

# Reforming accountability

**Jon Andrews**

January 2025

## About the Education Policy Institute

The Education Policy Institute is an independent, impartial, and evidence-based research institute that promotes high quality education outcomes, regardless of social background. We achieve this through data-led analysis, innovative research and high-profile events.

## Acknowledgements

This work has been kindly supported by AQA.

AQA, or Assessment and Qualifications Alliance, is an independent education charity that sets and marks exams for GCSEs, AS and A-levels, the Extended Project Qualification, Tech-levels and Technical Awards, and other qualifications in the UK. Find more about them [here](#).

## Background

It is now over 30 years since the introduction of school performance tables for schools England and the establishment of the school inspectorate, Ofsted.<sup>1</sup> The accountability system has its roots in the Citizen's Charter initiative of the early 1990s – though local school inspection goes back to the start of the 20<sup>th</sup> century. The aim of the initiative was to ensure that users of public services knew what level of service that they had a right to expect. In education this included the publication of schools results in each area and the regular and independent inspection of schools by a centralised inspectorate.

In 2022/23, Ofsted carried out a total of 3,720 graded inspections, 3,260 ungraded inspections of good and outstanding schools, and 260 urgent or monitoring inspections and over 15,000 primary schools and 3,000 secondary schools had results published in the 2022/23 school performance tables.<sup>2,3</sup> Both have become a significant feature of the school system in England. International evidence shows that accountability which allows direct comparison between institutions can have a positive impact on attainment.<sup>4</sup> England is not alone in its system, around 80 per cent of countries have inspection of some form although how it is carried out varies considerably with some countries relying far more on self-evaluation.<sup>5</sup>

But the school accountability system is not without its challenges or its critics. Even the best considered performance measures risk perverse incentives in terms of school admissions and moves out of schools, unfavourable Ofsted judgements can affect teacher retention and recruitment, and the death of headteacher Ruth Perry brought into sharp focus the potential effects on teacher and school leader wellbeing.

Over the past year, Ofsted carried out its “Big Listen” to gather feedback from the sector and parents, and the new government has announced that it plans to reform performance tables with a new school report card, as well as a new Ofsted framework from 2025/26.

It is therefore an opportune moment to consider how the accountability system is currently operating and reflect on what it is that it should be delivering.<sup>6</sup>

---

<sup>1</sup> Select Committee on Public Administration, ‘Twelfth report: the Citizen’s Charter’, (July 2008)

<sup>2</sup> Ofsted, ‘The annual report of His Majesty’s Chief Inspector 2022/23’, (November 2023). Note that the number of inspections was particularly high because Ofsted received additional funding to catch-up on inspections missed during the Covid-19 pandemic and to ensure all schools are inspected at least once between 2021 and 2025.

<sup>3</sup> <https://www.compare-school-performance.service.gov.uk/>

<sup>4</sup> Annika B. Bergbauer, Eric Hanushek, Ludger Woessmann, ‘Testing with accountability improves student achievement’, (September 2018)

<sup>5</sup> Jane Perryman et al, ‘Beyond Ofsted: an inquiry into the future of school inspection’, (November 2023)

<sup>6</sup> Whilst there are other ways in which schools are “accountable” (such as academy trust funding agreements with the Department for Education) we consider accountability here in terms of performance tables (currently known as [compare school performance](https://www.compare-school-performance.service.gov.uk/)) and Ofsted inspections.

In late 2024 we convened a roundtable to discuss the what the accountability system should be looking to measure and deliver, and we incorporate participants reflections throughout, although this discussion paper does not represent a position of consensus amongst our participants.

# Holding schools to account and providing information to parents about local schools

## The current system was set up for a different era

The current accountability system is largely based on the performance of individual schools and was developed in an era in which most schools were in some form of “local authority control”. The publication of performance data allowed direct comparison between schools and assess local provision as a whole, with accountability ultimately falling on local politicians. Therefore, there was a local democracy angle of schools being accountable to the local authority who were in turn accountable to the local electorate.

But this has become less relevant in the era of mass academisation and the rise of multi-academy trusts. By December 2024, 46 per cent of primary schools, almost half of special and alternative provision schools, and 83 per cent of secondary schools were academies, the majority being part of a multi-academy trusts.<sup>7</sup> This has resulted in different models of governance and, therefore, different models of accountability. Academy trust CEOs are accountable to their board of trustees, while headteachers of local authority maintained schools are accountable to their governing body and then to the local authority.

## Parents and carers are using accountability measures, but not necessarily getting the whole picture

Despite having different governance models, all schools should also be accountable to their local communities and there should also be mechanisms in place for the government to benchmark effectiveness, identify low-performance and have processes in place to intervene when necessary.

Around two-thirds of parents of school-aged children say that Ofsted is a valuable source of information about education in their area and around a third of parents say that the overall Ofsted judgement was a decisive factor in the final choice of school.<sup>8</sup>

But the proximity to school is a decisive factor for over half so, while Ofsted judgements inform parental choices, they are by no means the decisive factor for most. Even when used, analysis suggests that the overall effectiveness grade from Ofsted is generally not a good predictor of what a pupil will go on to achieve at a school.<sup>9</sup>

The accountability system does not adequately reflect how schools are now run. While Ofsted reports comment on the efficacy of individual schools in relation to quality of education and

---

<sup>7</sup> DfE, ‘Open academies, free schools, studio schools, UTCs and academy projects in development’, (December 2024)

<sup>8</sup> YouGov, ‘Ofsted parents annual survey 2021’, (April 2021)

<sup>9</sup> Christian Bokhove et al, ‘How useful are Ofsted inspection judgements for informing secondary school choice?’, (October 2022)

leadership and management, the reality is that decisions around financial management, workforce deployment, and the curriculum offer are now often taken at trust level rather than in individual academies.

Similarly, despite earlier efforts by the Department for Education, performance tables for academy trusts are simply aggregations of their individual schools and do not attempt to measure the impact that the trust has on pupil outcomes. MAT level inspections could be an important step in helping parents understand how schools are run. While the system is now mature enough to develop a MAT level inspection framework, such a framework may need to have some flexibility to recognise the considerable variation in the size and scale of MATs across the country.

The degree to which decision-making is delegated to individual schools is unlikely to be well understood by parents, and as such reports could help them understand how decisions relating to their child's education are being taken.

### **Judging the system as a whole by the same measures as schools has not always worked**

There should be mechanisms in place for the government and the public to assess the quality of the school system as a whole, although we recognise that there are risks to this.

The interpretation of data to make political points is one of those risks. For over a decade, the statistic on the proportion of schools that are rated as good or outstanding was a staple of politicians. But, due to the exemption on inspecting outstanding schools that was in place over that period, those statistics were not reliable. By spring 2024, the average outstanding school had last been inspected over seven years ago.<sup>10</sup> It resulted in a measure in which the proportion of schools rated as good or outstanding increased from 68 to 90 per cent – the validity of which was somewhat questionable, with the flaws in the interpretation known for some time.<sup>11</sup>

Therefore, inspection outcomes do not provide a reliable measure of the performance of the system as a whole. Furthermore, they are based on the performance and management of a school at the time of inspection and do not necessarily reflect the current circumstances. This is a particular issue where schools have been inspected under different governance arrangements or different Ofsted frameworks and is now made more challenging by the removal of the overall effectiveness grade.

Attainment measures, when used for this purpose, are not without weaknesses either. The measures used in performance tables in the late 2000s were on a consistent upward trajectory, but these improvements were not reflected in international measures that allowed more direct comparisons over time.

---

<sup>10</sup> Louis Hodge, 'Ofsted ratings – have standards been improving?', (May 2024)

<sup>11</sup> Jon Andrews, 'Does the claim of 1.9 million more children in good or outstanding schools stack up?', (July 2018)

## There was a deliberate decision not to recognise the context of schools

Whether or not to recognise the different characteristics of pupils and circumstances within a school has long been a subject of debate. Contextual value added was introduced to school performance tables in 2006 to account for different intakes with scores adjusted based on a range of different pupil characteristics such as disadvantage and ethnicity. Michael Gove, Secretary of State for Education between 2010 and 2014, derided the expectation that pupils from certain groups would inevitably achieve lower outcomes, describing such views as the “soft bigotry of low expectations”.<sup>12</sup> Contextual value added was dropped from performance tables in 2011.

The views at EPI’s roundtable were also mixed, with some arguing that it is only fair to acknowledge the different scale of challenges in schools and that some groups make less progress than others while others felt that it was precisely the role of public accountability to highlight where groups were underperforming and where standards needed to be improved. In essence, that the problem should be viewed through the lens of a fairness to children rather than the fairness to schools.

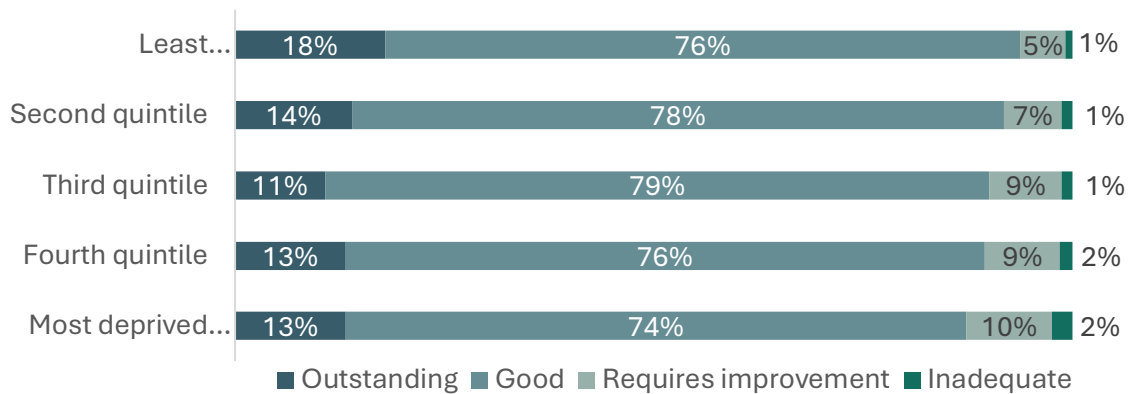
We consider it as a question as to the extent to which measures allow users to identify the effectiveness of schools, so that best practice can be disseminated across the system and weaknesses can be addressed early on.

Current measures of accountability are correlated with the characteristics of pupils in the school – for example, the level of deprivation. Figure 1 shows the latest inspection judgement by the level of deprivation in the area in which the school operates. In the least deprived quintile, 94 per cent of schools were rated as good or outstanding, compared with 87 per cent of schools in the most deprived quintile. There is however some evidence that this relationship has weakened under the current inspection framework.

---

<sup>12</sup> Michael Gove, ‘Speech: the civil rights struggle of our time’, (November 2013).

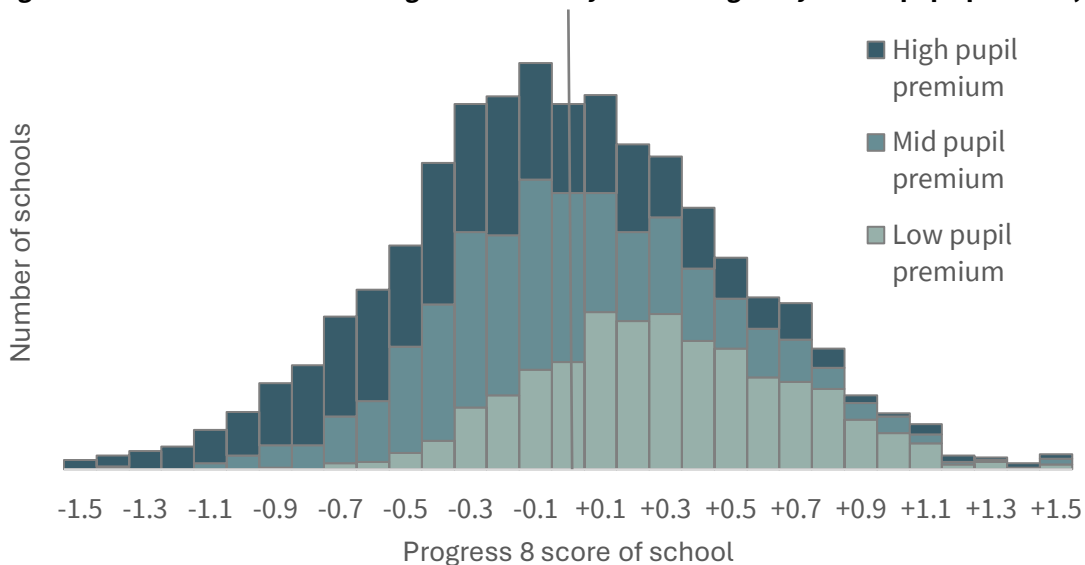
**Figure 1: Latest Ofsted judgement by deprivation (IDACI) of the school, September 2024**



Source: Ofsted "State-funded school inspections and outcomes: management information", September 2024. Includes state-funded nursery, primary, secondary, special, PRU/AP

We can also see this relationship in measures in performance tables. Figure 2 shows the distribution Progress 8 scores by level of disadvantage in the school (as measured by the percentage of pupils eligible for the pupil premium). Around one-in-three of the most disadvantaged schools are labelled as "well below average" by the Department for Education on this measure, compared with just one-in-fifty of the least disadvantaged.

**Figure 2: Distribution of school Progress 8 scores by level of eligibility for the pupil premium, 2023**



Source: Analysis of Department for Education "Compare the performance of schools and colleges in England", key stage 4 2023. Ever6FSM bandings relate to the lowest, middle and highest third of percentage ever6FSM in school. State-funded mainstream schools. Excludes 20 schools outside the range of horizontal axis.

There are also striking differences at key stage 2. In a third of the most disadvantaged schools, less than half of pupils end primary school at the expected standard in reading, writing, and



mathematics compared with in 1 in 11 of the least disadvantaged schools.<sup>13</sup> As a measure of progress that controls for the prior attainment of pupils, Progress 8 goes at least some way to control for the characteristics of the school population – since we know that those characteristics are also correlated with attainment in primary school. But there is still a gap in outcomes between disadvantaged and non-disadvantaged pupils that is equivalent to three-quarters of a grade in each subject at GCSE, even when comparing to pupils with similar prior attainment.<sup>14</sup>

There is an argument that relationships such as these are not a flaw in the measurement system and instead are highlighting the fact that pupils from disadvantaged schools are found disproportionately in less effective schools. For example, schools with higher levels of disadvantaged tend to have higher teacher turnover or struggle to recruit experienced teachers.<sup>15</sup>

But we know that most of the differences in pupil performance are explained by pupil level factors, rather than school factors – around nine-tenths of the variation in pupil outcomes occurs at the pupil level and around one-tenth occurs as variation between schools.<sup>16</sup>

In 2023, disadvantaged pupils outperformed non-disadvantaged pupils in just 3.6 per cent of secondary schools.<sup>17</sup> In other words, pupils who on the face of it have been exposed to, broadly, the same schooling, still achieve lower results. It is therefore inescapable that pupils from disadvantaged backgrounds make, on average, less progress than other pupils and the fact that this occurs within the same school – in almost every school – suggests that this is not simply a function of the school that they attend.

The analysis above is focussed on state-funded mainstream schools. However, Progress 8 is also published for special schools on the same basis. Unsurprisingly, Progress 8 scores for special schools are particularly low or, in a large number of cases, not published (this may be because of small numbers or because pupils do not have prior attainment data). The overwhelming majority of special schools that do have scores published are labelled as ‘well below average’ with scores below minus 1. Such a performance measure is unlikely to tell us anything meaningful about the quality of education and support received by pupils in these schools.

### **The risks of perverse incentives could be better managed**

Even in the best designed system, performance measures are vulnerable to the development of perverse incentives which, when combined with other factors such as funding pressures, can provide a disincentive to schools being inclusive for all pupils.

---

<sup>13</sup> Analysis of Department for Education "Compare the performance of schools and colleges in England", key stage 2 2023. Ever6FSM bandings relate to the lowest, middle and highest third of percentage ever6FSM in school. State-funded mainstream schools with at least 11 eligible pupils.

<sup>14</sup> DfE, 'Key stage 4 performance, academic year 2022/23', (February 2023).

<sup>15</sup> Education Endowment Foundation, 'Three takeaways from the evidence on improving teacher recruitment and retention', (June 2023)

<sup>16</sup> DfES, 'Variation in pupil progress 2023', (July 2023)

<sup>17</sup> Author's analysis of secondary school performance tables 2023.

In a 2019 study we considered the scale of “unexplained exits” – pupils leaving school rolls that did not seem to be instigated by family events.<sup>18</sup> We found that as many as 1 in 10 pupils in the GCSE cohort experienced an exit from secondary school. Three quarters of exits were experienced by vulnerable groups including 1 in 6 pupils from low-income backgrounds, 1 in 6 pupils with an identified special educational need, and 1 in 4 pupils with identified mental health needs.

Recognising the context of a school could go some way to help mitigate some of the risks associated with perverse incentives – essentially, recognising that it is, on average, more difficult to achieve the highest outcomes with more challenging intakes. In particular, school accountability should account for the needs of children with SEND and other additional needs. Ultimately, inflexible systems that do not recognise and adjust for the high level of additional needs in the child population are incompatible with an education system that works for all.

---

<sup>18</sup> Jo Hutchinson and Whitney Crenna-Jennings, ‘Unexplained exits from schools. Further analysis and data by multi-academy trust and local authority’, (October 2019)

# Identifying weaknesses and ensuring systems are in place if the school does not improve

As we have discussed in this brief paper, education policy has moved a long way from school accountability being about simply helping inform interested citizens. Even if schools are not directly accountable to government, accountability measures still provide a mechanism by which government can influence heavily what happens in schools.

## Accountability has been shaped by wider policy priorities

Performance tables have become both a way that ministers can signal their values, e.g. expectations of all pupils should be high regardless of background, and a lever by which they have been able to shape what happens in schools. Academies are not currently bound to follow the national curriculum (though this is likely to change under planned legislation) but primary academies do have the results of assessments in key stage 2 reading, writing and mathematics published every year. In secondary schools, school performance is assessed using the EBacc and its related measure, Progress 8, which means that schools are incentivised to offer and focus on a particular set of academic subjects.

And the reach of Ofsted judgements had also gone beyond accountability to children and parents. Duties introduced in the Education and Adoption Act 2016 meant that local authority maintained schools that were rated as ‘inadequate’ were required by law to become academies and for some time those that were ‘coasting’ – receiving two consecutive ratings that are less than ‘good’ – were also eligible for intervention though would not necessarily face academisation.<sup>19</sup>

## Direct intervention should not be the default position

It is our view that the government should have levers to pull if it becomes apparent that a school or a trust is failing its pupils, but there are two crucial features to this position. First, judgements about poor performance should be based on more reliable and empirical data, that acknowledges the wider circumstances of a school or trust.

Second, the starting point should be support to schools and trusts, not straight to intervention. And it is important that the support does not come from Ofsted. Participants in our roundtable discussion agreed that there needs to be a clearer distinction between accountability and school improvement and that giving Ofsted a role in school improvement undermines their independence.

---

<sup>19</sup> Education and Adoption Act 2016. Planned legislation will remove the duty to issue an academy order for inadequate schools.

## The current system can be a barrier to improvement

There is some evidence that measures of accountability can make it more difficult for schools to improve. Schools that receive a series of ratings that are less than ‘good’ often face more challenging circumstances, higher teacher turnover, and higher levels of disadvantaged pupils and pupils with special educational needs – though poor Ofsted outcomes are only a ‘modest’ contributory factor to lack of improvement over time and many other schools share these characteristics.<sup>20</sup>

Later this year, we will be publishing a follow-up paper on a new model for support and intervention in underperforming schools.

---

<sup>20</sup> Bernardita Munoz-Chereau et al, ‘Stuck schools: can below good Ofsted inspections prevent sustainable improvement?’, (June 2022)

# The building blocks for a better accountability system

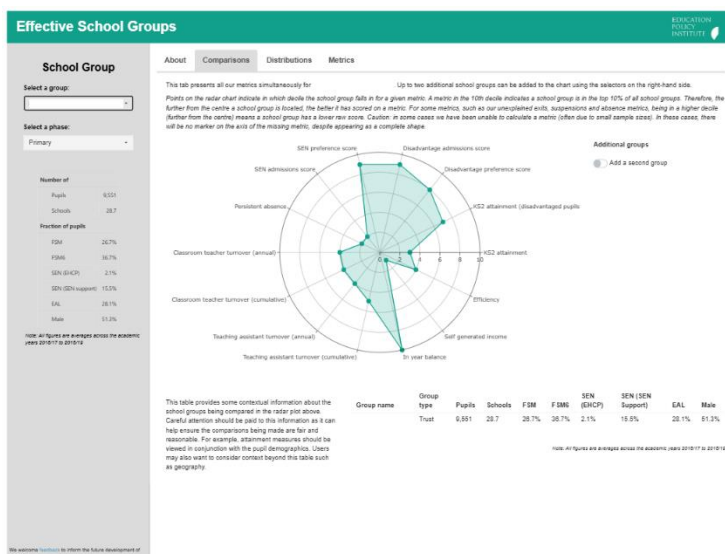
Ofsted’s response to its Big Listen and the earlier announcement of the ending of one-word judgements are both positive steps towards reforming a school accountability system that has made leadership of some of the most challenging schools even more difficult.

As we have set out above, the current accountability system creates perverse incentives, which can often include narrowing of the curriculum, reducing the inclusiveness of admissions practices, and removing pupils from school rolls. Ofsted’s plans to introduce a new focus on inclusion in its criteria for inspection has the potential to help counter these effects and make the school system work better, particularly for pupils from disadvantaged backgrounds or with special educational needs.

## Recommendation 1: Press ahead with report card to give a wider view of school performance; using EPI benchmarking tool as a blueprint

It is right that the Department for Education and Ofsted are reevaluating how school performance is measured. Measures such as pupil performance and inclusion are readily served by existing data collections. Our own benchmarking tool for multi-academy trusts and local authorities, could provide a blueprint for metrics in the new school report card. The tool, which we published in early 2024 includes measures of attainment, progress, inclusion, the school workforce, and school finances.<sup>21</sup>

Figure 3: Example from EPI’s Effective School Groups benchmarking tool



In this approach we have taken “inclusion” to mean how representative a school is of its local community (specifically the rate at which disadvantaged pupils and pupils with special educational needs apply to and are admitted to the school in comparison to other pupils). We

<sup>21</sup> <https://epi.org.uk/effective-school-groups-data-tool/>

recognise that this is just one dimension of inclusion, one that can be directly measured by data, and other aspects are probably better suited to inspection.

### **Recommendation 2: consider how to reflect children and young people’s wellbeing**

There are no current measures of pupil wellbeing available in centrally collected data.

The #BeeWell annual survey of pupil wellbeing has been in operation in Greater Manchester since 2021, and more recently across Hampshire, Isle of Wight, Portsmouth and Southampton.<sup>22</sup>

#BeeWell is an example of how this kind of data could be collected more widely, including providing information privately to schools, and publicly by where young people live. If school support for pupil wellbeing were to be part of the accountability system we would need to consider whether the burden it places on schools is reasonable and whether the information it generates should be publicly available or should remain private to the school (for example, with inspection reflecting whether the school has adequate arrangements in place to monitor and address pupil wellbeing).

### **Recommendation 3: implement plans for MAT inspection but in consultation with the sector and building in acknowledgement of different size and scale of MATs**

We agree that Ofsted should now be able to inspect multi-academy trusts. This reflects the fact that decisions around financial management, workforce deployment, and curriculum choices are now often taken at trust level rather than in individual academies. The degree to which decision making is delegated to individual schools is unlikely to be well understood by parents, and as such reports should help them understand how decisions relating to their child’s education are being taken.

### **Recommendation 4: Focus Ofsted’s role on the quality of teaching and learning and broad and balanced curriculum**

Ofsted should focus on the areas in which inspections can add value rather than on those areas which can be better served by more frequent empirical data. Its primary focus should be on the quality of teaching and learning and the extent to which pupils are offered a broad and balanced curriculum. Its emphasis on inclusion in the new framework should be the extent to which the curriculum and teaching is flexible to the needs of all pupils, in particular those with special educational needs, rather than on how representative a school is of its community (which can be better served by regular data collections). To maintain its independence, Ofsted should not have a role in school improvement.

### **Recommendation 5: separate safeguarding from other elements of the accountability system**

Placing safeguarding within a high-stakes framework that is currently targeted towards more disadvantaged schools can have unintended consequences. Some schools might lack sufficient

---

<sup>22</sup> A collaboration between the University of Manchester, the Gregson Family Foundation and Anna Freud. <https://beewellprogramme.org/>

focus, whereas others, by virtue of their intake, receive regular feedback. The high-stakes nature of inspection may also have the unintended consequence of incentivising people to 'hide' issues rather than acknowledge them and seek help. The nature of safeguarding risks is constantly evolving, and all schools require regular training and feedback.