contributor to the national press (in particular to *The Guardian*, *The Observer*, *The Times Educational Supplement* and *Schools Week*). He tweets at @colinsparkbridg.

References

- Coffield, F. (2017) *Will the Leopard Change Its Spots? A new model of inspection for Ofsted*. London: UCL Institute of Education Press.
- Headteachers' Roundtable (2019) *Doorstep Manifesto – General Election 2019*. Online. <https://schoolsweek.co.uk/wp-content/uploads/2019/11/HTRT-Manifesto-A4.pdf> (accessed 15 August 2020).
- Monk, R. (2005) *Wittgenstein*. London: Granta.
- Ofsted (2019) *Education Inspection Framework*. London: Ofsted.
- Richards, C. (2016) 'Dear Amanda Spielman...'. *Times Educational Supplement*, 18 November. Online. www.tes.com/magazine/article/dear-amanda-spielman (accessed 18 October 2020).
- Vickers, G. (1983) *The Art of Judgment: A study of policy making*. London: Longman.