



WHAT COMES NEXT

A FAIRER APPROACH TO PRIMARY ASSESSMENT

**Conor Carleton and
Ellie Harris**

October 2026

ABOUT IPPR

IPPR, the Institute for Public Policy Research, is an independent charity working towards a fairer, greener, and more prosperous society. We are researchers, communicators, and policy experts creating tangible progressive change, and turning bold ideas into common sense realities. Working across the UK, IPPR, IPPR North, and IPPR Scotland are deeply connected to the people of our nations and regions, and the issues our communities face.

We have helped shape national conversations and progressive policy change for more than 30 years. From making the early case for the minimum wage and tackling regional inequality, to proposing a windfall tax on energy companies, IPPR's research and policy work has put forward practical solutions for the crises facing society.

IPPR
8 Storey's Gate
London
SW1P 3AY
E: info@ippr.org
www.ippr.org
Registered charity no: 800065 (England and Wales),
SC046557 (Scotland)

This paper was first published in October 2026. © IPPR 2026

The contents and opinions expressed in this paper are those of the authors only.

CONTENTS

Summary.....	5
1. The case for reform.....	6
The background and purpose of this paper	6
Primary school assessment within the wider assessment and accountability architecture.....	7
Methods.....	7
Defining disadvantage.....	7
2. Immediate changes from the 2027/28 academic year	8
Recommendation 1: The Department for Education should establish an independent ‘statutory accessibility and bias audit’ for all key stage 2 assessments	8
Recommendation 2: The Department for Education should remove the Spelling, Punctuation and Grammar (SPaG) test	9
Recommendation 3: The Department for Education should change how results are published and reported	9
Recommendation 4: The Department for Education should introduce a competency framework for pupil results in place of scaled scores	11
3. A blueprint for further change.....	12
Recommendation 5: The Department for Education should introduce a random sampling approach to assessment across years 3–6	12
References.....	14

ABOUT THE AUTHOR

Conor Carleton is a children and young people's expert, providing policy consultancy for IPPR.

Ellie Harris is a principal research fellow and head of children and young people's policy at IPPR.

ACKNOWLEDGEMENTS

The authors would like to thank the experts who gave up their time to speak with us. This includes Tim Oates, Pete Whitehead, Agnes Fitzpatrick, Christopher Such, Professor Gemma Moss, Professor Dominic Wyse, Professor Mick Waters, Loic Menzies, and Kirsti Hadley. While their engagement was invaluable in the policy development process, the recommendations are the views of the authors alone.

The authors would also like to thank More Than A Score for generously supporting this work. More Than A Score is a coalition of parents, teachers, school leaders and education experts.

www.morethanascore.org.uk



Download

This document is available to download as a free PDF and in other formats at:
<http://www.ippr.org/articles/what-comes-next-primary-assessment>

Citation

If you are using this document in your own writing, our preferred citation is:
Carleton C and Harris E (2026) *What comes next: A fairer approach to primary assessment*, IPPR.
<http://www.ippr.org/articles/what-comes-next-primary-assessment>

Permission to share

This document is published under a creative commons licence:
Attribution-NonCommercial-NoDerivs 2.0 UK
<http://creativecommons.org/licenses/by-nc-nd/2.0/uk/>
For commercial use, please contact info@ippr.org



SUMMARY

England's primary school accountability and assessment system is not working well for disadvantaged children. By the end of primary school, disadvantaged children are, on average, 10 months behind their peers, and that gap is growing (EPI 2025). In 2026, 4 per cent of disadvantaged pupils met the higher standard in reading, writing and maths (combined), while 11 per cent of other pupils met the higher standard (DfE 2026b). In other words, non-disadvantaged pupils are nearly three times more likely to achieve the highest marks in SATs than their disadvantaged peers.

Change is needed. Building on the evidence outlined in our IPPR paper *“Primary is what comes first”: How end of key stage 2 exams impact disadvantaged children* (Carleton et al 2026) and a series of roundtables and interviews with experts, in this paper we outline recommendations to build a more equitable, accessible, functional and reliable system.

We set out four recommendations for immediate implementation from the 2027/28 academic year.

- **Recommendation 1:** The Department for Education should establish an independent ‘statutory accessibility and bias audit’ for all key stage 2 assessments.
- **Recommendation 2:** The Department for Education should remove the Spelling, Punctuation and Grammar (SPaG) test.
- **Recommendation 3:** The Department for Education should change how results are published and reported by:
 - ending the publication of school-level key stage 2 results, which would also end the publication of school-level league tables
 - using rolling averages (for instance, three- or five-year averages) of attainment data from the department and from Ofsted for internal monitoring and inspection.
- **Recommendation 4:** The Department for Education should introduce a competency framework for pupil results in place of scaled scores.

However, reform must go further to build a truly equitable system that properly supports disadvantaged pupils. Therefore, we also set out a further recommendation to act as the foundation of a longer-term, more significant reform to primary school assessment.

- **Recommendation 5:** The Department for Education should introduce a random sampling approach to assessment across years 3–6.

1.

THE CASE FOR REFORM

THE BACKGROUND AND PURPOSE OF THIS PAPER

This paper is the second in our series exploring the impact of end of key stage 2 assessments (commonly referred to as standard assessment tests; SATs) on disadvantaged pupils. The first, *“Primary is what comes first”: How end of key stage 2 exams impact disadvantaged children*, articulated the problem, presenting the results of a nationally representative poll of primary school senior leaders and in-depth qualitative interviews with primary school senior leaders and year 6 teachers (Carleton et al 2026). The paper argued that SATs do not work well enough for disadvantaged pupils, in three areas.

1. **Perverse incentives and curriculum narrowing.** The primary school assessment and accountability framework incentivises schools to narrow their curriculum and focus on tested subjects (‘teaching to the test’), particularly during SAT preparation periods. The loss of wider curriculum time and enrichment most significantly affects disadvantaged pupils, who are often less likely to have access to wider or extra-curricular opportunities, such as art, drama and music, outside school (Carleton et al 2026). Of those polled, 93 per cent felt that preparation for SATs has an impact on the breadth of curriculum in primary schools, with 77 per cent reporting more time being spent on tested subjects when preparing for SATs.
2. **An accountability framework that does not reflect school context.** Primary schools are judged against a single national benchmark – without meaningful account taken of the communities they serve and the cohorts they have each year. We found this can obscure genuine school performance, especially in disadvantaged areas, and leave senior leaders facing scrutiny that they feel bears little relationship to what their school has delivered.
3. **Increased teacher pressure that makes its way into the classroom.** SATs-related accountability pressures drive teacher stress, affect pedagogy and are a factor in teachers’ decisions to leave the profession. These effects are felt most acutely in schools with the least capacity to absorb them, with consequences for the quality of teaching and pastoral support that teachers can provide pupils, and impacts falling principally on disadvantaged children. Of those polled, 79 per cent of senior leaders in the most disadvantaged primary schools reported a negative impact on day-to-day school life due to SATs.

Only 7 per cent of senior leaders polled believed that SATs should be kept as they are. As a result, the paper found a strong case for reform of the primary assessment and accountability system that follows the tests.

This paper represents a first step towards that reform. It outlines immediate steps that the government can take to improve the current system and mitigate against some of the areas in which the system is currently falling down. However, minor reform will not be sufficient in the long term. This paper therefore also articulates a blueprint for more significant reform, providing an outline for what a more equitable, accessible and functional system could look like in practice in the medium term.

PRIMARY SCHOOL ASSESSMENT WITHIN THE WIDER ASSESSMENT AND ACCOUNTABILITY ARCHITECTURE

Primary school assessment does not operate in a vacuum and the issues the recommendations of this paper are designed to address are not exclusive to end of key stage 2 exams. In primary school, the performance of the phonics screening check and the reception baseline assessment are contested (Goldstein et al 2018, Wyse et al 2022). In secondary school, the reform or replacement of GCSEs remains a topic of debate, as does the utility of assessing school performance using a Progress 8 measure, which many see as faulty (Leckie and Goldstein 2019, Menzies et al 2023, Harris et al 2025).

This paper deliberately focusses on key stage 2 assessment. Rethinking England's approach to phonics, reception baseline assessments or GCSEs is, therefore, out of scope. Yet primary school assessment cannot be fully understood, or indeed fixed, in isolation from the wider school assessment and accountability architecture, and more work is needed to design the large-scale, long-term change that this architecture needs. This paper's recommendations should be read as a necessary and deliverable first step within that wider piece of work on reform.

METHODS

Building on the evidence set out in *“Primary is what comes first”* (Carleton et al 2026), the recommendations outlined in this paper are the product of a series of roundtables and interviews with nine experts from across primary teaching, assessment and accountability. While these roundtables and interviews were invaluable in the policy development process, the recommendations are the views of the authors alone.

DEFINING DISADVANTAGE

In the qualitative interviews and polling that were carried out as part of this work, and which were reported in *“Primary is what comes first”*, ‘disadvantage’ was defined as:

“pupils who face additional barriers to learning compared with their peers, including pupils eligible for free school meals, pupils with special educational needs and disabilities, pupils with English as an additional language, pupils with long-term health conditions and pupils from ethnic minority backgrounds”.

For key stage 2 assessment, the Department for Education (DfE) defines disadvantaged pupils primarily as those eligible for free school meals in the previous six years or pupils who have been in care (DfE 2026a).

2.

IMMEDIATE CHANGES FROM THE 2027/28 ACADEMIC YEAR

In this chapter we set out four recommendations for immediate implementation from the 2027/28 academic year.

RECOMMENDATION 1: THE DEPARTMENT FOR EDUCATION SHOULD ESTABLISH AN INDEPENDENT ‘STATUTORY ACCESSIBILITY AND BIAS AUDIT’ FOR ALL KEY STAGE 2 ASSESSMENTS.

The Department for Education (DfE) should establish an annual independent ‘statutory accessibility and bias audit’, commissioned by the department and delivered independently of the Standards and Testing Agency (STA), to review key stage 2 test papers against defined criteria, including:

- cultural bias
- linguistic complexity
- cognitive load
- accessibility for pupils with additional needs.

Findings should be published annually, and the STA should be required to respond, with findings directly feeding into the design of the following academic year’s tests.

End of key stage 2 test content routinely assumes shared cultural experience and background knowledge that pupils do not equally have (Carleton et al 2026). In previous years, tests have set pupils writing tasks centred on activities such as skiing or foreign travel. Tests that do not distinguish between pupils who do not have the relevant experience or who lack the underlying skills, and pupils who do have these things, are not measuring what they are designed to measure. Our polling found that 67 per cent of senior leaders do not think that schools have the support and resources they need to ensure that all pupils experience SATs fairly and equally, and 82 per cent believe SATs disproportionately affect disadvantaged pupils.

Designing tests around assumed cultural touchpoints that not all pupils can relate to has an impact on disadvantaged children most significantly. Current access arrangements – such as giving pupils extra time, a separate room or a reader – are centred on how tests are administered. Whether the underlying content or format of the tests is fair in the first place is not currently adequately addressed when the accessibility of tests is evaluated.

The evaluation of test accessibility must also consider pupils with special educational needs and disabilities. In our qualitative interviews, senior leaders described access arrangements such as those mentioned above as “cosmetically inclusive but substantively inadequate” (Carleton et al 2026). One leader described a pupil with special educational needs who had been performing well on practice papers, but who suffered a panic attack under exam conditions and was unable to finish the exam. The school was still required to submit the incomplete paper. Of senior leaders polled, 90 per cent said that

pupils with special educational needs and disabilities experience the greatest negative impact from end of key stage 2 exams.

The STA has its own internal test and development framework, which consists of an ‘inclusion panel’ that reviews tests for accessibility to pupils with special educational needs and disabilities, and an external ‘equity advisor review group’ that reviews proposed reading tests to check “the subject matter is appropriate for all pupils” (STA 2024). It also runs ‘differential item functioning’ analysis to examine where performance differs by group (STA 2024). However, these mechanisms have not been sufficient to remove the structural disadvantages that some groups of children face. No fully independent, annual, public-facing audit exists to specifically assess the cultural bias, linguistic complexity and content accessibility of end of key stage 2 exams. The establishment of such an audit would create a low-cost, long-standing accountability mechanism to ensure that test content is appropriate and accessible to all pupils.

RECOMMENDATION 2: THE DEPARTMENT FOR EDUCATION SHOULD REMOVE THE SPELLING, PUNCTUATION AND GRAMMAR (SPaG) TEST.

The Spelling, Punctuation and Grammar (SPaG) test – sometimes referred to as the Grammar, Punctuation and Spelling (GPS) test – is not fit for purpose. Rather than have it reformed or replaced, the DfE should completely remove it, with the focus on spelling, punctuation and grammar kept through the existing English assessments. Alongside this, the DfE should review relevant aspects of the National Curriculum to remove grammatical terminology content (for instance, ‘fronted-adverbial instruction’) where there is limited evidence of benefit to pupils’ writing (Wyse et al 2026).

While the DfE’s 2025 curriculum and assessment review acknowledged that “the [SPaG] test has led to a disproportionate amount of classroom time being spent on teaching grammatical constructs in isolation, often at the expense of a more integrated approach” (DfE 2025: 130), it suggests amending the test to keep “those elements of the current test that encourage effective teaching” (ibid), but it gives no further detail.

Rather than reforming the SPaG test or replacing it with something different, the DfE should go further and completely remove the test. The test assesses grammar, punctuation and spelling in isolation from pupils’ own writing. While pupils’ grammatical understanding is associated with the quality of their writing, evidence suggests grammar teaching benefits pupils’ writing ability mainly when it is contextualised, or taught together (Marjokorpi 2023). Where it is taught as abstract rules and terminology, its effect on writing quality is weak (Wyse 2001, Andrews et al 2006). Time spent studying abstract grammar rules in ways that do not improve pupils’ comprehension, understanding or application of English means they have less time to spend on other aspects of the curriculum.

Removing the test does not remove the need for schools to teach spelling, punctuation and grammar, nor does it remove the usefulness of a diagnostic mechanism for assessing school performance in this area. The assessment of spelling, punctuation and grammar should be kept within the wider key stage 2 exam assessments, particularly the English assessment. Removing the SPaG test removes the standalone, high-stakes, faulty test, not the imperative to teach these concepts.

RECOMMENDATION 3: THE DEPARTMENT FOR EDUCATION SHOULD CHANGE HOW RESULTS ARE PUBLISHED AND REPORTED.

The DfE should change how results are published and reported by:

- ending the publication of school-level key stage 2 results, which would also end the publication of school-level league tables
- using rolling averages (for instance, three- or five-year averages) of attainment data from the department and from Ofsted for monitoring and inspection.

Our previous paper in this series highlighted two issues with the current approach of publishing school-level data (Carleton et al 2026).

First, school-level data creates inter-school competition and high-stakes accountability. Headteachers and classroom teachers are judged on their results, which can drive perverse behaviours around admissions, ‘teaching to the test’ and competition between schools rather than collaboration. Given the attainment gap between disadvantaged and non-disadvantaged pupils, this disproportionately has an impact on schools working in the most deprived areas.

As one primary school senior leader reflected in an interview:

“the local authority will hold you to account. The DfE will hold the local authority to account. Ofsted will challenge you around that data picture when it’s not true.”

Primary school senior leader

Furthermore, it is not only what happens within the school walls that defines pupils’ outcomes: other parts of pupils’ lives also shape their performance in assessments. Reporting outcome data at a school level can mean that people consider the data as an accurate measure of a school’s effectiveness when in reality it reflects what is happening in pupils’ lives as well.

Second, failing to account for high levels of cohort variability (that is, the changes in ability between one year’s cohort and another) encourages schools to focus on ‘cuspy’ children – children who, with some extra attention, may meet the expected standard – but at the expense of both other pupils and a varied curriculum. One primary school senior leader we spoke to articulated how this plays out in how their results are received:

“I’ve got 45 per cent SEND [children with special educational needs and disabilities] in my current year 6, and only 7 per cent next year. So next year my results are going [to look very different] and everyone is going to go ‘Isn’t that amazing, the job they’ve done?’ But it’s just a different cohort of children.”

Primary school senior leader

Competition between schools does not necessarily drive standards and, outside of major cities, parents often do not have meaningful choice when it comes to selecting schools for their children (Gorard 2014, Burgess et al 2019). The publication of school-level data and school-level league tables raises the stakes of key stage 2 assessment for primary schools, creating perverse incentives for schools, which has a negative impact on disadvantaged pupils (Carleton et al 2026). The need to attract pupils to secure school funding, which is vital to offer the support pupils need, and the risk that they may lose their job if results fall below what is expected, raise the stakes for headteachers even further.

Ending this practice would retain the important features of the accountability system while removing some of the incentives for schools to act in ways counter to their own pupils’ interests. Many of the biggest challenges facing schools and the wider education sector – for instance in relation to attendance or the provision of support for pupils with special educational needs and disabilities – require a collaborative approach across schools, other services and local areas (Centre for

Young Lives and Mission 44 2026). Primary school assessment and accountability infrastructure should therefore be designed to encourage collaboration between schools, not competition.

The introduction of a rolling average does not remove accountability or change the work of the DfE or Ofsted. Instead, it makes sure schools are judged on sustained performance, and not on a single-year snapshot. A school may do exceptionally well in its end of key stage 2 exams one year, and much less well the following year with the same approach and teachers. As alluded to above, far from the teaching worsening, this may reflect the differing abilities of the two cohorts. While natural cohort variation is less of a factor at the national level (each year, some schools do better, some do worse), this does not occur at the school level. A rolling average still allows the DfE and Ofsted to monitor the direction of travel of a school but smooths out some of the irregularities that come from natural cohort variation (Jerrim and Menzies 2019, Menzies and Jerrim 2020). Jerrim and Menzies (2019) suggest that this benefit is likely to be greatest in primary schools, where cohorts are smallest.

RECOMMENDATION 4: THE DEPARTMENT FOR EDUCATION SHOULD INTRODUCE A COMPETENCY FRAMEWORK FOR PUPIL RESULTS IN PLACE OF SCALED SCORES.

Rather than using a binary approach for key stage 2 results, the DfE should introduce an asset-based competency framework. This competency framework should consist of a set of progressively more sophisticated statements describing positively what a pupil *can* do. Under the framework, no pupil would receive a single number or a pass/fail label, removing the harm that a 'fail' label causes.

Results would provide information about what children are capable of, providing them, their parents and their future teachers with information about their strengths, competencies and areas for further development. For instance, in place of the binary met/not met expected standard in maths, pupils would receive statements summarising their abilities in fractions, decimals, percentages, statistics, geometry and more. Secondary schools would then have useful data on and insight into their incoming cohorts.

Currently, at the end of key stage 2, pupils are told whether they have met the 'expected standard' at primary school. If they do not meet this standard, it frames the result as something they have personally failed. This is despite SATs nominally measuring school performance (Carleton et al 2026) and that, regardless of the result, they will make their transition from primary to secondary school anyway. The current binary approach also treats a pupil who narrowly misses the expected standard the same as a pupil who is operating much further below the threshold. There is no evidence that ability grouping itself harms disadvantaged pupils (Jerrim et al 2025), but telling pupils that they have "not met the expected standard" at a key transition point can be damaging for their confidence and self-esteem at a moment of important developmental transition (Carleton et al 2026).

Furthermore, the utility of raw key stage 2 pupil performance data for secondary schools is a subject of debate. Teacher Tapp research suggests that while secondary schools that do use attainment groupings (setting or streaming)¹ from year 7 (the first year of secondary school) initially base these off key stage 2 exams, "[secondary] school-based assessments and teacher judgements of students' suitability for specific classes play an increasingly significant role" from the beginning of year 8 onwards (Allen 2025).

¹ Setting and streaming are methods used to group pupils by attainment. Setting involves grouping pupils by attainment for specific subjects, and streaming involves grouping pupils for most, if not all, of their classes.

3.

A BLUEPRINT FOR FURTHER CHANGE

The changes set out in the previous chapter would make the current accountability and assessment system fairer and more reliable in the immediate term. Taken together, they would reduce biases in the testing approach and lower the stakes of assessment, while retaining important accountability infrastructure and producing better-labelled results for pupils. However, these changes would leave the underlying architecture of the system unchanged. The fundamentals of this architecture – a single, high-stakes assessment window, concentrated at the end of key stage 2 – are the primary driver of the perverse incentives identified in our previous paper in this series (Carleton et al 2026).

The impacts of curriculum narrowing and a loss of enrichment opportunities fall disproportionately on disadvantaged pupils (ibid). Removing the publication of school-level data and introducing rolling averages for accountability might make the resulting judgements fairer, but neither would remove the incentives for schools to ‘teach to the test’, narrow their curriculum to focus on tested subjects and reduce their extra-curricular offers. Furthermore, the disadvantage gap is already substantial by the end of key stage 2 (EPI 2025). An assessment architecture built around a week of high-stakes exams in the final term of primary school does little to respond to this gap earlier, before it has augmented so substantially.

Therefore, while the changes set out above would represent a meaningful step forward in the short term, more significant, structural change is needed to build an improved, more equitable system that both meets pupils’ needs and holds schools accountable for the quality of their teaching.

A new framework is needed that performs the following four functions:

- lowers the stakes of any single assessment while retaining the validity and reliability of assessments
- is accessible to all pupils and schools, but particularly those facing the greatest disadvantage – for instance, pupils with special educational needs and disabilities, looked-after children, children eligible for free school meals and children from ethnic minority backgrounds
- captures genuine school and system performance, including the early identification of where a school needs support, without driving inter-school competition
- treats primary school as a developmental stage in its own right and not simply as preparation for secondary school.

Our fifth recommendation, therefore, is as follows.

RECOMMENDATION 5: THE DEPARTMENT FOR EDUCATION SHOULD INTRODUCE A RANDOM SAMPLING APPROACH TO ASSESSMENT ACROSS YEARS 3–6.

The single year 6 test window should be replaced with a random sampling model spanning years 3–6, across a broader range of subjects. While assessment would

nominally be grouped by subject (for instance, history or science), the assessment tool would be designed to assess whether a pupil is developing appropriately for their age alongside their subject-related competencies.

The DfE would issue assessments to a sample (for example, 25 per cent) of schools each year, across the whole of key stage 2, testing a broad range of pupil knowledge and underlying child development. This data would be used to identify national, regional and local trends, as well as monitor cohort trends for specific groups of pupils, for example those in receipt of free school meals. The DfE should set a maximum time limit between assessments (for instance, three years). The DfE and Ofsted could use school-level data to monitor performance, with schools identified as consistently under-performing being offered additional resources and support.

Which year group and subject are sampled, and in which schools, would be unpredictable from year to year. In a particular year, one school in an area may have their year 4 cohort assessed in geography, while a neighbouring school has their year 6 cohort assessed in maths and music. A third school, meanwhile, may not be assessed at all. This sampling would ensure schools are not incentivised to narrow the curriculum or teach to the test, as they cannot predict the content of assessments.

This framework would provide a genuine solution to tackle the issues of the current system, while retaining all the important accountability functions the system needs to function appropriately. Because any subject, and any year group, could be assessed, schools would be incentivised to maintain a broad and varied curriculum, rather than narrowing towards specific tested content in English and maths. A system based around child development, furthermore, captures genuine developmental progress rather than rewarding narrow test preparation.

The sampling approach across three different year groups would provide data for the DfE to see how the primary system is functioning at a system level. Moving to a random sampling model would mean Progress 8 could not exist in its current form, however. Progress 8 is itself a contested measure (Leckie and Goldstein 2019), and the DfE should work on its reform, alongside reform of the secondary school assessment and accountability infrastructure. Whether, and how, sampled or aggregated data from this framework could inform a reformed measure of secondary school accountability is a question for that further work.

Taken together, the reforms set out in this paper address the challenges we identified previously with the current key stage 2 accountability and assessment arrangements (Carleton et al 2026). But building and quality-assuring this system at primary school level would be a substantial undertaking, and further work is required to develop the detail needed to support implementation. The purpose of this paper has been to set out the principles for a more equitable, sampling-based system, and begin the work of articulating what this could look like in practice.

REFERENCES

- Allen R (2025) 'Organisation of classes in secondary schools', Teacher Tapp blog, 6 January 2025. <https://teachertapp.com/articles/organisation-of-classes-in-secondary-schools>
- Andrews R, Torgerson C, Beverton S, Freeman A, Locke T, Low G, Robinson A and Zhu D (2006) 'The effect of grammar teaching on writing development', *British Educational Research Journal*, 32(1): 39–55. <https://doi.org/10.1080/01411920500401997>
- Burgess S, Greaves E and Vignoles A (2019) 'School choice in England: Evidence from national administrative data', *Oxford Review of Education*, 45(5): 690–710. <https://doi.org/10.1080/03054985.2019.1604332>
- Carleton C, Harris E and Morjaria A (2026) "Primary is what comes first": How end of key stage 2 exams impact disadvantaged children, IPPR. <http://www.ippr.org/articles/primary-is-what-comes-first>
- Centre for Young Lives and Mission 44 (2026) *Ambitious about inclusion: A guide to implementing the schools white paper*. https://media.cfyl.org.uk/6a3169dd5d331789446a80ca_Ambitious-About-Inclusion-Report_16-June-2026.pdf
- Department for Education [DfE] (2025) *Curriculum and assessment review: Building a world-class curriculum for all: Final report*, GOV.UK. https://assets.publishing.service.gov.uk/media/690b96bbc22e4ed8b051854d/Curriculum_and_Assessment_Review_final_report_-_Building_a_world-class_curriculum_for_all.pdf
- Department for Education [DfE] (2026a) 'Key stage 2 attainment: methodology', GOV.UK website, 10 September 2026. <https://explore-education-statistics.service.gov.uk/methodology/key-stage-2-attainment>
- Department for Education [DfE] (2026b) 'Key stage 2 attainment: academic year 2025/26, provisional', GOV.UK website, 10 September 2026. <https://explore-education-statistics.service.gov.uk/find-statistics/key-stage-2-attainment/2025-26-provisional>
- Education Policy Institute [EPI] (2025) *Annual report 2025*. <https://epi.org.uk/annual-report-2025-disadvantage>
- Goldstein H, Moss G, Sammons P, Sinnott G and Stobart G (2018) *A baseline without basis: The validity and utility of the proposed reception baseline assessment in England*, British Educational Research Association. <https://www.bera.ac.uk/publication/a-baseline-without-basis>
- Gorard S (2014) 'The link between academies in England, pupil outcomes and local patterns of socio-economic segregation between schools', *Research Papers in Education*, 29(3): 268–284. <https://doi.org/10.1080/02671522.2014.885726>
- Harris E, O'Brien C, Gill K, Menzies L, Poku-Amanfo E, Rooney K and Graham J (2025) *Who is losing learning? Finding solutions to the school engagement crisis*, IPPR. <https://www.ippr.org/articles/who-is-losing-learning-solutions>
- Jerrim J, Lopez-Agudo LA and Marcenaro-Gutierrez OD (2025) 'Does within-school between-class ability grouping harm the educational outcomes of socio-economically disadvantaged children? International evidence', *European Sociological Review*, 41(5): 722–736. <https://doi.org/10.1093/esr/jcae050>
- Jerrim J and Menzies L (2019) 'Accountability reformed: The case for multi-year measures', Centre for Education and Youth blog, 29 November 2019.
- 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): 518–537. <https://doi.org/10.1002/berj.3511>
- Marjokorpi J (2023) 'The relationship between grammatical understanding and writing skills in Finnish secondary L1 education', *Reading and Writing*, 36(10): 2605–2625. <https://doi.org/10.1007/s11145-022-10405-z>

- Menzies L and Jerrim J (2020) *Improving headline school performance measures: Multi-year averages should be used in school league tables if they return in 2021*, Centre for Education and Youth, UCL Institute of Education and Pearson. <https://cfey.org/wp-content/uploads/2020/07/Improving-headline-school-performance-measures.-Multi-year-averages-should-be-used-in-school-league-tables.-Menzies-and-Jerrim-2020-1.pdf>
- Menzies L, Yates W and Huband-Thompson B (2023) *Balancing act: Navigating the tensions in our school system*, IPPR. <https://www.ippr.org/articles/balancing-act-tensions-in-our-school-system>
- Standards and Testing Agency [STA] (2024) 'Test development handbook 2024', GOV.UK website, last updated 8 May 2025. <https://www.gov.uk/government/publications/national-curriculum-test-development-handbook/test-development-handbook-2023>
- Wyse D (2001) 'Grammar for writing? A critical review of empirical evidence', *British Journal of Educational Studies*, 49(4): 411–427. <https://doi.org/10.1111/1467-8527.t01-1-00185>
- Wyse D, Bradbury A and Trollope R (2022) *Assessment for children's learning: A new future for primary education*, Independent Commission on Assessment in Primary Education. https://www.icafe.org.uk/reports/NEU2762_ICAPE_final_report_A4_web_version.pdf
- Wyse D, Aarts B, Anders J, de Gennaro A, Dockrell J, Manyukhina Y, Sing S and Torgerson C (2026) 'Teaching grammar and writing: A randomised controlled trial and implementation process evaluation of Englicious', *Journal of Writing Research*, 18(1). <https://doi.org/10.17239/jowr-2026.18.01.04>



GET IN TOUCH

For more information about the Institute for Public Policy Research, please go to www.ippr.org

You can also call us on +44 (0)20 7470 6100,
e-mail info@ippr.org or tweet us @ippr

Institute for Public Policy Research

Registered Charity no. 800065 (England & Wales),
SC046557 (Scotland), Company no, 2292601 (England & Wales)