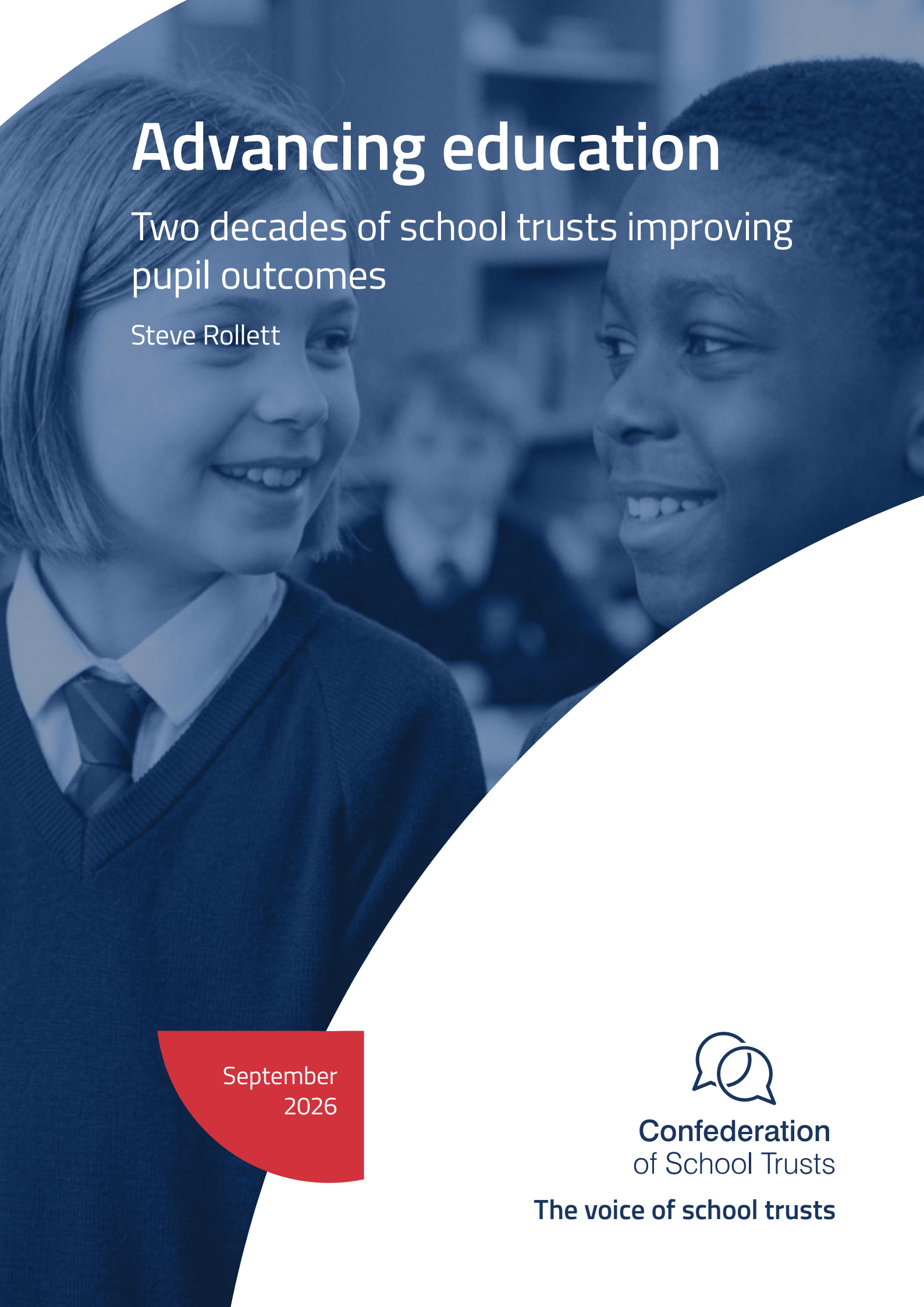


Advancing education

Two decades of school trusts improving pupil outcomes

Steve Rollett

September 2026

Confederation of School Trusts

The voice of school trusts

Confederation of School Trusts

The Confederation of School Trusts is the national organisation and sector body for academy trusts. Our members educate more than four million children through the dedicated work of their 600,000 staff.

We believe school trusts offer the best model for improving education because they uniquely unite schools in organisations with a strong, single purpose: education. This encourages the sharing of expertise, ideas, and resources to give pupils a better start in life.

We advocate for trusts by helping politicians, policy makers, and the wider public better understand their power to help children. We support trusts through practical advice, strategic guidance, and consultancy. We connect them with each other, and with expert thinking and research.

Published in September 2026 by the Confederation of School Trusts. Registered offices c/o Stone King LLP, Boundary House, 91 Charterhouse Street, London, EC1M 6HR. Registered charity number 1107640. © 2026 Confederation of School Trusts. All rights reserved. No part of this document may be reproduced or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without prior written permission of CST.

How to cite this publication: Rollett, S (2026). *Advancing education: Two decades of school trusts improving pupil outcomes* London: Confederation of School Trusts

Foreword

Since the first academy schools – in the form we recognise them today – opened their doors in 2002, there has been much debate about their power to improve education.

Firm conclusions on the benefits for pupils have been vexed by changes over time to assessment and inspection.

This paper examines nearly two decades of national data, from 2007 to 2025, to ask a straightforward question: did schools improve their position in national attainment rankings after joining academy trusts, relative to schools that remained with local authorities?

Across the analyses, the answer points consistently in the same direction. Schools in trusts have made substantial gains in their national ranking relative to their peers. The gains are larger for schools previously judged to be failing (sponsored academies) and for schools serving the most disadvantaged communities.

The analysis also points to the importance of the school trust functioning as a group: schools joining such groups performed better than comparable schools operating on their own, whether as standalone academies or as local authority maintained schools.

These findings matter for both the assessment of past policy and decisions about the future shape of the system.

They provide good evidence that the creation of academies by the Labour government in the early 2000s, and the subsequent extension of that policy, has on balance improved educational outcomes, particularly where underperformance had become entrenched.

They also strengthen the case that the continued movement – set out in the 2026 Schools White Paper – towards a system in which all schools belong to a high-quality school trust is the right one, and consistent with the direction of the evidence.

Leora Cruddas CBE
Chief Executive, Confederation of School Trusts

Executive summary

Core findings

The improved attainment in trust schools

Equivalent to **5,400** more primary children reaching the expected standard in reading, writing, and maths every year

3x the primary pupil-outcome gain in the most disadvantaged communities, against the national trust figure

Equivalent to **500,000** more pupils gaining 1-2 GCSE grades higher in a single subject

OR

Equivalent to **16,000** more grade 4+ GCSEs in English and maths

Increase in national attainment rankings for trust schools

+5.6 percentile places up the national attainment rankings for the average trust school – other schools stood still

+4.0 of those percentile places associated with trust membership itself, on the strongest available statistical tests

The reform of England's school system has been remarkable: from the first mainstream academy opening in 2002, 64% of pupils at state-funded schools now attend academies; for secondary age pupils, the shift is even more pronounced at 84%. Drawing on national data from 2007 to 2025 – every state school in England – this research asks whether schools that joined trusts improved more than schools that did not.

The answer is consistent across every test applied: Schools that joined trusts rose substantially in national attainment rankings – on average from below the national middle to above it.

Schools that stayed with local authorities, on average, did not improve their rankings. The analysis in this paper is not a criticism of local authority schools. Nor is it evidence that local authority schools got worse in real terms over this period. Rather, the evidence in this paper finds that trust schools improved their attainment ranking while local authority schools, on average, did not. This finding suggests collaboration between schools in trusts brings an additional capacity for improvement.

The improvement remains strongly associated with trust membership after adjusting for intake and testing against alternative counterfactuals: the headline causal estimate is +4.0 ranking places, rising to +6.1 when the comparison group is matched on prior trajectory and intake, and the result survives placebo testing.

The finding is not explained by the available measures of intake. It remains on progress measures that account for pupil starting points (+3.6), after adjusting for free school meal rates, and there is no evidence of differential cohort-size change.

The biggest turnarounds are schools sponsored into trusts after being judged inadequate: +15.8 ranking places against comparable schools that would later face the same judgement but had not yet been sponsored – the largest effect in the research.

What it means for children

We estimate the impact of the effect to be around 5,400 more children reach the expected standard in reading, writing, and maths each year. In the poorest fifth of communities, primary gains are three times the national figure.

At secondary, around half a million pupils sit GCSEs in trust schools annually, and the improvement is equivalent to roughly 1–2 additional GCSE grades in a single subject for each of them – or, counted at the threshold instead, around 16,000 of those same pupils reaching a grade 4 or above in English and maths (the two are the same improvement expressed two ways, not separate gains).

How it was measured

Every state school in England, 2007 to 2025. Each school was ranked against all other schools nationally, each year, on its headline published attainment results for that year. Changing tests and grading systems over the period do not distort the comparison.

Improvement is each school's change from its own pre-trust starting point. One rule of interpretation: a fall in the rankings does not necessarily mean falling results – schools can be overtaken by improving schools while their own performance holds steady. Rises and falls are movements in relative position.

The power of the group

The stronger improvement effect is associated with membership of a multi-academy trust: new governance combined with collective capacity. Standalone academies changed less than academies in groups, though it's important to note they were starting from a much higher baseline average.

School type	Change in percentile place in national rankings	
Academies in trust groups	+6.2	Rose substantially
Standalone academies	+1.2	Improved slightly from a high start
LA maintained schools	-0.3	No observable improvement

Like-for-like: schools in groups outperformed comparable schools outside any group by +1.4 ranking places (95% CI 0.9 to 1.9), rising to +2.7 when the comparison group is weighted on prior trajectory and intake – with sponsored-school recoveries excluded, so the result is not driven by recovery from a low point.

99% of sponsored academies are in groups: in practice, the school system's formal turnaround work has overwhelmingly been delivered through trusts.

Why this evidence can be trusted

- Covers every school in England with available data, over nearly two decades – not a sample.
- Findings triangulated across independent methods: before/after comparison, causal difference-in-differences under five specifications, a progress-measure test, placebo tests, a common-baseline comparison (including school-by-school matching: like-for-like advantage +1.7), and rate-of-improvement analysis. Each answers a different limitation; all point the same way.
- Independently validated by an external academic expert; full methodology published alongside.
- Limitations of the methodology are stated openly: rankings measure relative rather than absolute position, and sector averages conceal some variation.

The mechanisms inside trusts require further research.

What this means for policy

- Structure: the evidence suggests that trust membership creates greater capacity for improvement. The next phase of reform should seek to extend that capacity to schools that remain outside trusts.
- Equity: the largest gains are found among schools serving the most deprived communities – making trust growth relevant to educational equity as well as system structure.
- Form: the evidence points more strongly to trust groups than to academy status alone. The next phase should therefore focus on schools joining effective groups, while recognising that the analysis does not imply every trust improves every school.

Across two decades of national data, schools that joined trusts improved substantially more than schools that did not. Independently validated and fully replicated, the evidence materially strengthens the case for continuing the transition towards a system of school trusts.

Independent review

QUEEN'S UNIVERSITY BELFAST

SCHOOL OF SOCIAL SCIENCES, EDUCATION AND SOCIAL WORK

Independent review by Professor Daniel Muijs, Head of School, School of Social Sciences, Education and Social Work, Queen's University Belfast

The question of the impact of academisation and of joining a multi-academy trust is one that has been the subject of considerable discussion, but in many cases less actual empirical research. While existing studies tend to show positive impacts (e.g. Machin & Vernoit, 2011; Eyles & Machin, 2015; McDool, 2016), this is not universally the case with some showing no effect, (e.g. Bernadelli et al, 2018) so the evidence has to date remained mixed. In part this is due to issues of comparison, selection, and data limitations which mean most extant studies have remained somewhat limited in time and scope.

This uncertainty is not surprising, as measuring and evidencing the impact of a large-scale policy change such as academisation is complex. A real-world policy setting is not exactly a randomised control trial, which means that a variety of factors, from differential starting points to competing policies and societal developments interfere with attempts to measure impact cleanly.

This means that there isn't, and can't be, a cast-iron method to look at the impact of a major policy like academisation and the formation of multi-academy trusts. Nevertheless, it is imperative that we look at the available data to measure outcomes as accurately as possible, as otherwise we will be forever subject to discussions that rely on ideology and opinion rather than evidence, and that therefore don't take us further in developing evidence-based policies that can help us improve outcomes for children and young people.

That is why the analyses presented here are so valuable. The scale and longitudinal nature of the data analysed from 2007 to 2025 lends a robustness to the analyses that is particularly valuable, and, for example through the ranking measure, intelligent choices have been made to standardise over time as much as is possible in light of changes to attainment measures over time. Careful thought has been given to the limitations of the data, with a range of approaches used to try and address issues such as different baselines, using a range of approaches and robustness checks of the analyses.

It is this breadth of analyses that lends particular strength to the findings. While all have limitations, the overall direction of results is clear and robust.

Though the exact size of any effect is harder to determine, with different comparisons producing somewhat different findings depending on assumptions and controls, the direction of effects is entirely consistent across analyses. While the different pre-joining trajectories of maintained schools means that causality remains subject to some residual uncertainty, these analyses therefore provide substantial evidence consistent with a positive causal effect of trust entry.

This is important, as it provides one of the clearest answers we have to date regarding the effectiveness of academies and school trusts. And that answer is a positive one.

Of course, questions remain. The exact size of the positive impact is less clear, and differs across school types. We can of course also not assume a universal effect. Differences within the broad categories of academies and maintained schools are large, and further analysis is needed to shed a light on the differential impact of trusts and on the factors that lead to positive outcomes.

The relative rather than absolute nature of the ranking measure also means that it would make sense to do further work using different methodologies to try to standardise scores over times, such as using replicate the principal models using underlying outcomes expressed as year-specific z-scores, and using raw metrics within periods where the same measures operate.

As large-scale policy reforms are not randomised, there will always be some limitations in terms of evidence of impact. Nevertheless, this important study provides compelling and convincing evidence in this contested field, and suggests that our system of multi-academy trusts deserves policy support.

Professor Daniel Muijs
Head of School, School of Social Sciences, Education and Social Work
Queen's University Belfast

Queen's University Belfast is part of the Russell Group of universities, combining excellence in research and education with a student-centred ethos. 99% of its research environment is world-leading or internationally excellent (REF 2021). It is in the UK Top 5 for intellectual property research commercialisation and No 1 in the UK for knowledge transfer partnerships.

1. A half-reformed system, an unanswered question

School trusts now educate 64% of pupils in England's state-funded school system, but a substantial minority of schools remain outside trusts. The system has therefore reached a useful point at which to take stock: 54% of schools are in trusts and 46% remain linked to local authorities. This 'half-reformed' position allows us to look back over the experience of the two parts of the system and ask what the longitudinal evidence tells us about improvement. Examination and test outcomes cannot capture everything that matters in judging a school's impact, but they are an important dimension of school improvement and are subject to robust national validation. They therefore provide the principal outcome measure for this study.

Making a meaningful comparison between trusts and other school types, including local authority maintained schools, is methodologically difficult. Government policy has itself changed the composition of the two sectors over time. Most obviously, the requirement for schools judged inadequate by Ofsted to be sponsored into trusts transferred a large group of struggling schools into the academy sector. At the same time, reforms to qualifications and national performance measures have repeatedly altered the basis on which school outcomes are reported. Both features complicate any attempt to compare performance over a long period.

There is an established literature on the impact of academy trusts, but much of the underlying data is now around a decade old and some of it is confounded by various changes to tests and performance measures which make longitudinal analysis hard.[1]

Given the scale of change in the school system since those studies were undertaken, their findings cannot by themselves answer the question of trust impact in the system as it exists today. Recent debate has therefore been working with an important evidential gap. At times, the absence of up-to-date evidence about trust impact has been treated as though it demonstrates evidence of absence. It does not.

Our purpose here is to close that gap. We trace school performance across almost two decades, ending with the most recent published results in 2025, and test whether the pattern changes when alternative analytical approaches are used.

Every stage of the analysis is documented in the methodology appendix and has been independently validated by an academic expert.

1 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), pp. 268–284; Andrews, J (2018) *School performance in academy chains and local authorities*. Education Policy Institute; Hutchings, M and Francis, B (2018) *Chain effects: The impact of academy trusts on low-income pupils*. The Sutton Trust.

2. What we measured, and how

We measure outcomes primarily through a school's percentile position in the national distribution of the headline Key Stage 2 or Key Stage 4 attainment measure. This is termed the Normalised Achievement Score for Attainment (NASattain).

For primary schools, the relevant Department for Education headline measure is used in each year, such as the percentage reaching the expected standard in reading, writing and maths.

For secondary schools, the calculation uses the corresponding GCSE attainment measure, such as Attainment 8.

Because national assessment and accountability measures changed during the period, the raw measures cannot simply be compared year by year. The ranking method avoids that problem.

In each year, schools are ordered only against other schools assessed under that year's national framework. The resulting score therefore answers a consistent question across the whole period: where did this school sit relative to its national peers? Changes in that position can then be followed even when the underlying headline measure changes.

A score of 50 represents the national median in a given year; scores above 50 indicate a higher position relative to national peers. Improvement is defined as the change between a school's pre-trust baseline and its current score.

How the evidence is structured

The analysis treats attainment as the primary outcome (NASattain). Progress measures are used as a further instrument to mitigate the potential effects of differing starting points in pupil intake and to test findings (NASprog).

Additionally, a composite measure (NAS comp) is used to blend the two together as a tertiary holistic measure, noting the analytical constraints of doing so (including that doing so can conceal divergent findings).

Unless otherwise stated, the figures throughout the paper refer to NASattain.

The methodology appendix sets out all three measures in full.

The scores must be read as measures of relative position. If a school's score falls, it has moved down the national distribution; its underlying results need not have fallen. A school can, for example, maintain the same level of attainment but lose

ranking places because other schools have improved faster.[2]

Baseline

Trust and local authority baselines differ by design: trust schools are measured from the two years before joining their trust (an average window of around nine years), while LA schools – for whom there is no joining event – are measured from the mean of their two earliest available results, typically giving a window of around 17 years.

This is the most important methodological limitation of this paper, and its implications are set out in the footnote below. It is also unavoidable, and is the very reason the analysis matters: one group was subject to an intervention the other was not.

However, two analyses are conducted to mitigate the effects of this difference (the causal analysis and the common-baseline analysis). The results of these indicate there is genuine signal among the methodological noise.

2 Two issues follow from the differing baselines. First, improvement may not be linear over time: schools converting from poor performance might show rapid early gains that then plateau, and measuring from the point of conversion may capture the fastest-improving period, which may overstate the trust effect relative to a longer measurement window. Second, the two measures answer slightly different questions – trust improvement approximates the effect of an intervention event, while LA improvement approximates a long-run trend from an arbitrary starting point – and comparing them directly risks assuming LA schools would have shown the same pattern had they been measured from a more recent baseline, an assumption that cannot be tested with this data. The common-baseline analysis, which measures every school from the same starting year, addresses this from a different direction and reaches consistent findings.

3. What we found: trust schools rise in national rankings – children benefit

At sector level, the movement is clear. Schools that entered academy trusts gained 5.57 percentile points on average from their pre-trust baseline. The corresponding average change for schools remaining with local authorities was –0.27. The difference in movement between the two groups is therefore almost six places in the national distribution. Looking only at current position, the 2025 mean is 51.33 for trust schools and 49.73 for maintained schools. This tells us both that trust schools now sit slightly higher on average and, more importantly for this study, that their relative position has improved substantially over time while the maintained sector has been broadly flat to slightly negative.

- Trust schools currently sit marginally higher in national rankings than local authority schools
- Schools in school trusts have, on average, risen significantly in national rankings over time – beyond the change seen in local authority schools (which on average have declined slightly over the period).

What this means for children

The ranking changes can also be translated into more tangible pupil-level equivalents. Doing so gives an indication of the number of children represented by an upward movement of this size in the underlying primary and secondary attainment measures.

- Secondary schools in trusts rose in national rankings by the equivalent of around 1–2 GCSE grades per pupil in one of their eight subjects. Around 509,000 pupils sit GCSEs in trust schools each year, so this is an improvement reaching roughly half a million children annually.
- For primary schools in trusts, the equivalent shift represents approximately 2 additional percentage points of pupils meeting the expected standard in Reading, Writing and Maths – equivalent to around 5,400 additional children nationally reaching that standard every year.

This translation matters because the policy significance is not the movement of a statistical rank in itself. The movement represents better relative outcomes for pupils attending trust schools than those schools historically achieved before joining. At national scale, that corresponds to improved attainment for thousands of children at both primary and secondary. Academisation should no more be presented as a universal remedy than any other policy intervention, but an improvement of this magnitude across the system is significant.

Accounting for disadvantage

Raw attainment is strongly associated with the characteristics of the communities schools serve. Schools working in more disadvantaged areas tend, on average, to record lower attainment, reflecting wider socio-economic barriers as well as differences in school effectiveness. The causal analysis deals with this by comparing trust and local authority schools within the same deprivation quintile and over the same calendar years.

Trusts have made the greatest difference in the

communities that need it most. Disaggregating the causal estimate by deprivation produces a clear gradient. Within every Income Deprivation Affecting Children Index (IDACI) quintile, trust membership is associated with a positive effect when schools are compared with maintained schools in the same deprivation group over the same years. The estimate is +2.17 percentile points in the least deprived fifth and increases to +4.57 in the most deprived fifth. The latter is more than twice the former, with non-overlapping confidence intervals at the two ends of the distribution. The evidence therefore points not only to an average trust effect, but to a larger one in the communities where educational disadvantage is greatest.

Sponsored academies account for an important part of that distributional story. Nearly two-thirds of sponsor-led academies are located within the most deprived 40% of areas nationally, reflecting the concentration of schools judged inadequate in those communities. Their distinct improvement trajectory is considered in Section 4. Restricting the analysis to converter academies alone removes the gradient (+1.23 in the most deprived fifth against +1.63 in the least), so it is substantially a composition effect. This is likely to mean that the sector's largest gains have been in the most disadvantaged communities, principally through sponsored schools, rather than that deprivation amplifies the effect of trust membership for any given school.

In the most deprived quintile, the pupil-level equivalents are correspondingly larger. Q1 secondary trust schools move from approximately the 37th to the 43rd percentile for attainment, a shift equivalent to around 2 GCSE grade points per pupil in one of eight subjects. Among Q1 primary trust schools, the equivalent is about 6 additional percentage points reaching the expected standard in Reading, Writing and Maths. That is three times the equivalent for trust primary schools nationally and corresponds to around 4,200 additional pupils reaching the standard each year in Q1 trust schools alone.

Nor is there an obvious cohort-size pattern that would explain the result. There is no relative fall in cohort size of the kind we might expect if the improvement were being produced by systematically removing lower-attaining pupils from reported results. To reconcile stable cohort numbers with such an explanation would require lower attainers leaving to be replaced, at national scale, by higher-attaining pupils entering recently converted schools. The data offer no evidence for such a shift. Further work on cohort composition would be worthwhile, but it is not a plausible explanation generated by the evidence presented here.

4. Testing the finding: would improvement have happened anyway?

A stronger test is whether these schools would have improved anyway. That requires an estimate of the outcome they might have experienced without joining a trust and a comparison between that counterfactual and what actually happened. We estimate this using the difference-in-differences method developed by Callaway & Sant'Anna; the full specification is set out in the methodology.[3]

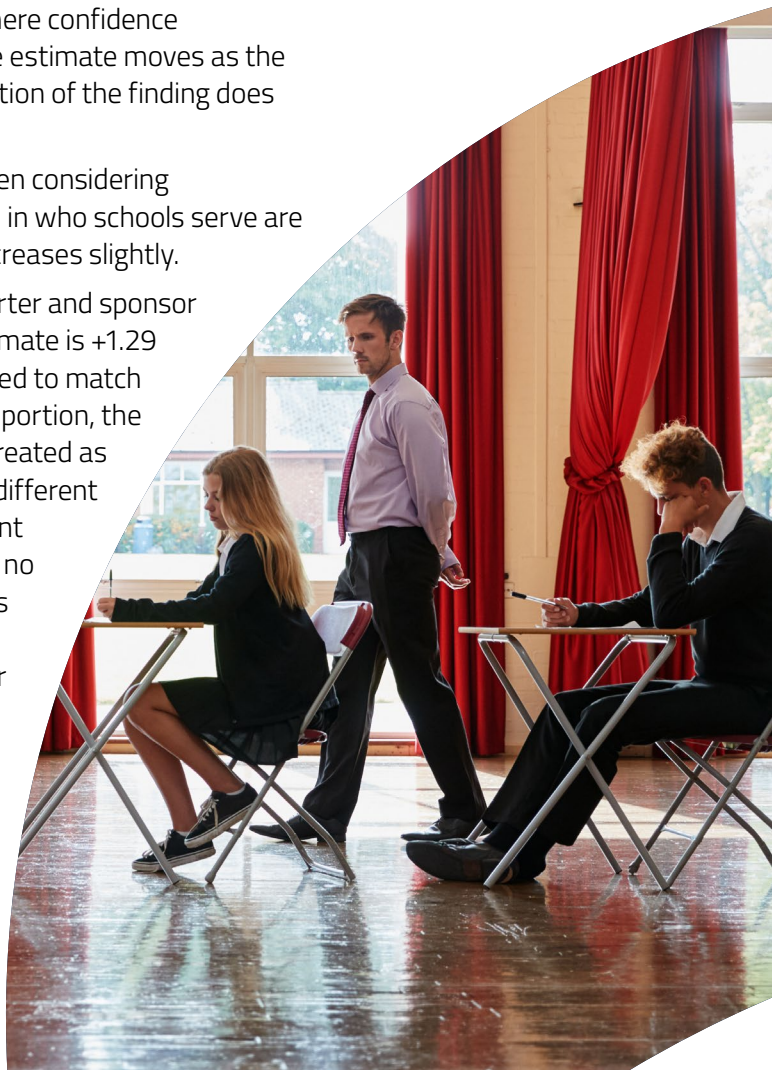
The resulting estimate is +4.00 ranking points. In other words, the model suggests that joining an academy trust raised average attainment ranking by four points relative to the outcome expected had the same schools remained with local authorities. This estimate is designed to separate the trust effect from system-wide improvement affecting all schools over the same years.

The relationship between the two headline figures is important. The descriptive analysis records a total post-baseline improvement of +5.57 points among trust schools. The causal model attributes +4.00 points of that movement to trust membership. The remainder captures improvement the model estimates could have occurred without conversion, including some recovery among schools joining after a difficult period. The estimates are not interchangeable, but they tell a consistent story.

There is little ambiguity about the statistical direction of the result. The +4.00 estimate has a 95% confidence interval of +3.48 to +4.52. We then vary the specification. Controlling for free school meals (FSM) eligibility produces +4.14 (95% CI 3.64 to 4.63); outcome regression gives +4.34 (95% CI 3.84 to 4.85); and reweighting local authority comparators so that they more closely resemble the trust cohort in both pre-joining trajectory and FSM proportion gives +6.08 (95% CI 4.38 to 7.85). Removing the transition year produces +4.42, although no confidence interval was calculated for that supplementary variant. Where confidence intervals are available, all sit clearly above zero. The precise estimate moves as the comparison is made more or less demanding, but the direction of the finding does not.

The disadvantage-adjusted models are also important when considering admissions as an alternative explanation. Once differences in who schools serve are taken into account, the estimated effect does not fall. It increases slightly.

Entry route matters. There are differences between converter and sponsor academies. Among converters, the unweighted causal estimate is +1.29 points (95% CI 0.76 to 1.80). When comparators are weighted to match the converter cohort on pre-joining trajectory and FSM proportion, the estimate is +3.59 (95% CI 1.06 to 6.05). Neither should be treated as a lower or upper bound. Sponsor-led academies present a different identification problem. Since 2016, an inadequate judgement has required a maintained school to join a trust, so there is no continuing pool of otherwise equivalent inadequate schools remaining under local authority designation. A credible comparator is instead the group of schools that would later be sponsored but had not yet reached that point. Across the 2011–2015 joining cohorts, sponsored academies improved by +15.82 points relative to this group (95% CI 14.79 to 16.94). Their own position rose by +10.90 while



3 Callaway, B & Sant'Anna, P H C (2021) 'Difference-in-Differences with multiple time periods'. *Journal of Econometrics*, 225(2), pp. 200–230.

the not-yet-sponsored comparator schools continued to move down. No other effect in this study is as large.

The hardest test: progress measures

There is, however, a tougher way to test the intake explanation. Attainment measures inevitably reflect both the work of schools and pupils' starting points. National progress measures, such as Progress 8, are constructed specifically to account for those starting points. If the trust effect were principally the product of a changing intake, we should expect it to disappear, or at least reduce sharply, when progress rather than attainment is used as the outcome.

That is not what happens. Trust schools rise by +5.04 points on the progress ranking from their pre-trust baseline; local authority maintained schools rise by +1.12. The causal estimate is +3.63 points (95% CI 2.80 to 4.37), and all five corresponding specifications are positive, ranging from +3.2 to +4.0. The progress series is shorter because published data end in 2023 for primary and 2024 for secondary schools, so fewer years and schools are available. The methodology paper reports the analysis in full.

The fake dates test

We also asked the model to find an effect where there should be none. Trust schools are assigned a fictitious joining date three years before they actually entered their trusts. During that period there was no treatment: governance had not changed. A method that simply manufactured positive effects around these schools ought therefore to return a similar result. Instead, the fake-period estimate is 0.002 times the real effect - effectively zero. The placebo test therefore behaves as it should.

Schools at a low point, and recovery that would have happened anyway

There is another reason to be cautious about apparent post-conversion improvement. Many sponsored schools entered trusts precisely because their performance had deteriorated. In the years before conversion they were often falling through the national rankings, frequently culminating in an inadequate Ofsted judgement. This creates the familiar evaluation problem described as Ashenfelter's Dip: interventions can begin when performance is unusually weak.[4] It also raises the possibility of regression to the mean, whereby an extreme low point is followed by some recovery regardless of what intervention follows.

The causal design is intended to reduce this risk. It does not infer an effect simply because a school falls before conversion and rises afterwards. It asks whether its post-conversion path differs from that of local authority schools observed over the same years. Both sets of schools therefore face the same national conditions. If ordinary recovery from a low point were sufficient to explain the result, similarly declining maintained schools should recover in much the same way. The estimated trust effect arises because they do not. The placebo result provides an additional safeguard: the same method finds virtually no effect when the joining date is moved to a period in which nothing actually changed.

Measured from the same starting gun

A different robustness check gives every school the same starting date. Taking 2016 as the baseline and measuring through to the latest result eliminates the

4 Ashenfelter, O (1978) 'Estimating the effect of training programs on earnings', *The Review of Economics and Statistics*, 60(1), pp. 47–57.

unequal observation windows used in the main descriptive analysis. Under this common baseline, relative movements are exactly zero-sum. Schools that are now in academy trusts gain +2.53 attainment-ranking points; those that remain with local authorities lose 2.70. The groups were differently composed at the outset. In 2016, future trust schools averaged 49.6 and schools that would remain maintained averaged 53.2. By 2025, their positions had crossed: 52.1 for trust schools and 50.5 for local authority schools.

We therefore repeat the common-baseline analysis using matched schools rather than the two sectors in aggregate. Each trust school is paired with maintained comparators from the same phase, the same FSM quintile and within two points of its 2016 ranking. The resulting trust advantage is +1.71 points (95% CI 0.93 to 2.47; 9,946 matched schools). It survives tighter matching and the addition of pre-2016 trajectory to the matching criteria. Matching explains much of the raw difference between the sectors, as we would expect given their different starting composition, but it does not remove the positive and statistically significant trust advantage.

There is one additional technical qualification. School type is assigned using current status, meaning that years before conversion appear in the historical series for the present-day trust group. Whether that pushes the estimate upwards or downwards depends on the trajectories those schools followed before conversion. The methodology paper considers this issue in more detail.

The rate of improvement accelerated after joining

A final check uses a different approach altogether. It looks within trust schools and asks whether conversion coincided with a change in the speed of improvement. For schools with enough data points before and after joining, the annual movement in attainment rank averaged +0.26 points before conversion. That was already faster than the long-run rate among maintained schools, which was close to zero. After conversion, the rate increased to +0.49 points a year - approximately twice as fast. A simple continuation of the pre-existing trend would not produce that acceleration. Nor is it what we would normally expect if the whole pattern were only a short-lived rebound from an unusually low starting point.

No single one of these checks is decisive on its own. The raw longitudinal comparison, the five causal specifications, progress outcomes, the placebo exercise, the common-baseline tests and the analysis of the rate of improvement each have different weaknesses and answer different challenges. Their results nevertheless converge.

5. The power of the group

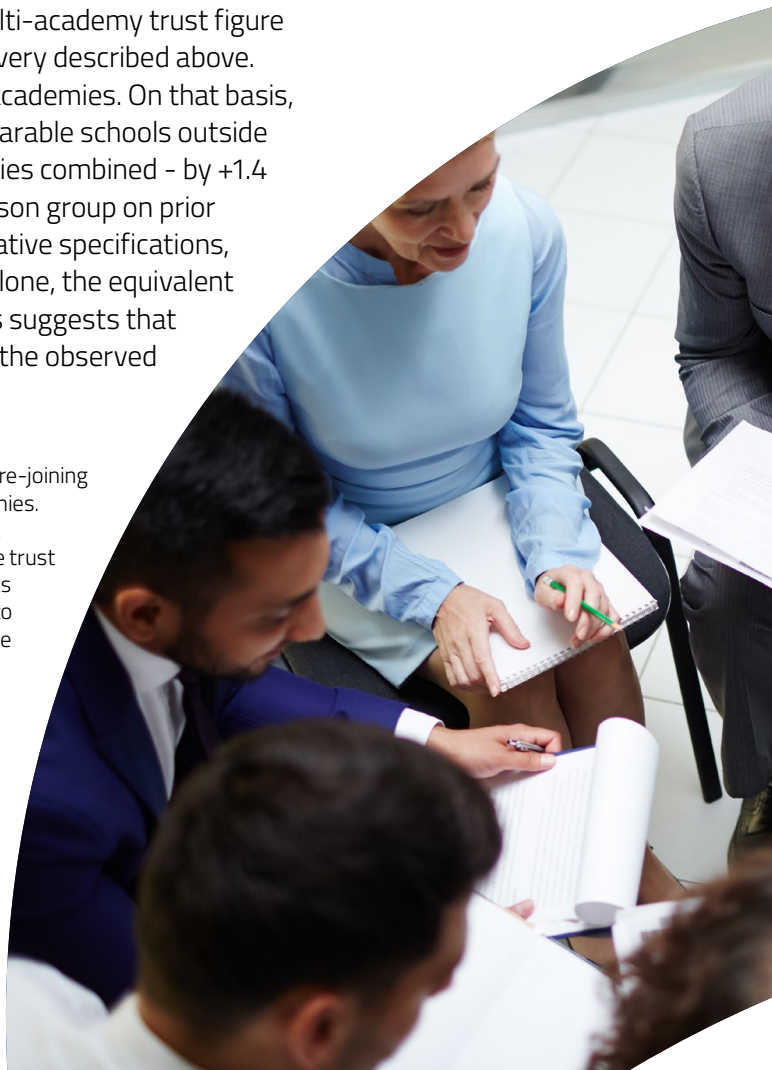
Having established the overall trust effect, the next question is organisational. Is the relevant change simply becoming an academy, or does being part of a multi-academy trust matter? The distinction is substantive. A maintained school generally continues as a single institution within local authority arrangements. Many standalone academies also converted with considerable continuity in their existing governance. Joining a multi-academy trust brings both a governance change and entry into a wider organisation with shared capacity and accountability. Comparing these routes allows us to test whether the group itself appears to matter.

The raw comparison is set out below.[5]

Group	No of schools	Raw change	Mean starting position
Academies in multi-academy trusts	8,333	+6.19	46.8
Standalone academies (single academy trusts)[6]	638	+1.17	67.9
LA maintained (benchmark)	8,260	–0.27	50.0

Two features of this table are striking. Multi-academy trust schools rise by more than six ranking places on average, while standalone academies remain broadly stable from a much higher starting point. And the system's work with schools formally identified as failing has overwhelmingly taken place in groups: 99% of sponsor-led academies are in multi-academy trusts. In practical terms, almost all of the turnaround component of the trust effect has therefore occurred within multi-school organisations.

That does not make the raw comparison causal. Standalone academies were a highly selected population, beginning at the 68th percentile on average compared with the 47th percentile for schools joining groups. The multi-academy trust figure figure also contains much of the sponsored-academy recovery described above. We therefore restrict the causal comparison to converter academies. On that basis, converters joining multi-academy trusts outperform comparable schools outside trust groups - maintained schools and standalone academies combined - by +1.4 ranking places (95% CI 0.9 to 1.9). Reweighting the comparison group on prior trajectory and FSM proportion gives +2.7. These are alternative specifications, not lower and upper bounds. Against maintained schools alone, the equivalent estimates are +1.5 and +3.1. The closeness of those results suggests that standalone academy status accounts for relatively little of the observed advantage compared with membership of a group.



5 Mean change in national attainment percentile rank from each school's pre-joining baseline (two-year window); mainstream converter and sponsor-led academies. Schools are classified by their current structure: a school that converted as a standalone academy and later joined a multi-academy trust is counted in the trust group, with its improvement measured from the two years before it joined its current trust – so improvement during its standalone period is not credited to the group. Starting positions differ substantially between the groups, and the reading rule in Section 2 applies.

6 The number of standalone academies is small because the population itself has declined and is relatively small: the large majority of academies belong to multi-academy trusts, and many schools that originally converted as standalone academies have since joined groups. The analysis includes every mainstream standalone academy with sufficient published results; the only exclusions are the standard ones applied across the study (special schools, alternative provision, 16–19 institutions and similar, whose performance measures are not comparable), not any sampling from the standalone population.

This finding needs three qualifications. We cannot make a reliable direct league table of multi-academy trusts and standalone academies: there are relatively few standalones, their selection history is unusual, and their own causal estimate is inconclusive. Nor does this analysis tell us which features of multi-academy trusts generate the effect. Finally, joining a group and changing governance happen together, so the data cannot cleanly allocate the estimate between those two changes.

Even with those limits, the balance of evidence places the group at the centre of the story. Improvement is more strongly associated with joining a multi-academy trust than with holding academy status while operating alone. Shared school improvement capacity, governance and organisational resource are plausible mechanisms, but this study does not claim to have isolated them. Standalone conversion was, in any case, concentrated among a smaller set of schools that tended already to be strong.

6. Limitations

The analysis draws on a near-complete national population of maintained and academy schools. That breadth gives the results considerable descriptive power, but it does not remove the limitations inherent in the design.

The most important are these:

- Trust and LA baselines measure different time windows. Trust schools are measured from the point of joining their trust – on average roughly nine years of data. LA schools use their earliest available result, typically from 2007, giving an average window of around 17 years. The window lengths differ systematically, and the two groups enter the analysis through different mechanisms: trust schools around a conversion event, maintained schools simply where data become available. Two analyses address this directly: the causal analysis compares each joining cohort with comparison schools over the same calendar years, and the common-baseline analysis measures every school from the same starting year – including on a matched, school-by-school basis against individually similar schools.
- Improvement is unlikely to be linear over time. Schools converting from a low base often show rapid early gains that then level off, so measuring from the point of joining captures the fastest-improving part of a school's trajectory; this may flatter the trust figure relative to a longer window. Expressing change as a per-year rate would not resolve it, since that assumes approximately linear change and leaves the difference in starting points intact.
- The ranking measure ordinalises the underlying data and can carry ceiling and floor effects. A robustness re-run using normal scores in place of percentile ranks produced materially identical standardised effects; analysis of actual scores and z-scores is identified as future work.
- Free school meal eligibility, used in the causal analysis to test whether intake explains the improvement, is a limited proxy for pupil intake and captures one dimension of disadvantage, not all of them.
- COVID-19 caused disruption to the national tests and exams for 2020 and 2021, which means there is a gap during this period. School-level secondary data returned in 2022 and primary returned in 2023. 2022 contains secondary results only, and the remaining local authority maintained secondary sector is small, so single-year comparisons for 2022 rest on a small and unrepresentative sample and should be treated with caution.
- Averages conceal variation. Not every trust school improved its ranking, and not every local authority school saw its ranking decline; the findings describe sector averages across thousands of schools, not a guarantee for any individual school or trust.

7. What this means for policy

The policy question is now different from the one faced when academies were first created. England has a mature trust sector educating more than half of the system, alongside a substantial maintained sector. We can therefore examine what happened to schools on either side of that structural divide rather than arguing only from theory. Across the analyses in this paper, schools joining trusts move further up the national attainment distribution than comparable schools that do not. The result survives causal testing, remains after accounting for disadvantage and pupil starting points, and is strongest in the communities facing the greatest disadvantage.

Three conclusions for policy follow. First, on system structure, the evidence supports continuing the transition towards school trusts. This should not be caricatured as a claim that every school could one day sit above the national average. The ranking measure makes that impossible. The point is instead historical and comparative: at the moment when the system has just tipped into a majority-trust form, we can observe how the two sectors have moved. Trust schools have improved more. That is evidence of an improvement capacity within trusts which policy should seek to extend to more schools.

Second, the distribution of the gains matters. The largest effects occur in the most deprived communities, so the question of trust growth is also a question of educational equity.

Third, the evidence has a clear boundary. It establishes an average relationship between trust membership and improved national standing; it does not imply that structure is sufficient, that every trust succeeds with every school, or that we yet understand precisely which trust practices produce the effect. That mechanism question is the next important research task.

The claim is consequently both bounded and important. Using two decades of national data, with the analysis independently reviewed, schools entering trusts improved their national ranking materially more than schools remaining outside them. That does not settle every question about trusts. It does, however, change the evidential basis on which the next stage of structural reform should be judged.

Use of AI assistance

In line with Resnik & Hosseini, we declare for transparency the use of AI in this work.[1] The analysis pipeline for this paper, including the Python scripts used to compute all figures from source data, was developed with the assistance of Claude, an AI assistant produced by Anthropic. The code used to generate the underlying school and trust level data, and the ranking calculations, were validated by an external independent expert validator specialized in AI coding. Subsequent calculations were independently verified by an academic expert using manually reviewable files (Excel) prior to publication, including the replication of calculations from scratch. Methodological decisions, interpretation of findings, and editorial judgements were made by the author. As such, while this significant body of work has been supported by AI, it has been subjected to verification and has been human led at each stage.

1 Resnik, D B, & Hosseini, M (2026) 'Disclosing artificial intelligence use in scientific research and publication'. *Accountability in Research*, 33(2).

Confederation
of School Trusts

cstuk.org.uk

© 2026 Confederation of School Trusts