

Improving the way Ofsted inspects education consultation: United Learning response

Sir Jon Coles

April 2025

This paper provides a response on behalf of United Learning to Ofsted's consultation, *Improving the way Ofsted inspects education*. We are submitting our response in writing in this format rather than via the online portal as the key overarching points we wish to make go beyond the specific questions there.

United Learning is a national group of over 100 schools founded 130 years ago, now with 75,000 pupils and 12,000 staff, which aims to provide excellent education to children and young people across the country. As with our contribution to Ofsted's previous 'Big Listen' exercise, this response is based on and informed by our experience of around 150 Ofsted inspections of our schools and training in the last decade. We believe this is the largest available sample of direct experience of inspection as a provider of education.

It would be better to prioritise stability in the inspection framework than make major change

Our position is clear: it is not desirable to change the inspection framework at present.

As we set out in our policy paper last year, *Supporting Schools: how school inspection could be improved*¹, the current framework has been about as good an inspection framework as we have had. We would not be in a rush to change it significantly. The current framework has been coherent, conceptually sound and rooted in evidence about what leads to good education and outcomes. In starting from a school's intentions, it is as un-prescriptive as is reasonable in a public service as to what those intentions should be; and it takes seriously the need to look at aspects of provision which cannot be assessed through data alone.

In addition to it being a sound framework, there are also significant inherent benefits to stability. An established framework is well understood, familiar and relatively predictable. This improves consistency and reduces stress. Over decades, schools which are doing reasonably well have rarely had more than one full inspection under any one framework: this discontinuity is a major cause of uncertainty and stress – schools can't trust that it will be 'like last time'.

We were also not convinced that there was a strong case for removing overall effectiveness judgements – we described some of the challenges that we envisaged with this in our policy paper and those concerns stand.

We recognise that the decision to make change was not in Ofsted's control. But if change is a given, and we're already set to forego the benefits of stability, we should make sure that any changes made address contribute to addressing the actual problems with inspection.

The problems of the inspection system are not mainly about the framework

So, what are the actual problems that need to be solved?

In our view, the biggest issue for the future of inspection is that Ofsted has recently become too 'soft'. There has been a catastrophic collapse in the confidence of some inspection teams in the last couple of years, meaning that some weak schools have been judged 'good' while in some cases, lead inspectors have not been prepared to countenance any judgement lower than RI, even for schools with glaring inadequacies.

¹ <https://unitedlearning.org.uk/Portals/0/unitedthinking/improving-inspection.pdf>

This has exacerbated the problem that the overwhelming majority of schools are in just two grade categories. We agree with the Labour manifesto's statement that: 'A system which declares nine in ten schools are good or outstanding fails to provide sufficient information on school performance.' In our view, there needs to be a complete reset of standards, and greater differentiation between schools currently judged to be 'good'.

Second, as discussed in greater depth in our previous policy paper on inspection, those responsible for schools need to reduce the emphasis that we place on Ofsted. This would reduce the 'stakes' associated with inspection and reduce the pressure on reliability and consistency. Ofsted should provide useful information to parents and a check on the performance of those running schools, but we should not pretend that Ofsted judgements are the only or main source of information available to those of us running schools – any effective local authority, governing body or trust will have far more information, data, firsthand knowledge and direct experience on which to base a much deeper and broader understanding of the strengths and weaknesses of their school(s). Every governing body, Trust Board and local authority can and should know their school or schools well enough that Ofsted will not identify unexpected problems during an inspection – and should commit that they will never therefore make significant decisions about a school or its leadership solely on the basis of an inspection.

Third, we previously highlighted issues of variable inspector quality and the need to increase Ofsted resources. There are two factors which in our experience make it more likely that the inspection process is less good or the judgements less sound: the first is inexperienced inspectors with limited school leadership experience; the second is short inspections. We proposed that higher minimum expectations of experience and seniority should be set for inspectors on appointment (which requires the ability to pay enough to attract people of this calibre) and moving to longer or better resourced inspections. Ofsted's resources need to be increased to facilitate both these things.

Last, as reflected in our 'design principles' for an effective inspection system, the system needs to, and does not currently, command sufficient 'professional consent' from those inspected. A lack of trust in the process of inspection, a lack of faith in the outcome being fair and a strongly negative perception of the system as a whole seem to be the drivers of stress – despite the fact that the majority of teachers are satisfied with their own most recent direct experience of inspection.

The proposals make sense as a response to government policy but do not address the problems as fully as they could

There are important aspects of the proposals which we support but which do not fully address key problems.

Firstly, we agree that there needs to be a change in the number of grades and the words used for them and that five grades makes sense. There needs to be a re-set in standards and in the distribution of grades, because the vast majority of schools are now clustered in the top two grades. In order to do that, there needs to be a change in the number of grades – we can't have some schools which hold a 'good' judgement for the next three years while 'good' becomes a much higher standard. We agree that it is better to have five grades than three: something approximating to a 'bell curve' distribution of schools across 5 grades, the middle one of which is fully satisfactory, would better represent the reality of school performance in the country on all dimensions.

However, while we support a change to five grades *as part of a significant re-setting of standards*, we are not yet reassured that there is enough elsewhere in the proposals to deliver that reset – a new grading scale needs to be accompanied by new expectations. While the standards set out in the consultation could be seen to set a reasonable standard for a ‘secure’ judgement, everything depends on how they are interpreted and used in practice, the quality of inspectors and their training, and critically, the messages inspectors receive about the level of performance needed to reach ‘secure’ or ‘strong’ standards. The messages and tone of the consultation do not yet suggest a major raising of the bar.

Secondly, we support the idea proposed that Ofsted should be drawing wherever possible from pre-existing standards and not inventing its own. This is a good idea, which had not been proposed elsewhere. It could usefully contribute to reducing the sense that Ofsted has its own agenda, reducing incentives on schools to do things ‘for Ofsted’, and could contribute to better schools for children provided that the standards are themselves soundly evidence-based as factors in providing well for children. Ofsted should satisfy itself and others that this is the case.

Thirdly, we support the end of ‘short inspections’. These are insufficiently resourced to provide sufficiently reliable inspection results and it is right to focus resources on full inspection.

However, taken as a whole, these proposals appear to increase rather than reduce pressure on inspection time. Unless there is a substantial increase to Ofsted resource, which has not been signalled, then keeping the same inspection cycle, replacing s8 with full inspections, adding multiple inspections of weak schools and potentially adding inspections of MATs and of safeguarding cannot be achieved alongside improved inspection quality.

A practical solution for improving inspection within existing resources

Whilst we would not be prioritising inspection change at this point, given that changes *are* being made, there is an opportunity to consider a different approach that might better address the most pressing issues set out above. On the assumption that the current state of the public finances means that the total resource available for inspection is, at best, fixed, we think that it would be worth exploring a model incorporating the following key features:

- Greater use of **risk-based inspection**, with a **longer inspection cycle for effective schools**. This would not realistically be longer than once every 6 years, but where a school is at least secure overall and not flagging concern on a data-led risk assessment (key indicators not significantly declining since a previous inspection), a longer cycle should be acceptable.
- With a longer inspection cycle, the level of information to parents should then be increased by the introduction of a **single definitive annual report card** as proposed in the Labour manifesto. This could be along the lines previously proposed by DfE under the previous Labour Government², incorporating outcomes from the most recent Ofsted inspection, headlines from the annual peer review, and the latest performance data all in one place.
- In years 4-6 of an inspection cycle, the **annual report card would provide the basis for the annual risk assessment** of whether earlier inspection might be required. An earlier re-inspection would be triggered by significant decline in key data or information since the previous inspection. This would be a transparent, widely understood risk assessment mechanism, which would command professional and public confidence.

² <http://data.parliament.uk/DepositedPapers/Files/DEP2009-1851/DEP2009-1851.pdf>

- The money saved through reduced inspection volume due to the lengthening of the inspection cycle for most schools should be reinvested to:
 - raise inspector pay to attract people with sufficient school leadership experience; and
 - increase the resource (number of inspector days) per inspection to ensure that inspections are valid.
 The increased resource would also enable slightly shortened inspection days, benefiting both leaders in schools being inspected and the appeal of being an inspector, especially to serving Heads.
- To support ongoing challenge and continuous improvement in already secure schools during these longer inspection windows, there could be a new **requirement for annual or biennial peer review in the interim** from an organisation accredited or quality-assured by Ofsted – which could be the responsible body for the school or another organisation.
- Additionally, a future process of **inspection of trusts, groups of schools and responsible bodies** could provide additional information or assurance to parents. Such a process (and one of regular peer review) could justify further elongation of the formal inspection cycle.
- The risk-based approach would guarantee **a greater proportion of limited Ofsted resource focused on schools where there is a risk that children are not being well served** but without the excessive focus on schools ‘in a category’ proposed in the current consultations.

A more risk-based approach, with longer windows for secure schools not flagged as high risk, would make better use of finite Ofsted resources. That would enable steps to be taken to improve inspector quality and resource inspections better. More use of accredited peer review and a high profile, authoritative annual report card drawing on a range of data and information could help to reduce the reliance on or prominence of inspection. These measures have the potential to improve professional buy-in to the accountability system as a whole.

Suggested improvements to the current proposals if they go ahead

In the event that a version of the consultation proposals does go ahead, we set out in the remainder of this response our commentary on the specific proposals. It is welcome that Ofsted have avoided some of the most significant potential pitfalls we highlighted in our previous policy paper:

- The proposed presentation and grading of outcomes is clear enough that parents will still get accessible information and an easy to understand sense of how effective Ofsted judges the school to be. Some in the sector seem to have been hoping for solely narrative-style reports, but the responses Ofsted gathered from parents demonstrate clearly that this would not have met their needs.
- It is good that there will remain a clear judgement on safeguarding. In general, we are pleased to see that many of the aspects we have argued do not need to change have been kept in these proposals.
- The judgements made and outcomes reported still enable government to identify on an objective basis with clear criteria, schools in need of further support or urgent change. This is reflected in the proposals set out in the parallel DfE accountability consultation. Without

an overall grade, DfE has to make decisions overtly about when to intervene which were once internalised within Ofsted.

- We are pleased to see the abolition of ungraded inspections, which were too short and insufficiently resourced. We welcome a rethink of the deep dive methodology in primary schools, where it was less effective and less manageable, though we do think that a clearer inspection methodology needs to be developed and articulated.

There are several things we would want to see refined or changed in current proposals if the plans retain broadly their current shape.

The inspection framework falls into the same trap as recent national leadership qualifications in giving insufficient attention to the **effective leadership of people**. In many respects, this is the most important aspect of leadership, crucial to all other aspects of quality and effectiveness in a school, as well as to improving recruitment and retention in the profession. Inspection should be examining whether there is clarity of vision and purpose, whether a clear direction is set for staff, whether staff are aligned with the school's vision, direction, approach and values and whether staff have appropriate opportunities to be involved in decision-making. There is a passing reference to workload under Leadership and Governance – though this is not well constructed and does not make a connection to whether work is meaningful; there are references to professional development in relation to 'Developing Teaching', but none in relation to support staff development. There is no consideration of effective leadership and line management of staff or their broader career and personal development, and no reference to staff wellbeing (only leaders').

We would like to see more focus on provision **beyond the classroom**, which is fundamental to the inclusion of children, their affiliation to school and sense of belonging and to developing a pupil as a whole person. We think that it is a very poor state of affairs that a school with weak extra-curricular provision, and little by way of sport, music, drama and so on, can be judged to be an excellent school overall. The framework has become overly technocratic and narrow. This is not solely Ofsted's responsibility, and goes beyond inspection, reflecting a wider issue in the schools system – but including it in inspection could signal a change and make a big difference. This – like other aspects of development of the whole person and their character – is an important aspect of a school that is hard to assess with data, and where therefore inspection has a particularly important role to play. These are all currently underplayed.

Further work is needed on the **details of the toolkits**. In particular there are instances where small and undefined differences in adjectives appear to be having to carry the weight of a distinction between secure and strong. This seems to be a recipe for confusion, inconsistency and unnecessary stress. We do not think that every part of every toolkit necessarily needs to show progression – it may not be credible that every aspect of practice has five distinct levels within it. What matters is that there is clarity about what differentiates each grade level as a *whole* within each judgement area – that it is clear and easy to understand what the difference between the grades is supposed to be overall. So it might be acceptable to have the same sentences at different levels of the toolkit (i.e. schools at level 3 are now as good as it gets in this one respect – the difference between grades 3 and 2 is found in other parts of this domain) rather than pretending every aspect gets better in every respect. This would certainly be preferable to changing an adjective to give the impression of pointing to an identifiable difference where there is not one, or without being clear what it is.

As above, we do not think there is sufficient capacity in Ofsted for what is proposed in terms of the frequency and volume of **monitoring** of schools causing concern. Nor would this be a good use of that resource if it existed. Termly monitoring of a school would not be conducive to effective school improvement.

The process described in the consultation for **determining 'Exemplary' grades** is cumbersome. There needs to be quality assurance and broad consistency if this grade is intended to be rare and exceptional, but the benefits of the process described do not appear be worth the cost. No grade should be so rare that the scale effectively reverts to being a four point one – we do want to see all grades used in a meaningful number of schools. Identifying the strongest practice for the benefit of the system is much more useful if it is identifying the best 10% of practice than the best 1%; and this is not such a compelling and critical school improvement lever that a very resource-heavy quality assurance process is justified.

The relationship between Ofsted's report cards and DfE's school profile is confused at the moment. Two separate products are proposed, with similar purpose and audience, based on overlapping information sources, but not actually matching or aligned. It sounds as though there is an intention for performance data and some kind of 'other information' to feature in both the report card and a separate school profile – potentially drawing on the same datasets but reflecting different points in time, which could be confusing for parents and the public. If done well, we are supportive of a *single* report card being produced for each school, and that would be a component of our proposed alternative approach as set out above. But the opportunity to produce a single annual report card as proposed in Labour's manifesto is in danger of being missed.

We do not think that sufficient attention has been given to the **risks of including more contextual information**. It is well known that there are parts of the country where educational performance is low; and there are groups of children – especially poor, white children – who fare very badly. But our improvement goal must be to change this. So, if contextual information brings nuance to headline performance measures, and highlights the successes of 'schools achieving against the odds', making it more attractive to teach in places which struggle to find strong teachers, then that is to the good. But there is an at least equal and opposite risk that 'contextual data' lowers expectations: if in a desert of success, mediocrity is lauded then that is nothing other than returning to a 'what can you expect from these kids' mentality.

We are **less concerned about the exact number of judgement areas**. While we understand the technical issues of decreased reliability in any one judgement arising from an increased number, the more important question from the perspective of a parent (as the main audience) is whether the overall picture is broadly right, which relates to the distribution of judgements overall – and may actually be helped by having more judgement areas. In this context, we caution both Ofsted and DfE against their current tendency to over-state the possibility of granular detail in inspection driving school improvement activity. Inspection is and should remain a 'best fit' activity.

Finally, our reflection on the pilot experience so far is that it has felt very different to inspection under the current framework, with less time for discussion with leaders and more focused directly on criteria. This could be because it is a new framework and inspectors are familiarising themselves with new content and detail, but we are concerned that this should not be the experience of schools facing inspection for the first time under a new framework and think that this requires some attention.