

Written Evidence from the Professor Gemma Moss, Professor of Literacy at UCL, Institute of Education and Director, ESRC Education Research Programme.

This submission draws on work I have undertaken as Chair of the BERA Expert Panel on Assessment. The panel considered alternatives to the English high stakes testing and accountability system. The panel's report was published in 2021 and can be found here:

<https://www.bera.ac.uk/publication/high-standards-not-high-stakes-an-alternative-to-sats>.

My submission is also informed by research undertaken with funding from ESRC on the impact of COVID on primary schools, their pupils and their communities. These were:

1. *A duty of care and a duty to teach: educational priorities in response to the Covid-19 crisis*, funded by UKRI/ ESRC. Reference number: ES/V00414X/1. Data were collected and analysed between May – Sept 2020
2. *Learning through disruption: rebuilding primary education using local knowledge*, funded by UKRI/ ESRC. Reference number: ES/W002086/1. Data were collected and analysed between May – Sept 2021. As part of this project we prepared a briefing note for Ofsted on resuming inspection.

The usefulness of Ofsted inspections and inspection reports.

At the end of the COVID pandemic, Ofsted recognised it needed to listen and learn from the profession about the difficulties schools had faced during a period of extensive disruption and the unpredictable impacts that followed for schools and their communities (Moss et al, 2022). A good example of Ofsted informing itself through listening is provided by the report, Governing in unprecedented times, jointly published with the National Governance Association. It is a record of focus group discussions with governors, structured around open questions, that asked governors to reflect on how schools had managed during the pandemic, the challenges they had faced and the solutions they had found.

In adopting this approach, Ofsted recognised that it did not itself have ready-made answers to a rapidly changing set of circumstances and therefore could not pre-judge what would be the most appropriate response. By listening instead of judging, Ofsted were able to gather new system intelligence to inform policy going forward. Gathering additional intelligence directly from the profession should be an integral part of system oversight. Yet at the moment there are few formal channels through which this happens. This neglect of professional knowledge developed from dealing directly with conditions on the ground weakens the system as a whole. In the case of COVID the anticipated impacts on children's learning predicted have not materialised at the depth expected, but there are indeed long-lasting impacts on children's behaviour, well-being and attendance that schools recognised first (Roffey, 2023; HarmeY and Moss, 2023).

The idea of listening to learn directly from the profession should be at the heart of the inspection system. Too often it is not. Ofsted relies on metrics that are too thin to really enable them to understand the school in its context: FSM is an inadequate measure of the impacts of pupil poverty on educational attainment (Taylor, 2019). The more significant impacts of persistent poverty on

pupils are not yet captured adequately in the data ([Tuckett et al, 2022](#)). The rise in child poverty that this signals is indeed part of a much bigger societal problem that schools alone cannot solve. Absence post-COVID has spiked, yet there is no clear understanding of why, and to date little research on what is driving this. As with many other characteristics of the community, key indicators are missing or not readily available in the data record inspection asks schools to produce. This hampers understanding of what is really going on.

Attainment data without contextualisation do not tell the whole story of the quality of education in the school. Schools that recruit more affluent pupils will always outperform those who recruit pupils from the most deprived communities. Without taking this into account, the basic comparison tells us nothing about the quality of the school. These points are well known, but value-added, the methodology used by statisticians to account for these differences, has been largely unwound from current Progress-8 measures ([Leckie and Goldstein, 2017](#); [Leckie and Goldstein, 2019](#)). In recognition of the bias inherent in the attainment data, when given too much weight, Ofsted replaced the inspection framework with deep dives into key areas of the curriculum. But this has created a different set of difficulties in passing reliable judgements on the efficacy of curriculum delivery in a context it poorly understands.

Schools do need to be held to account. But there are other more efficient ways of doing this than the current reliance on infrequent short visits by inspectors, geared to delivering final summary judgements on the basis of what they have seen. The current approach disproportionately impacts on schools serving the most disadvantaged communities. This leads to higher teacher turnover in these schools, more likelihood of a change of school management, yet little evidence that this really leads to decisive improvements in a school's performance. A system weighted in favour of schools with the most advantaged student intakes needs to change ([Hutchinson, 2016](#)).

An alternative approach.

In 2021, the BERA expert panel on assessment proposed a different way of overseeing how the school system in England is functioning. We proposed collecting much richer data from a nationally representative sample of schools and their pupils, taking into account their regional context. A broader range of assessment instruments would capture vital aspects of children's learning right the way across the curriculum, not focus on a narrow set of skills that too often leads to teaching to the test. Parent, pupil and teacher questionnaires collected from the sample schools would add a much more rounded picture of the very different contexts in which schools operate, and those aspects of community life inside and outside school that most impact on children's learning. By recruiting a cohort of students to be followed over time a fuller picture would emerge of children's progress and the different routes they take through education.

National Sample data would create a more accurate and fairer picture of the challenges schools face without the need for whole population testing. This would give politicians a richer understanding of where they need to invest. Analysis of the system-level data placed in the public domain would fuel conversations in local communities and clusters of schools about how they compare with others in similar circumstances, the kinds of educational outcomes they value most and what could be done to achieve them. Make the data fully accessible to the research community, and it would allow them to identify and explore new avenues of inquiry. In Germany a similar system exists already: The National Education Panel Survey, which recruits successive pupil cohorts to the sample and follows them through key transition points in their education pathways.

A National Sample would lead to a new role for Ofsted in which inspections would be informed by what is already known about whole system functioning, using its data and findings. Inspections could investigate more closely those issues causing most concern, with reviews targeting different subjects and domains in different years. Reports to parliament would provide feedback for policy development, identifying areas for investment looking across the system as a whole. A National Sample would save money – the costs of testing all pupils as regularly as we do are high, the benefits modest. Indeed, the damage done by tracking pupil progress as minutely as some schools do, using repeated testing to identify flightpaths to test success, would be recognised for what it is – a law of diminishing returns, which narrows the curriculum and reduces opportunities to learn.

To maintain local accountability, schools would be required to develop their own school improvement plans, adapted to local circumstances, but informed by relevant findings from the National Sample. Drawing on the National Sample data would create a much deeper research-informed culture at school level. School improvement plans would enable parents to gain the reassurances on school quality that they need. They would also give parents a much clearer understanding of the school's values and approach to meeting challenges than they currently obtain from an external report on quality produced by Ofsted after a brief visit. Progress on school improvement targets could be overseen on a regional basis, using the successful model of peer-driven school improvement developed through the London Challenge. This could usefully replace a regional commissioning architecture primarily designed to move schools requiring improvement into Academy chains, without the necessary building blocks in place to support their transformation.

In all these ways, moving to a National Sample would create a more reliable, informative and fairer system of intelligent accountability, without any of the downsides of the current approach. It would enable system monitoring appropriate to policymakers' needs. It could be achieved at a considerable saving to the education system as a whole and with the promise of much richer data, it would be better able to support system improvements over the medium to long term.

References

Harmey, S. and Moss, G. (2023) Learning disruption or learning loss: using evidence from unplanned closures to inform returning to school after COVID-19. *Educational Review*, 75:4, 637-656, DOI: [10.1080/00131911.2021.1966389](https://doi.org/10.1080/00131911.2021.1966389)

Hutchinson, J. (2016), *School inspection in England: is there room to improve?*, Education Policy Institute, <https://epi.org.uk/wp-content/uploads/2018/01/school-inspection-in-england-web.pdf>

Leckie, G. and Goldstein, H. (2017) The evolution of school league tables in England 1992–2016: 'Contextual value-added', 'expected progress' and 'progress 8'. *British Educational Research Journal*, 43(2), 193–212 <https://doi.org/10.1002/berj.3264>

Leckie, G. and Goldstein, H. (2019) The importance of adjusting for pupil background in school value-added models: A study of Progress 8 and school accountability in England. *British Educational Research Journal*, 45 (3) pp. 518-537, <https://doi.org/10.1002/berj.3511>

Moss, G., Goldstein, H., Hayes, S., Chereau, B. M., Sammons, P., Sinnott, G., & Stobart, G. (2021). *High standards, not high stakes: An alternative to SATs that will transform England's testing and school accountability system in primary education and beyond*. British Educational Research Association. <https://www.bera.ac.uk/publication/high-standards-not-high-stakes-an-alternative-to-sats>

Moss, G., Bradbury, A., Braun, A., Duncan, S., Levy, R. and Harmey, S. (2022) *Research evidence to support primary school inspection post-COVID*. UCL Institute of Education.

Roffey, S. ASPIRE to a Better Future: The Impact of the Pandemic on Young People, and Options for Schools Post-COVID-19. *Educ. Sci.* 2023, 13, 623. <https://doi.org/10.3390/educsci13060623>

Taylor, C. (2018) The Reliability of Free School Meal Eligibility as a Measure of Socio-Economic Disadvantage: Evidence from the Millennium Cohort Study in Wales, *British Journal of Educational Studies*, 66:1, 29-51, DOI: 10.1080/00071005.2017.1330464

Tuckett, S., Hunt E., Robinson, D., and Cruikshanks, R. (2022) *COVID-19 and Disadvantage Gaps in England 2021*. Education Policy Institute

October 2023