Department for Education

Multi-academy trust performance measures: England, 2014 to 2015

Methodology document for SFR32/2016

July 2016

Contents

About this release

This document provides an overview of the data used in the production of the Multi-academy trust performance measures statistical working paper.

It provides information on the data sources, their coverage and quality and explains the methodology used in producing the data.

It is based on the Office for National Statistics' guidelines for measuring statistical quality.

Feedback

We are changing the way our releases look and welcome feedback on any aspect of this document at:

Academies.DATA@education.gsi.gov.uk

After 1 October 2016 this email address will be changed to:

Academies.DATA@education.gov.uk

About the output

Data Collection

The underlying data is that used for the performance tables at key stages 2 and 4 and Edubase, the department's database of school records.

Punctuality

This is a new publication this year, because this is a developing policy area.

The figures in 'multi-academy trust performance measures' are used as internal management information before the performance data has been published. The pre-announced publication date is chosen in order to allow enough time to produce and quality assure all the tables in the release, and ensure that the data used is the most up-to-date.

Changes to the tables

This is a new publication.

A measure at key stage 4 was published last year, which covered MATs with at least five schools with results at key stage 4. This year we cover MATs with at least three schools with results at key stage 4, as well as publishing a measure at key stage 2 for the first time.

We have changed the presentation of the tables to be more consistent with how data on the new progress 8 measure at key stage 4 is published on the department's performance table website.

Further changes to the accountability measures are planned in 2015/16 and these tables will reflect those changes in future years.

Accuracy and reliability

Data coverage, quality and validation

There are no planned revisions to this Statistical First Release. However, if at a later date we need to make a revision, this will comply with the departmental revisions policy.

An analysis of key stage 4 multi-academy trust performance measures was published last year as a Statistical Working Paper (SFR 09/2015). This methodology note draws on that work. For more details, please consult: Measuring the performance of schools within academy chains and local authorities

The working paper described the search for a measure to capture two things: how effectively schools within MATs are currently performing; and how that performance has changed over time.

It should be acknowledged that the overall performance of MATs has many dimensions including pupil outcomes, financial management, quality of leadership, value for money, workforce management and capacity to expand. Performance can also be impacted by a number of contextual factors including, for example, start point and pupil make up.

No single measure is ever likely to capture every element of performance or impact. This should be borne in mind when considering the outcomes reported in this paper. It is also for this reason that we are providing extensive additional contextual data alongside these outputs.

There are strong arguments for using published data from the performance tables to underpin any new measure. It means that performance at MAT level is directly linked with measures of accountability at school level. However, simple aggregations of existing measures, such as the proportion of pupils within a MAT who achieved five good GCSEs, risk giving a misleading account of the performance of the MAT as a whole. They may, for example, merely reflect that the MAT has recently taken over poorly performing schools. Such a situation could also introduce perverse incentives into the system that would drive undesirable behaviours. For example, high performing MATs might be reluctant to take on more challenging schools if they feel this will be reflected in measures as poorer performance.

The aim was to develop measures which would avoid these potential pitfalls – demonstrating both current performance and improvement over time, taking into account start point and showing relative progress.

The first measure captures the current performance within MATs by taking an average of the current value added in each relevant school. Value added is an estimate of school performance that measures the performance of pupils in comparison to pupils with similar prior attainment nationally. In calculating the overall score, a weighting is applied based on both the size of the school (i.e. pupil numbers) and on how long the school has been part of the MAT in question.

This measure can be considered to be an attempt to answer the question: How much progress are the pupils in the schools in this MAT currently making in comparison to average (based on pupils with similar prior attainment nationally)? This will recognise those MATs that have historically driven improvements in performance and that are now maintaining that higher level. It also means that the performance of new schools that do not have historic performance data is recognised.

The second measure captures the relative improvement in the performance of the schools in a MAT over time. It examines changes in value added measures across years in comparison to schools with a similar starting point. The performance of each school is compared to other schools that started, in terms of value added, at a similar level. This is then aggregated to MAT level to get a measure of the overall level of improvement of schools within the MAT. Again, in

calculating this aggregation, a weighting is applied for both school size and length of time in the MAT.

This measure can be considered to be an attempt to answer the question: How much has the performance of the schools in this MAT improved compared to schools with a similar starting point?

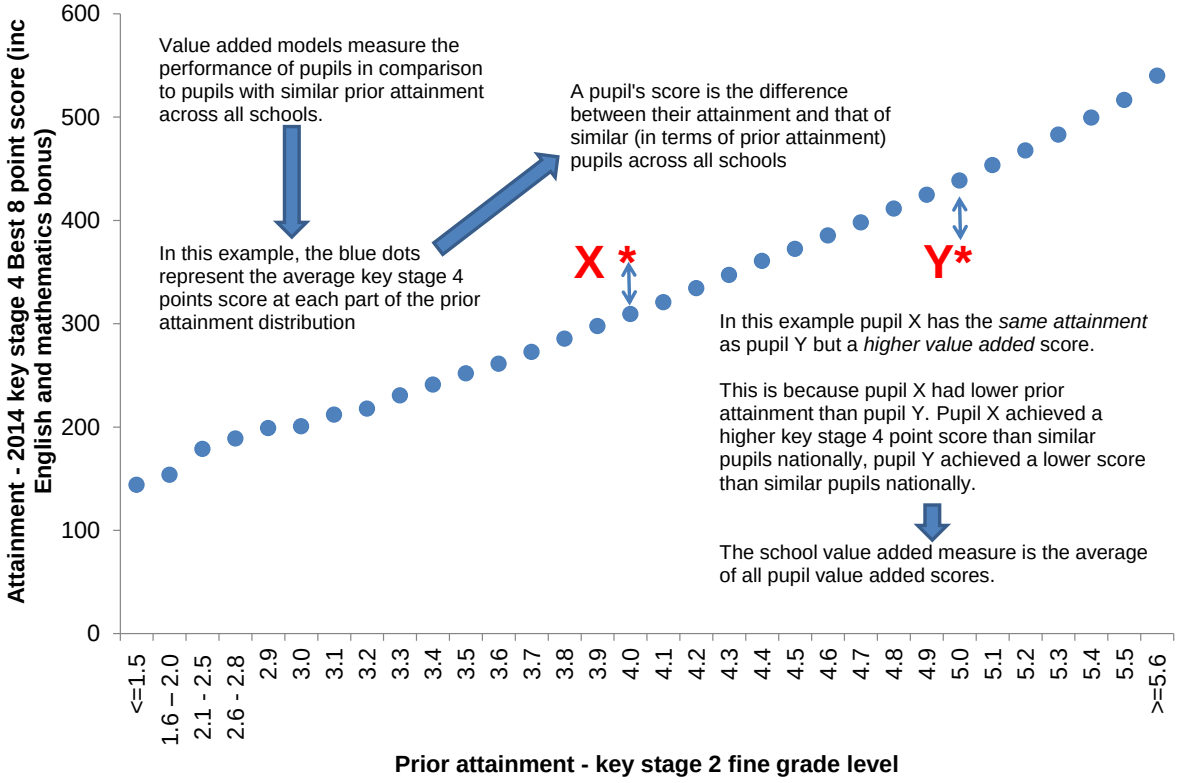
We believe value added provides a fairer comparison than simple aggregates of attainment since it controls for pupil intakes. However, measures of value added are estimates with a degree of uncertainty which should be recognised in any measures derived from them, particularly when looking at changes over time. The nature of value added means that two schools with the same score can have very different characteristics which may affect rates of improvement. To aid interpretation of the scores we are including additional contextual information, including number of schools included in the measures, the types of school and the average levels of disadvantage, special educational needs and prior attainment.

No measure can fully capture the range of individual circumstances in every school, academy MAT or the full breadth of their activity. Similarly, the measures are based on current performance data. As the performance tables evolve these measures will also evolve (for example with the introduction of progress 8 and new key stage 2 progress measures).

Results are presented for MATs with at least three schools that had results at key stage 2 or key stage 4. We have chosen to use this threshold as this focuses on larger groups who typically are more established in their roles. It also means that the results for a group are less likely to be disproportionately affected by the results of one school, which may have a very small number of pupils.

To control for prior attainment, attainment can be replaced by value added. Value added is a measure of the performance of pupils relative to those with similar prior attainment nationally. An illustration of how value added is calculated at key stage 4 is given in figure 1.

Figure 1: Assessing value added at key stage 4



Whilst there are various approaches to calculating value added, they generally share this same

principle that attainment for a pupil is compared to pupils who are 'similar'. School scores are created by taking the average of all pupil scores. The resulting scores are an estimate of the performance of a school and allow comparisons between schools with different intakes. Value added measures compare the performance of pupils to those with similar prior attainment nationally and therefore provide an estimate of the performance of a school.

In secondary schools, value added is a measure of the progress of pupils between the end of key stage 2 and the end of key stage 4 (GCSE and equivalent qualifications). The measure is presented as a score centred around 1000. In schools with scores above 1000, pupils make more progress than similar pupils nationally and conversely scores below 1000 mean that pupils make less progress than similar pupils nationally. On this scale, 6 points represents one GCSE grade in one subject. Hence a score of 1006 means that pupils achieve one grade higher in one GCSE subject than similar pupils nationally.

In primary schools, value added is a measure of the progress of pupils between the end of key stage 1 and the end of key stage 2. The measure is presented as a score centred around 100. In schools with scores above 100, pupils make more progress than similar pupils nationally and conversely scores below 100 mean that pupils make less progress than similar pupils nationally.

Such measures can be aggregated across a MAT to provide a summary of the current performance of schools, having allowed for the prior attainment of pupils attending those schools.

When assessing a school's performance, it should be noted that it is based on a given set of pupils' results. A school could have been equally effective and yet the same set of pupils might have achieved slightly different results and the school would almost certainly have shown different results with a different set of pupils.

For these reasons, the department presents value added measures with confidence intervals. Confidence intervals are provided as a proxy for a range in which users can be confident that the true value added score lies. The size of a confidence interval is determined by the number of pupils included in the value added measure and the spread of pupil scores nationally. Smaller schools have wider confidence intervals because their value added score is based on a smaller number of pupils.

When measuring the change in value added at school level, there are three components:

- the current value added of the school;
- the historic value added that improvement is measured against; and
- the average change in value added in the comparison groups.

Each of these will have a level of uncertainty and so any measure based on value added needs to take this level of uncertainty into account.

Given the relative strengths and weaknesses of measures of attainment and value added, this paper proposes two measures of the performance of schools within MATs that are based on value added.

The first measures "current performance" by taking an average of the current value added in each relevant school. This measure can be considered to be an attempt to answer the question: How much progress are the pupils in this MAT currently making in comparison to average (based on pupils with similar prior attainment nationally)?

The second measures "relative improvement" in a MAT over time. It does this by comparing changes in value added measures across years to schools with a similar starting point. This measure can be considered to be an attempt to answer the question: How much has the performance of the schools in this MAT improved compared to schools with a similar starting point?

The overall score for a MAT at each key stage is the weighted average of the individual school scores. The weight is based on:

- the number of pupils in the cohort (so that a school's contribution to the overall score is proportional to its size); and
- the length of time a school has been with a MAT (so that those that have been there the longest are given the greatest weight).

Schools that have been with the MAT for one year are given a weight of 1, those with the MAT for two years are given a weight of 2 and so on up to a maximum weight of 5.

The second measure captures the change in school level value added (VA) scores between a baseline year and the current year in comparison to schools with similar value added in the baseline year. The baseline year is taken as the last year as the predecessor school (if applicable) or five years ago whichever is more recent.

The resulting score is a 'standardised' score that is unit free. Therefore the next stage of the process translates this score back onto a common scale (i.e. GCSE and equivalent points at key stage 4).

The calculation used for a school's improvement score is:

$$= \frac{(\text{VA score in current year} - \text{VA baseline score}) - \text{average change in VA comparison group}}{\text{Standard deviation of changes in VA comparison group}}$$

The 'VA baseline score' is the average of the VA in the baseline year and the previous two years (where available.) Value added measures are subject to volatility, so the aim of taking an average over several years is to protect against a school having an atypical result in the baseline year that then informs its improvement score for a number of years.

In some cases, schools have multiple predecessor schools open in the baseline year. In these cases the results of predecessor schools are merged together.

The steps in the calculation of the measure of change in value added are:

- group schools by their value added baseline score;
- calculate the change in value added for each school as the difference between current value added and its baseline score;
- calculate the average change in value added within each group;
- calculate the school's improvement relative to the average improvement in their value added group; and
- divide the improvement score by a measure of spread of scores within the group.

Given a MAT value added measure at key stage 2, its confidence interval is given by:

$$\text{MAT current value added measure} \pm 1.96 \times \sqrt{\frac{W \times B}{nB + W}}$$

The MAT improvement in value added measure at key stage 2 has a confidence interval of:

$$\text{MAT improvement in value added measure at KS2} \pm 1.96 \times \sqrt{\frac{W_1 \times B_1}{nB_1 + W_1} + \frac{W_2 \times B_2}{nB_2 + W_2}}$$

Where W represents the total variation in value added scores at national level that is within schools (i.e. occur at the pupil level), B represents the variation in value added schools that occurs between schools, and n represents the number of pupils in the chain or local authority that are included in the value added measure. The subscript denotes the year to which the data refers (i.e. y_1 or y_2).

Given a MAT current value added measure at key stage 4, its confidence interval is given by:

$$\text{MAT current value added measure} \pm 1.96 \times \sqrt{\frac{\sigma^2}{n}}$$

The MAT improvement measure at key stage 4 has a confidence interval of:

$$\text{MAT improvement in value added measure at KS4} \pm 1.96 \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}$$

Where, σ^2 represents the variance of pupil value added scores across all pupils nationally and n represents the number of pupils in the MAT that are included in the value added measure. The subscript denotes the year to which the data refers (i.e. y_1 or y_2).

Any methodology that examines school performance over time is affected by changes to assessment and accountability. In recent years this has included changes to accountability at key stage 4 in 2014 (including the implementation of the recommendations of the Wolf review which limits the range and number of qualifications that are included and the use of first entry rather than best entry) and also changes to the assessment of English at key stage 2.

The use of value added mitigates against the impact of such changes to a large extent. This is because such measures capture performance relative to other schools (rather than to an absolute standard) and they are hence less affected by year-on-year changes than headline measures of attainment. However, use of value added does not eliminate the issue of changes to the accountability measures entirely since some schools can be disproportionately affected by reforms.

The impact of changes over time is further mitigated by ensuring that the underlying value added methodology is consistent through the period of assessment. Value added methodologies have evolved over the time period being considered. In fact, for many schools their baseline year in this methodology falls when school performance tables used contextual value added (a measure of performance that accounted not only for prior attainment but also pupil and school characteristics). Therefore, for the purposes of this measure, historic school scores have been re-calculated using the value added methodology that was used in the 2014 performance tables. This means that the underlying data may differ slightly from published measures.

In the performance tables, the school type (e.g. sponsored academy or community school) is taken at the start of September in the academic year to which the results refer. This means that results are only attributed to the school if it has been open for a full academic year. For example, schools that convert to academy status part way through the year have their results published against the predecessor school.

The new measures are consistent with this approach. In order for a school to be included within the results of a MAT, it has to have been with that MAT at the start of the academic year. In order to be included within an academy measure in 2015, a school had to join that MAT by 11 September 2014. Schools that move between MATs part way through the academic year will have their results included within their original MAT.

As set out above, measures of performance in primary schools are likely to be based on the progress of pupils between key stage 1 and key stage 2. This means that primary schools with a highest age that is less than 11, including infant schools and first schools, will be excluded from the analysis. This is the same approach as in performance tables where key stage 1 assessments are not published at school level.

Not all academies replace existing institutions. Since September 2010, about 70 new provision academies have opened. Similarly, free schools are generally new provision (unless replacing existing independent schools). Many of these schools are 'growing schools', meaning that rather than having pupils in all year groups, they fill from the lowest year group (e.g. a secondary school may have year 7 pupils in its first year, year 7 and year 8 pupils in its second year and so on).

This means that, although these schools have been open for a number of years, they do not yet have pupils at the end of a key stage and hence do not yet have published results.

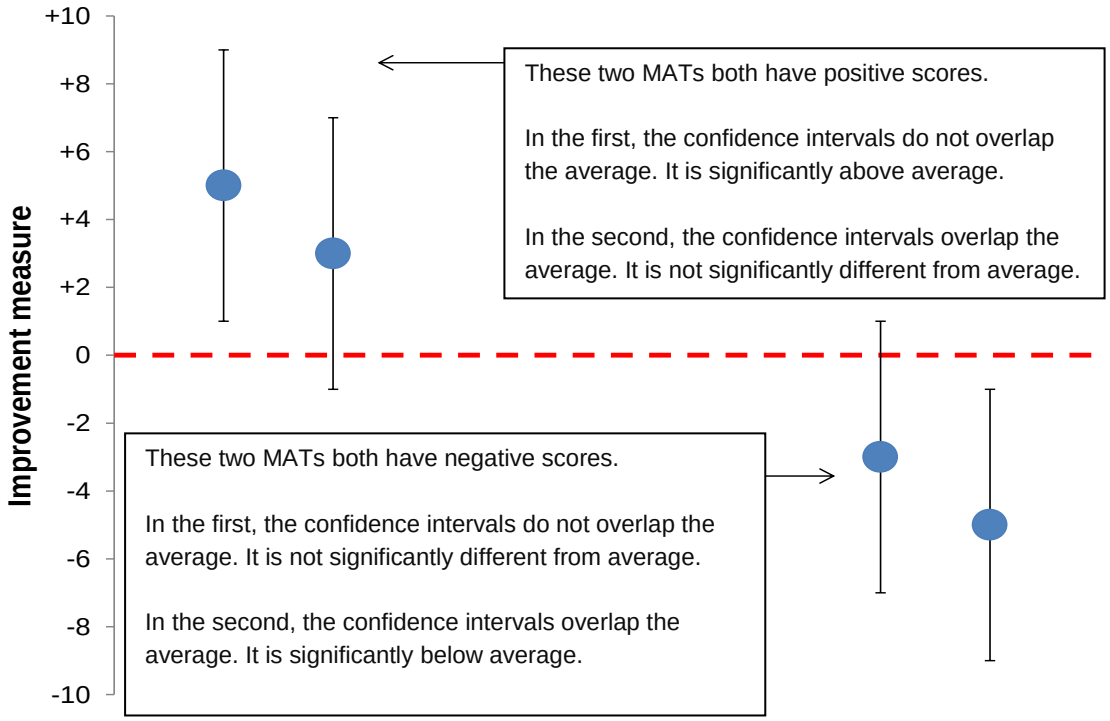
For MATs there is additional data showing:

- the total number of schools in the MAT that cover key stage 2 or key stage 4; the total number included in the performance measures; the number of schools by type of academy; and the number of academies by length of time open.
- the proportion of their key stage 2 or key stage 4 cohort that are: disadvantaged; recorded as having special educational needs at school action plus or with a statement; or have a first language other than English. In addition, the average prior attainment score across the MAT is included.

Many MATs will have scores that are not significantly different from average. As a rule of thumb:

- if the confidence intervals of one MAT do not overlap the confidence intervals of another, then they are significantly different from each other. (Note that this is not a necessary condition. Situations where there is overlap of confidence intervals but the results are significantly different from each other are possible.)
- if the confidence intervals for one MAT overlap with the score of another MAT, then they are not significantly different from each other;
- if the confidence intervals of one MAT overlap the confidence intervals of another (but does not overlap the score itself), then the two scores are unlikely to be significantly different from each other.

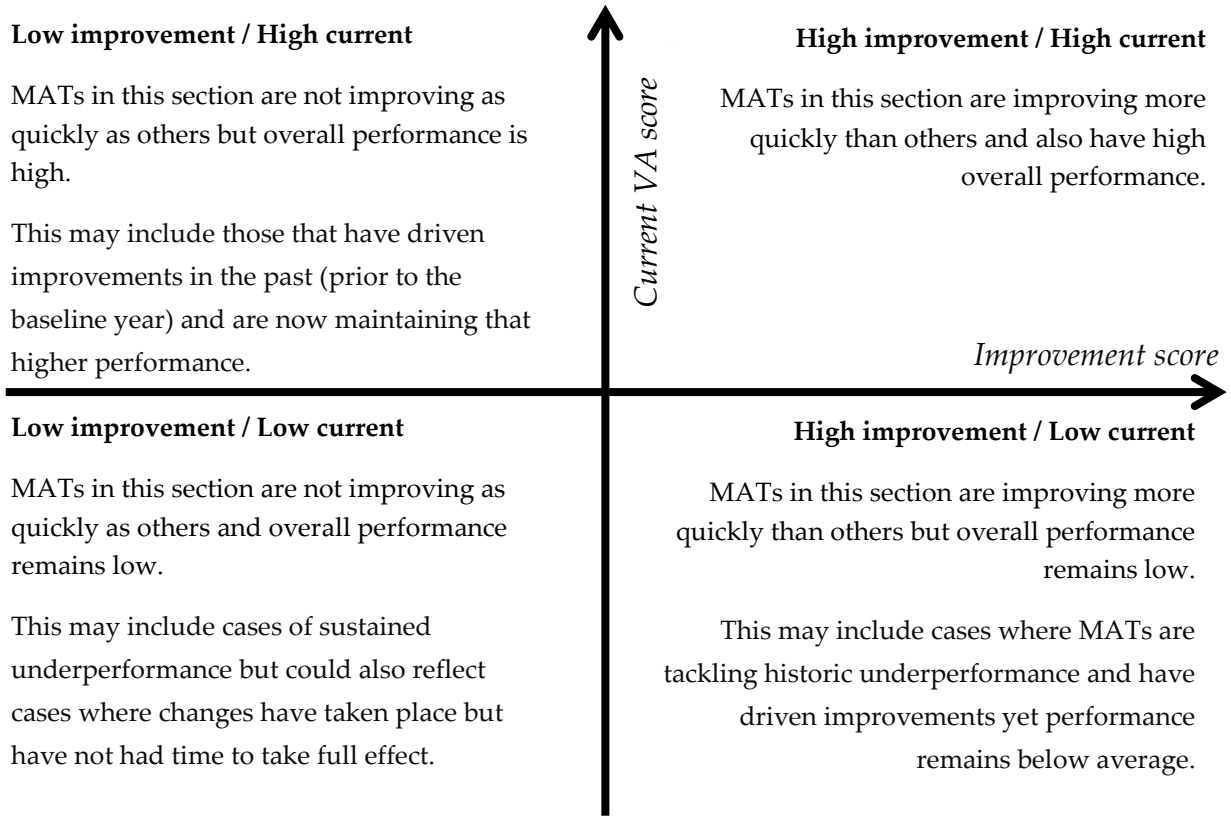
Figure 2: Assessing whether a score is significantly different from average



How should users interpret the improvement in value added measure alongside the current value added measure?

There are two aspects to measures of performance within multi-academy trusts – current value added and improvement in value added. Whilst these scores are understandably correlated, it is possible for a MAT to have a high score on one measure and a low score on the other. Figure 3 below sets out how this might be interpreted.

Figure 3: Interpreting a combination of current value added and improvement in value added measures



Disclosure Control

The Code of Practice for Official Statistics requires us to take reasonable steps to ensure that our published or disseminated statistics protect confidentiality.

The data published in this release does not reveal the identity of individuals. We have suppressed school level results for any year where the VA scores were not published for a school or its predecessor.

Accessibility

The statistical working paper text is published in pdf format so that it is accessible to all users irrespective of their choice of software. Care is also taken to ensure that the document meets accessibility guidelines. Key figures are highlighted in the text which also draws out the key messages such as changes over time. Small tables or charts illustrating key figures are also included in the text.

The statistical working paper is accompanied by formatted Excel tables with clear titles which allow users to find more detail than can be provided in the text. Any important limitations or inconsistencies in the data are mentioned in footnotes so that users don't have to refer to the text or this document.

Definitions and coverage

Definitions	The pupil characteristics data in this release are submitted to the department by schools as part of the school census collection. The definitions of all the characteristics are listed in the school census guidance.
Coverage	The coverage is all multi-academy trusts with at least three schools with eligible performance at either key stage 2 or key stage 4 in 2014/15, where the school was in the MAT at September 2014, unless the results were suppressed in line with the publication of that performance measure. For the previous year's data, the coverage is all multi-academy trusts with at least three schools with eligible performance at either key stage 2 or key stage 4 in 2013/14, where the school was in the MAT at September 2013, unless the results were suppressed in line with the publication of that performance measure.

Uses and limitations of the data in this publication

Data Item	Uses	Limitations
Value added measures at key stages 2 and 4	Comparison of performance between schools and MATs	These figures may not have been widely used historically by schools and MATs to compare performance or to set internal performance targets, compared to absolute measures of performance like percentage of 5 good GCSEs or English Baccalaureate or Progress 8, or the percentage of pupils achieving level 4 at key stage 2.

Further information is available

Previously published figures	Figures before 2014 are not available.
Previously published methodology	Figures for 2014 at key stage 4 are analysed in the Statistical Working Paper: Measuring the performance of schools within academy chains and local authorities (SFR 09/2015)
Edubase	Edubase statistics
Key stage 2	Key stage 2 performance documents
Key stage 4	Key stage 4 performance documents

Get in Touch

If from the media	Press Office News Desk, Department for Education, Sanctuary Buildings, Great Smith Street, London SW1P 3BT 020 7783 8300
If non-media	Michael Williams, Infrastructure and Funding Directorate, Department for Education, Great Smith Street, London, SW1P 3BT Email: Academies.DATA@education.gsi.gov.uk After 1 October 2016 this email address will be changed to: Academies.DATA@education.gov.uk

Department for Education

About this publication:

enquiries www.education.gov.uk/contactus

download www.gov.uk/government/publications

Reference: [SFR32/2016]

Follow us on Twitter: @educationgovuk

Like us on Facebook: facebook.com/educationgovuk