

# **Headteachers' experiences of school inspection under Ofsted's January 2012 framework**

## Abstract

This paper focuses on headteachers' experiences of inspection under Ofsted's revised school inspection framework, their views of its principles and its implications for school leaders and leadership. The paper draws on findings from a mixed-methods study to show that inspections are more focused on pupils' attainment and progress. Headteachers intend to prioritise these and other judged areas over those no longer explicitly judged. Whilst broadly agreeing with the framework's principles, many headteachers report that inspection was less positive owing to variation in inspector quality and rigidity in the (interpretation of the) framework.

The paper argues that leaders of schools serving socio-economically disadvantaged areas might find it harder to obtain a good Ofsted rating, with implications for headteacher recruitment and retention. It argues for improved Ofsted inspector training; a broad, values-driven leadership agenda by headteachers and the recognition and promotion of contextual responses to educational challenges.

**Keywords:** Ofsted; framework; school inspection; headteachers; school leaders; leadership.

## Introduction

The influence of external inspection regimes on schools and their leaders has been charted from a number of perspectives since Ofsted's<sup>1</sup> creation in 1992 prompted a shift in how accountability was constructed and operationalised. Common to most has been the recognition of the headteacher as central to the process (Muijs and Chapman, 2009). This is partly an artefact of such regimes themselves, which require a hierarchical distribution of labour within schools in order to render accountability meaningful. Consequently, headteachers' views have been researched on a range of aspects related to inspection experiences, including post-inspection actions and outcomes (Ouston, Fidler and Earley, 1997), and the impact of years' experience (Ferguson, Earley and Ouston, 1999) and school context (Chapman, 2002) on headteachers' attitudes to inspection. Such studies were located predominantly in the field of school improvement since Ofsted had adopted as its slogan,

‘improvement through inspection’. As this claim was investigated and found wanting, scholars began to question more deeply what inspection was for, a question made more pertinent to an international audience by the ‘borrowing’ of the Ofsted model by other national jurisdictions (see, e.g. Morrison, 2009). Several researchers have since located Ofsted within a performative regime, which structures the managerialism it represents in the leaders whose compliance it demands (Ball, 2008; Perryman, 2006; Hoyle & Wallace, 2007; Troman, 1997; Gilroy & Wilcox, 1997). For these scholars, inspection is about control rather than improvement: school leaders internalise the requirements of the inspection regime and subject themselves and in turn their staff to surveillance through, for example, self-evaluation processes. School leaders may believe themselves to be autonomous, but in fact their practice is extrinsically determined (Ball, 2003; Wright, 2001). One principal mechanism for achieving this compliance is through the publication of the inspection framework, whose tenets must be followed if a school is to be judged favourably.

### **The research focus**

Ofsted introduced a new inspection framework for use from January 2012 following the new Coalition government’s education White Paper (DfE, 2010: 69). This declared that Ofsted should ‘focus on just four things—pupil achievement, the quality of teaching, leadership and management, and the behaviour and safety of pupils’. Ofsted accordingly reduced the number of inspection judgments from twenty-seven to these four areas along with a judgment of the school’s overall effectiveness. Further key changes include the exemption from routine inspection of schools currently rated Outstanding<sup>2</sup>; an increased focus on pupils’ progress, especially in Maths and English; and a change in the way that progress is measured. In the

2009 framework, inspectors had looked at a contextual value-added (CVA) measure which took statistical account of such factors as a child's socio-economic status, gender and ethnicity. In reverting to a simpler value-added (VA) measure, the Department for Education argued that all pupils should be expected to make the same amount of progress, regardless of background (DfE, 2010). The framework was further revised in September 2012; the research reported in this paper therefore captures a moment in which a new direction for inspection policy was established.

This paper contributes to the discussion of inspection through examining the introduction of the framework as a policy intervention, focusing on its impact and implications. It reports findings from an exploratory study into headteachers' experiences and views of the framework, supported by the Economic and Social Research Council [grant number ES/J500094/1].

## **Research design**

This mixed methods study adopted what Creswell et al. (2003) call a concurrent triangulation design. An online survey collecting Likert-scale data and semi-structured interviews were used to improve confidence in the findings and mitigate any shortcomings of either method. Triangulation was methodological (Cohen et al., 2007) and multilevel (between the individual and the national, see Youngs and Piggot-Irvine, 2012). The data were collected sequentially, with some survey respondents then volunteering for interview, but analysed concurrently to answer the following research questions:

1. What are headteachers' experiences of inspection under the January 2012 framework?

2. What are headteachers' views of the framework's provisions and objectives?
3. What do headteachers believe will be the implications of the framework?

The data were integrated at the interpretation stage. The survey results were analysed using SPSS and the interviews were transcribed and coded using NVivo. Data coding was influenced by the research questions, the literature and emergent survey findings.

All English secondary schools inspected under the January 2012 framework and whose report was published on Ofsted's website between 1 February and 31 March 2012 were located using that website's search function, producing 175 eligible participants. These were emailed a link to a web-based survey and reminded after a fortnight. Of the 36 (21%) respondents, four (11%) led schools graded Inadequate, fourteen (39%) were graded Satisfactory, fifteen (42%) were Good, and three (8%) Outstanding. This compared with 17%, 42%, 36%, and 5% respectively in each category amongst the 175 inspected schools eligible to respond. A Chi-Square Test revealed that this difference was statistically non-significant ( $\chi^2=2.3$ , DF=3,  $p>.5$ ). This, along with a non-significant gender difference between the responding and eligible sample, ( $\chi^2=1.3$ , DF=1,  $p>.25$ ) shows that statistically, headteachers of schools graded highly, or poorly, or female or male leaders, did *not* participate in the survey in significantly smaller numbers, and improves confidence in the representativeness of the sample. 25% of respondents reported an improved Ofsted rating; 39% retained their rating and 36% saw their school's rating worsen.

Surveyed headteachers could volunteer for a semi-structured interview. Twelve did so, of whom six were selected from across England, demonstrating the full range of inspection outcomes. Whilst avoiding obvious skewing, the group has no statistical claim to representativeness. One interviewee "came out" as an Ofsted lead inspector during interview. Further interviewee characteristics are withheld to avoid the possibility of identification. The

headteachers were interviewed in person for 45 minutes or by telephone for approximately 35. Surveyed headteachers are referred to as respondents, and interviewed headteachers collectively as interviewees and individually as HT 1-6.

## **Findings**

Results show that reducing the number of judgments has resulted in inspections which are particularly focused on pupils' VA progress in Maths and English. Varied inspector quality, rigidity of framework or its interpretation, and the difficulty of rating highly in schools serving disadvantaged communities contribute to these inspections being experienced less positively. Nevertheless, headteachers largely support the framework's objectives and intend to focus more on the areas it privileges at the expense of those no longer inspected.

### ***Headteachers' experiences of inspection***

The questionnaire first asked respondents about their inspection experiences under the framework, focusing on the extent to which its objectives were realised (see Table 1). Largely, inspections *were* more focused on the four judgment areas. For a majority of respondents, the framework does not, however, adequately capture schools' activities, contexts and achievements. Nor has the framework produced inspections which are experienced more positively.

*Table 1. Survey respondents' experiences of inspection under the 2012 framework.*

*Key to headings in tables 1-2.*

1=Strongly agree

2=Agree

3=Neither agree nor disagree

4=Disagree

5=Strongly disagree

6=Don't know

*Figures (discounting headings) represent percentages and may not sum owing to rounding.*

	1	2	3	4	5	6
There was more focus on pupils' progress	56	42	0	0	3	0
Pupils' behaviour mattered more	39	42	8	11	0	0
Nothing has disappeared: inspected areas have moved	0	42	3	42	14	0
Inspectors spoke to more children	31	39	6	17	0	0
Teachers were observed more	56	33	6	6	0	0
Ofsted wanted to know what my school is normally like	17	33	14	28	8	0
This inspection captured my school	17	30	6	25	22	0
I agree with the inspectors' judgments	19	36	6	25	6	0
This inspection was a more positive experience	3	22	25	22	28	0

All respondents agreed that there was more focus on pupils' progress, a finding corroborated by five interviewees. HT 6, for example, noticed 'the very, very sharp focus on achievement and progress'. 81% thought that pupils' behaviour mattered more. Respondents disagreed over whether the framework streamlines priorities; 56% believed some areas were no longer inspected, whilst 42%, felt nothing has disappeared. HT 4 was one of five interviewees noticing the absences:

All of the other stuff, if you like, which is important to schools and would have got recognition under the previous framework around community cohesion and all of that soft stuff that's not teaching and learning ... that's all gone... (HT4)

Twenty percent more respondents believed Ofsted wanted to know what their school is normally like (53%) than did not (33%). HT 6 supported this; 'Ofsted's key purpose is... [to ask], what do children typically get in a typical day, or a typical week in this school?' One way in which this typicality was assessed was through inspectors talking to more children; 70% of respondents agreed or agreed strongly that this happened.

#### *Importance of classroom observations*

The desired increased focus on teaching and learning has resulted in more lesson observations, as 89% of respondents report. This was corroborated by the qualitative findings, HT 5, for example, noticing 'a stronger focus on classroom observation', and HT 4 reporting that the inspection 'was completely focused on what they saw in the classroom', but feeling that the lead inspector's interpretation was unfair. Although the substantive issue here is inspector quality, the subjective way in which different forms of data are selected and interpreted by inspectors to match the emergent narrative created by attainment and VA progress measures in (a) core subject(s) is a key finding. For instance, HT 6 said the inspectors 'weren't absolutely convinced ... they were seeing evidence in the classroom to back up our [higher] judgments'. HT 6's school's weak VA data in Maths provided inspectors with an impression of the school which disinclined them to pay attention to dissenting evidence, including HT 6's arguments. During HT 2's inspection, a similar



dissonance occurred between teaching standards observed during inspection and those recorded by the school; here, however, 'because [their] ... value-added data was so good' the interviewee was able to persuade the inspectors 'that the dissonance was simply stuff happening on the day, and teachers over-talking a bit, rather than a real problem' (HT 2).

### *The use of data*

The use of data *per se* did not feature in the questionnaire because there was little indication in the framework that it would change. The interviewees disagreed, for them its increase and changed use under the new framework was significant. Two felt that the inspectors scrutinised English and Maths more than under the former framework and applied any weaknesses found there to their interpretation of the school's performance more generally. One further interviewee, HT 6, had the same experience, but thought the inspectors justified; 'we could not demur from their view that ... standards were inadequate'.

### *Inspector variability*

In the interviews, *level of collaboration with inspectors* became a significant theme expressing a continuum of experiences. At one end, HT 3 'didn't feel there was any imposed outcome on anything'. However, HT 1 found the inspection team 'too rigid', but attributed this largely to its inexperience with the framework, 'they were still very much working to strict guidelines'. HT 4 agreed that inspectors were 'very much following the wording'. HT 3, the Ofsted inspector, describes inspectors' training:

...in the guidance for us all ... you look for the general sense of a level, you don't go down and say, I see you haven't, der der, and therefore I can't give you a Good on that, for example. So it's a lot less prescriptive.

This contradicts most others' experiences; HT 6 is typical, 'they kept going to the Inadequate ... if any of these bullet points is placed yes, therefore the judgment has to be a four'.

Next, the survey sought respondents' views of the validity and reliability of the inspection. They were evenly divided regarding the former, with the interviewees reflecting this split:

[The inspectors said] '...can I tell you some of the things that I noticed and which you may wish to think about'. How fabulous is that? ... But they'd noticed it going round, wasn't on the framework, they knew because they're educationalists that it was important. (HT 3)

HT 5, however, felt strongly that the inspection purposefully overlooked the school's achievements, saying '...the school has got significant strengths. And we felt that those were ignored'. Again, the connecting theme is variability in the quality and judgment of inspectors, raised spontaneously by all six interviewees; even the inspector said, 'inspectors are human, they do all have their own hobby-horses, they do have a mixed background' (HT 3). This supports the survey finding that 56% of respondents did and 39% did not agree with the inspection judgment.

*Inspections are less positive*

50% of respondents found this inspection a less positive experience, with 25% finding it more positive. A Mann-Whitney test confirmed that schools whose Ofsted rating improved had a significantly more positive experience than those whose rating worsened, ( $U=18$ ,  $Z=-2.8$ ,  $p<.01$ ). Two of the six interviewees found their inspection a positive experience. One was the Ofsted inspector, who enjoyed ‘an honest and open relationship’ with the inspectors (HT 3). Another, HT 6, found the inspectors and the process ‘rigorous’, but thought that ‘the judgment was fair and accurate’. The others all found the inspection negative; for HT 1 it was ‘frustrating’, for HT 4 it was ‘deeply traumatic ... [and] unpleasant’ and for HT 5 ‘it felt like a completely negative experience’. The reasons for these experiences can be placed along a continuum with *framework-related* and *inspector-related* at either end. The framework was criticised for its rigidity and its assumptions; HT 1, for instance, thought they ‘had a fair team restricted by, really, the framework’, and HT 5 believed it to be ‘deficit-led’. Mid-way along the continuum are examples of *inspectors’ rigid interpretation of the framework*, including an argument with inspectors HT 4 reported over the interpretation of the phrase, ‘rapid and sustained’. Finally, inspection experiences can be negative because of the inspectors. HT 4, for example, found some ‘unfriendly ... confrontational and downright rude’, producing an ‘unnecessarily adversarial’ inspection.

### ***Headteachers’ views of the framework***

Next, the questionnaire sought to establish the extent to which respondents agreed with the provisions and objectives of the January 2012 framework (see Table 2).

*Table 2. Survey respondents’ views of the 2012 framework’s principles and objectives.*

Figures (discounting headings) represent percentages and may not sum owing to rounding.

	1	2	3	4	5	6
The last framework was too centred on checking compliance with government policies	8	42	17	22	3	8
VA is better at judging pupils' progress than CVA	3	28	25	25	19	0
The new framework is right to try to be more focused	6	50	17	19	3	6
Behaviour deserves to receive more focus in inspections	8	31	39	22	0	0
Schools rated Outstanding should be inspected less frequently	6	31	17	36	11	0
The increased focus on progress is welcome	8	72	6	8	6	0
The desired levels of progress are achievable	0	47	11	31	11	0
Inspections shouldn't assess pupils' diet	8	33	42	17	0	0

Table 2 shows that respondents broadly support the framework's principles, with most welcoming its tighter focus. The interviewees agreed; 'I do think that the narrow emphasis ... is easier to cope with than the dozens of areas that were there before' (HT 5). Only 17% of respondents thought that pupils' diets, exemplifying a former judgment area, should still be judged. *Behaviour and safety* and *pupils' progress* receive more attention in this framework. Over 80% of respondents agree about progress, though with a caveat concerning the rates expected. Behaviour's promotion is less endorsed, with under 40% in favour and a similar

number expressing neutrality, a statistic corroborated by behaviour's absence from the interviewees' data. Two features of the 2012 framework receive less support. Eleven percent more respondents disagreed (47%) than agreed (36%) that schools rated Outstanding should be inspected less frequently than others, and finally, the abolition of CVA is the least popular of the new framework's features, with 45% of respondents disagreeing that its replacement, VA, is a better way of judging pupils' progress against 31% agreeing. Both views were expressed by the interviewees:

But when you ... started to look at the levels of progress data for Maths and English, you started to realise... the CVA judgments are disguising some really quite significant issues with youngsters' progress in the core subjects. (HT 6)

HT 2, with excellent value-added outcomes, nevertheless highlighted the implications of abolishing CVA:

Standards for children *shouldn't* be related to their context, but the judgment they make on the school *should*, because it just is harder ... I've got no axe to grind because we got Outstanding ... but it is made easier by having enough children that come from quite motivated homes ...

### ***Headteachers' views of the framework's implications***

To answer the final research question, survey respondents were asked if they would change the focus they place on areas which are newly privileged in, which have survived in some

form in or which have disappeared from the framework. Broadly, respondents intend to focus more on those areas appearing in the framework, and less on those which no longer do so (see Table 3).

*Table 3. Headteachers' views of the implications of the 2012 framework for their practice.*

*All figures represent percentages.*

	More focus	No change	Less focus
Care, guidance and support will receive	11	58	31
The quality of teaching will receive	78	22	0
Pupils' literacy skills will receive	72	28	0
My school's community engagement will receive	0	36	64
Behaviour will receive	31	61	8
Pupils' safety will receive	8	86	6
Leadership and management will receive	50	44	6
Pupils' progress will receive	67	33	0
Pupils' health will receive	0	44	56

Two of the four judged areas; *quality of teaching* and *leadership and management*, survive from the former framework. 78% of respondents intend to focus even more than presently on teaching, and 50% will focus more on leadership and management. *Behaviour and safety* receive more emphasis in this framework, yet in a finding consistent with that showing that respondents are unconvinced about its "promotion", most indicated that they would not change the importance they attach to either constituent element. Pupils' progress is a newly

privileged component of the fourth inspected judgment, *achievement*, and two thirds of respondents will focus more on this. Whilst not separately judged, pupils' literacy features strongly in the framework: over 70% of respondents (supported by HT 2, HT 5 and HT 6) will focus more on it. Three items exemplified areas removed from the framework; of these, two; the school's community engagement and pupils' health, will be de-prioritised by a majority of respondents, and a third intend to focus less on the third; the care, guidance and support offered to pupils. HT 3 would be unsurprised by this, warning that 'there are other heads who actually needed that poke on the community thing to make them take notice of it'.

#### *Further implications of the framework*

The qualitative data provided two further themes, first; *career implications*. HT 6 is no longer a headteacher; 'convincing those higher in education that a headteacher who's led a school from being Good to being a Special Measure school in two years, you know, that's a bit of a tough sell, and I understand that'. A further two plan to retire, one intends to 'just carry on' despite becoming 'disillusioned' (HT 4). This leads to the second theme: *implications for the wider school system*. HT 2, like HT 6, notes that outstanding candidates might be dissuaded from becoming head of a school serving a disadvantaged community because of lower opportunities for Outstanding ratings, or the increased risk of job loss; 'it's just not fair ... if you lose really, really good heads from schools in challenging circumstances because the best you can hope for is Good'. HT 5 felt that the framework was part of 'quite a narrow, anti-school agenda', which seeks to create 'a climate of fear' rather than improve schools.

#### **Conclusion**

Inspections *are* largely more focused under this framework; this focus is predominantly on one of the four inspected areas— pupils’ achievement, defined as their attainment and decontextualised progress, especially in Maths and English. This produces particularly data-driven inspection outcomes and reduces the function of leadership to the attainment of higher scores in pupils’ standardised tests, despite evidence that this is incongruent with a world-class education system (see, e.g., Hargreaves and Harris, 2011).

CVA’s abolition has made it harder for schools serving disadvantaged communities to get Outstanding. The moral argument that background should not excuse low achievement has persuaded many heads, but to construct policy as if social background *has no effect* on achievement is quite different, and shifts the blame for low achievement onto school leaders. This may have consequences for headteacher recruitment and retention in such schools, and for social justice for their pupils.

The framework and variability in inspector quality have contributed to the large number of headteachers finding their inspection to be less positive than under former frameworks. Judgment formation is subjective (Gilroy and Wilcox, 1997); excluding headteachers belies the apparent transparency of a published framework and exposes their school to misdiagnosis. As Ferguson et al. note, ‘[t]here can be no genuine partnership between schools and inspectors while the balance of power is so uneven and the penalties for failing the inspection involve such high stakes’ (1999: 246).

Nevertheless, many headteachers are compliant, supporting the framework’s objectives and structuring their schools in response. The 64% of leaders who intend to spend less time



developing links with their community, seemingly because Ofsted no longer inspects it, will do so in defiance of Chapman and Harris' (2004) findings that these are vital for improving schools facing challenging circumstances (SFCCs). Rather than improving the validity of inspections by broadening their terms, validity here is improved instead by school leaders' narrowing their curriculum to suit the inspection model at the expense of providing a rich learning experience. This raises valid questions concerning the purposes of education, and what *vision* leaders are permitted to have (see Wright, 2001).

In light of these findings, I recommend that more practising headteachers conduct inspections to improve the inspectorate's quality and relevance. I recommend that researchers evaluate whether the latest framework penalises schools serving socio-economically disadvantaged communities, given its potential consequences for headteacher recruitment and teacher morale in such schools. I urge headteachers to re-engage with the moral purposes of educational leadership, rejecting any overly compliant narrowing of pupils' learning experiences promoted by the framework. Finally, I recommend that any future framework recognise the importance of context in evaluating schools' work, not as an excuse for poor performance, but as a starting-point for conversations about, for example, local priorities and leadership styles. It should actively foster contextual solutions through structural flexibility. Developing a framework which encourages the diversity necessary to engender change in a range of contexts may reinvigorate purposeful, values-driven leadership, especially in England's least-advantaged schools.

## **Acknowledgements**

I would like to thank the participating headteachers; Helen Gunter and those anonymous reviewers who suggested improvements to this paper; Chris Chapman for his guidance throughout the research and Jan Webber, inspections specialist at the Association of School and College Leaders.

## Notes

<sup>1</sup> The non-ministerial department which inspects schools in England.

<sup>2</sup> In the January 2012 framework, Ofsted rated schools Outstanding, Good, Satisfactory (since Requires Improvement) or Inadequate. I have capitalised these throughout to distinguish them from other usages of these words.

## References

- Ball, S.J., 2008. The Legacy of ERA, Privatization and the Policy Ratchet. *Educational Management Administration & Leadership*, 36(2), pp.185–199.
- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), pp.215–228.
- Chapman, C., 2002. Ofsted and School Improvement : Teachers' perceptions of the inspection process in schools facing challenging circumstances. *School Leadership & Management*, 22(3), pp.257–272.
- Chapman, C. & Harris, A., 2004. Improving schools in difficult and challenging contexts: strategies for improvement. *Educational Research*, 46(3), pp.219–228.
- Cohen, L., Manion, L. & Morrison, K., 2007. *Research Methods in Education Sixth.*, Abingdon: Routledge.
- Creswell, J.W. et al., 2003. Advanced Mixed Methods Research Designs. In A. Tashakkori & C. Teddlie, eds. *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage, pp. 209–240.
- DfE, 2010. *The Importance of Teaching: The Schools White Paper 2010*.
- Ferguson, N., Earley, P. & Ouston, J., 1999. New heads , OFSTED inspections and the prospects for school improvement. *Educational Research*, 41(3), pp.241–249.

- Gilroy, P. & Wilcox, B., 1997. Ofsted, Criteria and the Nature of Social Understanding: A Wittgensteinian Critique of the Practice of Educational Judgement. *British Journal of Educational Studies*, 45(1), pp.22–38.
- Hargreaves, A. & Harris, A., 2011. *Performance beyond expectations*,
- Hoyle, E. & Wallace, M., 2007. Educational Reform: An Ironic Perspective. *Educational Management Administration & Leadership*, 35(1), pp.9–25.
- Morrison, K., 2009. School inspection in small states and territories: an overview and case study of Macau. *Compare: A Journal of Comparative and International Education*, 39(6), pp.751–767.
- Muijs, D. & Chapman, C., 2009. Accountability for improvement: rhetoric or reality? In C. Chapman & H. Gunter, eds. *Radical Reforms*. Abingdon: Routledge, pp. 28–41.
- Ouston, J., Fidler, B. & Earley, P., 1997. What Do Schools Do after OFSTED School Inspections-or before? *School Leadership & Management*, 17(1), pp.94–104.
- Perryman, J., 2006. Panoptic performativity and school inspection regimes: disciplinary mechanisms and life under special measures. *Journal of Education Policy*, 21(2), pp.147–161.
- Troman, G., 1997. Self-management and School Inspection: complementary forms of surveillance and control in the primary school. *Oxford Review of Education*, 23(3), pp.345–364.
- Wright, N., 2001. Leadership, Bastard Leadership and Managerialism: Confronting Twin Paradoxes in the Blair Education Project. *Educational Management Administration & Leadership*, 29(3), pp.275–290.
- Youngs, H. & Piggot-Irvine, E., 2012. The Application of a Multiphase Triangulation Approach to Mixed Methods: The Research of an Aspiring School Principal Development Program. *Journal of Mixed Methods Research*, 6(3), pp.184–198.