Department for Education

Multi-academy trust performance measures: England, 2015 to 2016

Methodology document for SFR02/2017

19 January 2017

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 for the academic year 2015/16.

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.gov.uk

Contents

1. About the statistics

Academies are state schools directly funded by the government. Each one is part of an academy trust. Trusts can be standalone or multi-academy trusts (MATs) - trusts that are responsible for a group of academies.

This output provides data and analysis on the performance of multi-academy trusts (MATs) in England at key stage 2 (KS2) and key stage 4 (KS4) respectively. These statistics do not include all MATs, MATs included within these statistics are:

- Those with at least three schools that had results at either KS2 or KS4 as published in the 2016 school performance tables where;
- Those schools had been with the MAT for at least three academic years. This is a different approach to previous statistical working papers on MAT performance and this is covered in more detail within the comparability over time section.

These statistics cover state-funded mainstream schools only. Special schools and pupil referral units/alternative provision academies/alternative provision free schools are not included.

In the 2016 results schools are counted under the MAT they were with as of 12 September 2015.

Where an academy sponsor oversees a number of multi-academy trusts, results are presented under the sponsor rather than the individual constituent MATs.

At KS2 there are three separate measures of MAT performance for reading, writing and maths progress respectively. These are calculated from published school level performance data for the 2015/16 academic year.

This means that for KS2 the MAT measures are based on the new primary school accountability system, implemented from 2016. Wider background and a technical guide for the new primary school accountability framework can be found here:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/563176/Primary_school_accountability_in_2016_technical_guide.pdf

At KS4 there is one measure of MAT performance based on progress 8. This means that the MAT measures are based on the new secondary school accountability system, implemented from 2016. Wider background and a technical guide for the new secondary school accountability framework can be found here:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/561003/Progress-8-school-performance-measure-18-Oct.pdf.pdf

MAT performance measures are measures which give an indication of how well MATs are currently performing. This year this output does not include a measure for MAT

performance over time. The reasons for this are explained in further detail below. We will resume the publication of measures showing MAT improvement over time in a similar fashion to our July 2016 publication when we have two years of comparable data.

It should be acknowledged that the overall performance of MATs has many dimensions including pupil outcomes, financial management, governance, value for money, workforce management and capacity to expand. Performance can also be related to a number of contextual factors including, for example pupil characteristics.

No single measure is ever likely to capture every element of performance or impact of a MAT. This should be borne in mind when considering the outcomes reported in these statistics. It is also for this reason that we are providing contextual data (including disadvantage and prior attainment) and school level underlying data for the 2015/16 academic year.

Data sources

The underlying data sources for this output are the published school level data for KS2 and KS4 respectively which can be found here:

https://www.compare-school-performance.service.gov.uk/

Edubase, the department's database of school records which can be found here:

http://www.education.gov.uk/edubase/home.xhtml

and internal DfE management information on, for example, the lead Regional Schools Commissioner (RSC) for each MAT.

How the output is created

KS2

This output contains three separate measures of MAT performance at KS2:

Reading progress, writing progress and maths progress respectively.

This measure captures the progress that pupils make in each subject from the end of key stage 1 to the end of KS2. They are a type of value added measure, which means that pupils' results are compared to those of other pupils nationally with similar prior attainment.

The respective progress score for each MAT is the mean average of its pupils' respective progress scores, weighted for length of time a school has been with the MAT and school end of key stage cohort size (typically the size of year group 6 for primary and year group 11 for secondary).

Schools are only included from their third full academic year with the MAT. Schools that have been with MAT for three years are given a weight of 3, those with the MAT for four or more years are given a weight of 4 (the usual length of KS2).

Weighting by school cohort size means that a school's contribution to the overall score is proportional to its size.

An example is shown below to illustrate the calculation of the *reading progress* measure at KS2 for a MAT:

	(i) Reading progress score	(ii) Number of pupils in end of key stage cohort	(iii) Number of years with MAT	(iv) Total weight (ii) x (iii)	(v) weighted score (i) x (iv)
Academy 1	-4.3	25	4	100	-430
Academy 2	-2.5	59	3	177	-442.5
Academy 3	3.3	50	4	200	660
Academy 4	-1.5	22	3	66	-99
Academy 5	5	90	3	270	1350
	Total	246		813	1038.5
				Overall score (sum of weighted scores / sum of weights)	1.3

To calculate the writing progress score and maths progress score for a MAT the same calculation is used but (i) reading progress score is replaced with (i) writing progress score or (i) maths progress score respectively.

We have not produced a combined measure at KS2. The production of separate reading, writing and maths progress measures for MATs reflects the approach for school performance in the annual school performance tables.

KS4

Progress 8 MAT measure

This measure captures the progress that pupils make from the end of KS2 to the end of KS4. It is a type of value added measure, which means that pupils' results are compared to the actual achievements of other pupils nationally with similar prior attainment.

The progress 8 score for each MAT is the mean average of its school progress 8 scores, weighted for length of time a school has been with the MAT and school cohort size.

Schools are only included from their third full academic year with the MAT. Schools that have been with MAT for three years are given a weight of 3, those with the MAT for four years are given a weight of 4 and those with the MAT for five or more years are given a weight of 5 (the usual duration of key stage 3 and KS4).

Weighting by school cohort size means that a school's contribution to the overall score is proportional to its size.

An example is shown below to illustrate the calculation of the progress 8 measure at KS4 for a MAT:

	(i) Progress 8 score	(ii) Number of pupils in end of key stage cohort	(iii) Number of years with MAT	(iv) Total weight (ii) x (iii)	(v) weighted score (i) x (iv)
Academy 1	+2.5	30	5	150	375
Academy 2	-2.5	59	3	177	-442.5
Academy 3	-3.3	50	4	200	660
Academy 4	-1.5	22	3	66	-99
Academy 5	-1.5	90	3	270	-405
	Total	251		863	+88.5
				Overall score (sum of weighted scores / sum of weights)	+0.1

Interpretation

For each of the above four measures:

For all mainstream pupils nationally, the average progress score is zero. The MAT level progress scores will be presented as positive and negative numbers either side of zero:

- If a MAT has a score of zero this means that, on average, pupils within the MAT do about as well as those with similar prior attainment nationally.

- A positive score means that, on average, pupils within the MAT do better than those with similar prior attainment nationally.

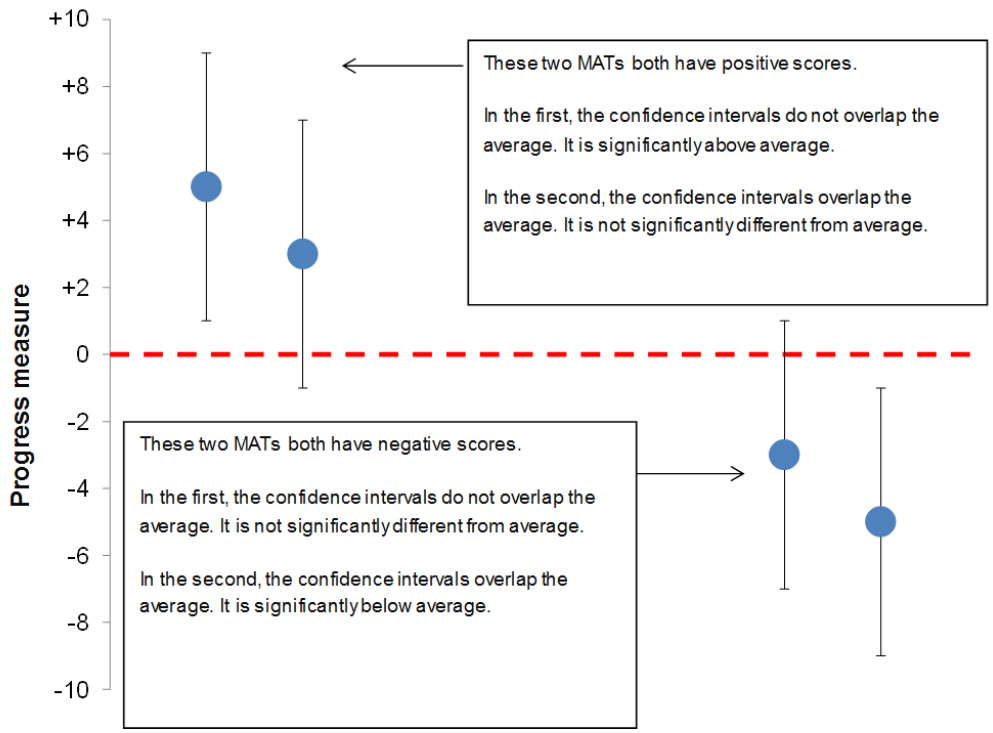
• A negative score means that, on average, pupils within the MAT do worse than those with similar prior attainment nationally. A negative score does not necessarily mean that any/all of the schools within the MAT are failing.

Confidence intervals and ranking

There is a level of uncertainty within our measures as they are based on a given set of pupils' results. MATs could have been equally effective and yet the same set of pupils might have achieved slightly different results and would almost certainly have shown different results with a different set of pupils. In recognition of this, the measures are presented with 95% confidence intervals. These provide a range in which users can be confident that the true progress score lies. Smaller groups have wider confidence intervals because their progress scores are based on smaller numbers of pupils. We can use the confidence intervals to identify MATs performing better than average or worse than average by a statistically significant amount, and close to average.

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.



Users should bear in mind that it is possible to be statistically above/below average anywhere within the distribution – not just at the extreme ends. In addition, the confidence intervals (that result from uncertainty) mean it is inappropriate to specify a precise performance-based ordering of all MATs.

Given a MAT progress measure, its confidence interval is given by:

$$\text{MAT progress score} \pm 1.96 \times \sqrt{\frac{\sigma^2}{n}}$$

Where σ^2 represents the variance of pupil progress scores across all mainstream pupils nationally and n represents the number of pupils in the MAT that are included in the progress measure. Each measure at KS2 and KS4 uses the same approach to confidence intervals.

Improvement measure

This year, due to the lack of comparable data over more than one year resulting from the new school accountability framework at KS2 and KS4, we cannot produce a new improvement measure. However, we advise users to look back at our measures published last year to get an indication of MAT improvement.

The improvement measure last year was based on the previous value added methodology. An explanation for the calculation and interpretation of this measure can be found in Annex A.

We intend to reintroduce the improvement over time measure from when we have two years of comparable data.

Relevance

This section describes the degree to which the statistics meet the current and potential needs of users.

Department for Education: it is important that we can hold MATs to account in a fair and transparent way. The data has been published as experimental statistics and is used to inform improvement.

Regional Schools Commissioners (RSCs): The data is used by Regional Schools Commissioners to support performance discussions with MATs, and to celebrate the success of MATs.

MATs: MATs use this information to benchmark their performance against others and help their efforts to improve.

Ofsted: could use the MAT measures to inform decisions about which MATs to 'batch' inspect.

Local Authorities: may be interested in performance of MATs within their area.

Schools: will be interested in seeing how their MAT is performing or to help them identify a prospective MAT to join.

Teachers and head teachers will be interested in seeing how their MAT is performing or to help them identify a prospective MAT to join.

Parents: could be used to compare the performance of a MAT that their own children are within. Could be used to hold their children's MAT to account and identify areas where they feel the MAT could improve. Could be used to inform choice of school for their pupil.

Others: could be used by researchers from this country and abroad.

The data generally meets users' needs well, but more can be done. The measures have been developed over a number of years and have been tested with MATs and others, and published as a statistical working paper (using the old value added methodology) in March 2015 and July 2016.

The output this year is in an adaptation phase where the measures are being aligned to the new school performance measures introduced from 2016. In future the Department will continue to engage with MATs to further improve and integrate these measures into performance tables, and explore alternative or additional contextual measures (including, for example, how to reflect progress in social mobility).

Timeliness

Timeliness refers to the lapse of time between the period to which the data refer and the publication of our measures

Our measures were published alongside the secondary school performance tables on 19 January 2017. Primary school performance data was published on 15 December 2016.

Schools are assigned to the MAT they were with as of 12 September 2015, as listed on Edubase.

Punctuality

Punctuality refers to the time lag between the actual and planned dates of publication.

The proposed month of publication is announced in advance on gov.uk and precise dates are announced in the same place at least four weeks prior to publication. In the event of a change to the pre-announced release schedule, the change and reasons for it would be announced.

2. Accuracy and reliability

Accuracy describes the closeness between an estimated result and the (unknown) true value.

Measurement error

Measurement error is the difference between the actual value of a quantity and the value obtained by a measurement. Repeating the measurement will reduce the random error caused by accuracy of the measuring instrument but not any systematic error caused by incorrect calibration of the measuring instrument.

For the steps taken to minimise measurement error in the school performance data please refer to the further information and guidance on the performance tables website.

The production team minimise measurement error by independently dual or triple running each output.

Validation and quality assurance of the data

The production team for the minimise measurement error and perform validation and quality assurance by independently dual or triple running each output. Any discrepancies in the data produced are discussed and more experienced staff involved as required. Additional checks are also carried out on the data produced.

Examples of additional checks include

Comparisons with previous figures

Check totals are consistent across tables

Check patterns in the data are as expected

Check figures against those produced for the performance tables

Data processing

Data is processed by individuals independently using Microsoft SQL and Microsoft Excel.

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 where the progress scores were not published for a school.

3. Accessibility and clarity

Accessibility is the ease with which users are able to access the data. It also relates to format(s) in which data are available and the availability of supporting information.

Clarity is the extent to which easily comprehensible metadata are available, where these metadata are necessary to give a full understanding of the statistical data.

The 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 text 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.

4. Comparability

Over time

The MAT measures produced for this output are not comparable with those from previous years. New performance measures were introduced for the 2015/2016 school performance tables. We have adapted our existing methodology accordingly; for KS2 this this year we have used the underlying school level reading progress, writing progress and maths progress headline measures respectively. At KS4 we have used the underlying school level progress 8 measure. In our publication of July 2016 and March 2015 we used the previous value added based performance measures, which are not directly comparable.

We plan to reintroduce improvement measures alongside future school performance tables publications. Users should refer back to our previous publication of July 2016 in order to gain an indication of MAT improvement over time.

We have made a change to the coverage of data in this publication. We now only include data for schools from their third academic year under a particular MAT. Previously we included data for schools with one academic year of results under a particular MAT. This brings these measures into line with inspection policy for new and rebrokered schools, recognising the amount of time needed for a MAT to have full impact on a school's results.

Our measures cover MATs with at least three schools in the relevant phase (this means that a MAT that has 3 schools with it for three years, but with two primary schools and one secondary school will not be included either the key stage 2 or key stage 4 MAT performance tables) with results in the 2016 school performance tables. This threshold is the same as in July 2016. However, in the output of March 2015 the threshold was at least five schools and only covered KS4.

These tables will continue to reflect any future changes in the school's accountability system.

Differences between school, Local Authority and National figures

Our MAT measures use the same school level data as published within the school performance tables on 15 December 2016 (primary) and 19 January 2017 (secondary).

We have not included measures for Local Authorities and have not produced a national figure.

Across different types of schools

We have included state funded mainstream academies within our MAT measures in other words; sponsored academies, converter academies, free schools, studio schools and University Technology Colleges.

We have not included special schools in our analysis. Even when comparing to other pupils with similar prior attainment, pupils in special schools generally make slower progress, and therefore types of value added measures can be a poor assessment of effectiveness.

Users should bear in mind that each MAT is different and they each operate under a variety of challenging circumstances. However, the measure does not fully account for the historic performance of schools, including the poor prior performance of schools that became sponsored academies

With other parts of the UK and internationally

Currently multi-academy trusts operate solely in England.

5. Got a query? Like to give feedback?

Media enquiries

Press Office News Desk, Department for Education, Sanctuary Buildings, Great Smith Street, London SW1P 3BT.

Tel: 020 7783 8300

Other enquiries/feedback

Adam Hatton, IFD Analysis, Department for Education, Sanctuary Buildings, Great Smith Street, London SW1P 3BT.

Tel: 020 7340 8364 Email: Academies.DATA@education.gov.uk

Annex A: Improvement measures (based on old 2015 value added)

For a full description of the improvement measure please refer to our **statistical working paper** of March 2015, in summary;

2015 Improvement measure

This 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.[1]

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

School level scores

The proposed approach for a school's improvement score is

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

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 in the baseline year 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[2];
- calculate the average change in value added within each group;

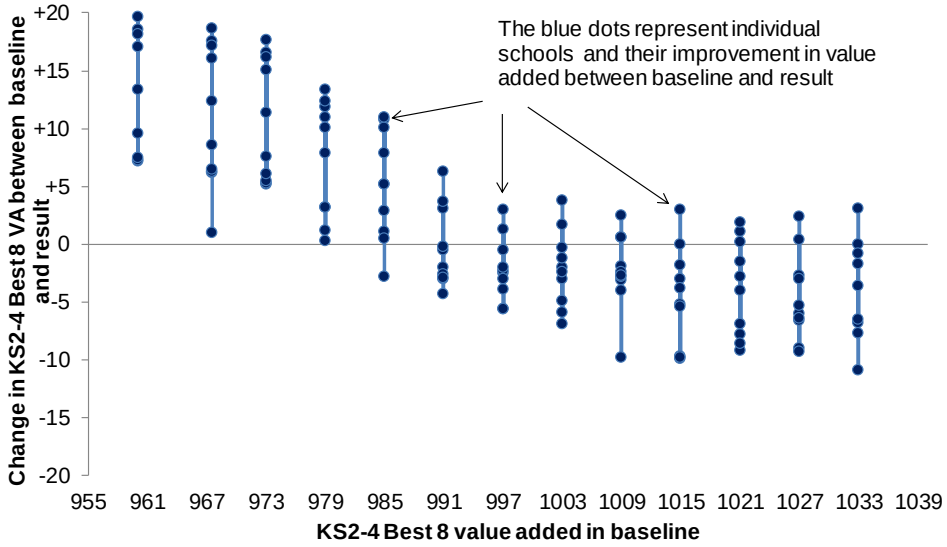
[1] This represents the usual duration of key stage 3 and key stage 4.

[2] Note that the current value added score is the value added in the current year and, unlike the baseline score, is not an average over several years. This is discussed further in section 5.

- 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.

The approach to this calculation is illustrated through a school that achieved a value added score of 995.0 at KS4 in the current year, having achieved a baseline score of 987.2. In 2015 the principles of this approach were also applied to KS2.

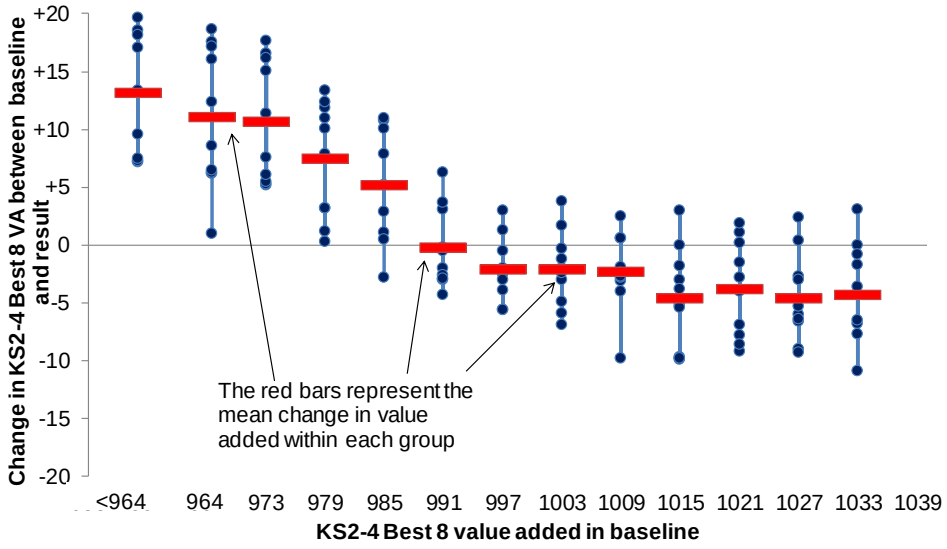
Step 1: Group schools by their value added baseline score. All schools with similar value added in the baseline year are grouped together[3]. The change in value added for each school between the baseline year and the result year is calculated.



Note: All data is illustrative

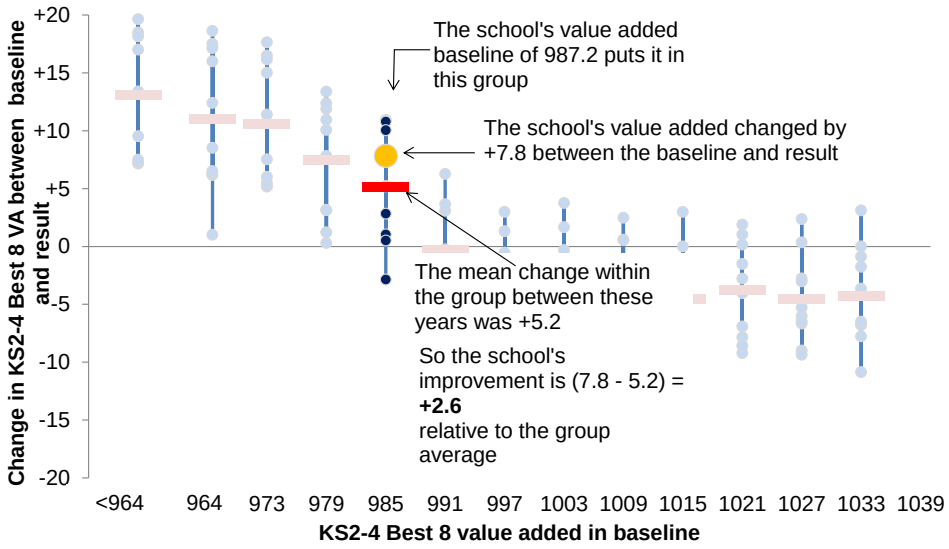
[3] Note that while each school has one "baseline year" it will appear in all comparator groups where it has the relevant data. For example, a school may have itself a baseline year of 2009 but its 2010 result will appear as a comparator school for schools with 2010 as a baseline and its 2011 results will appear as a comparator school for schools with 2011 as a baseline and so on.

Step 2: Calculate the average change in value added within each group. This is the arithmetic mean of all changes in school level value added scores within that group.



Note: All data is illustrative

Step 3: Calculate the school's improvement relative to the average improvement in their value added group. i.e. take the school improvement and subtract the group average. This is to control for the fact that those with the lowest value added tend to see the largest improvements and vice versa

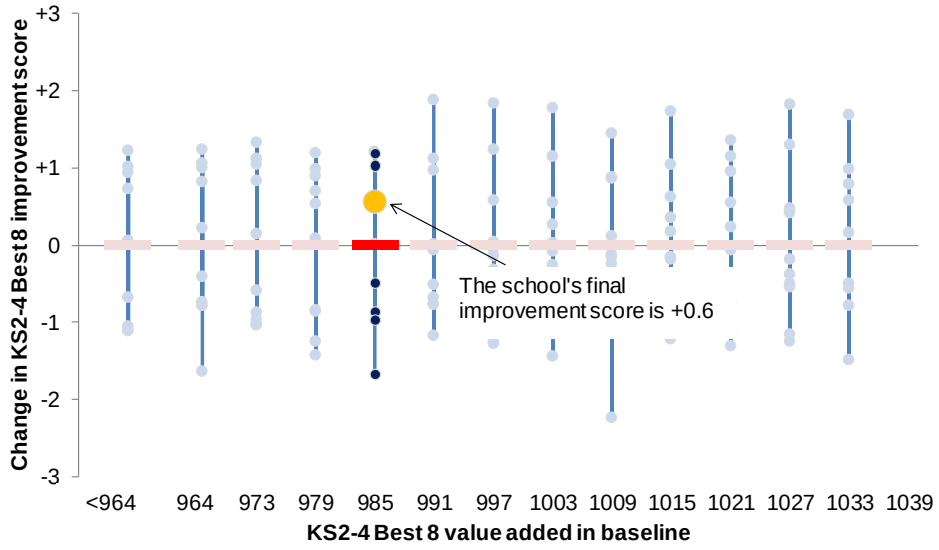


Note: All data is illustrative

Step 4: Divide the improvement score by a measure of spread of scores within the group.[4] This reduces the risk of bias caused by some groups having a wider spread of results than others (either comparing between start points, or by different lengths of time

[4]Spread is measured as the standard deviation of scores within the group.

open.) If this was the case, then extreme values from highly spread groups would carry more weight than those from other groups.



Note: All data is illustrative

Group averages and standard deviations for 2014 are provided in section 9.

MAT level scores

The overall score for a MAT 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.

At KS4 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 same is true for KS2 but up to a maximum weight of 4.

Figure 4.2 below provides an example MAT with five academies with an improvement score at KS4.

Figure 4.2: Example calculation of a MAT improvement score based on the individual school scores

	(i) Improvement score	(ii) Number of pupils in end of key stage cohort	(iii) Number of years with MAT	(iv) Total weight (ii) x (iii)	(v) weighted score (i) x (iv)
Academy 1	0.63	160	5	800	504.0
Academy 2	-0.15	150	2	300	-45.0
Academy 3	0.81	190	3	570	461.7
Academy 4	-0.32	210	4	840	-268.8
Academy 5	0.95	155	1	155	147.3
	Total	865		2665	799.2
				Overall score (sum of weighted scores / sum of weights)	+0.30

Note that this approach to weighting by length of time open gives a relative weighting *within* each MAT. If all of the schools joined the MAT at the same time, or had been with the local authority for the same length of time, the weighting would have no effect.

The result is a point estimate for the relative change in school effectiveness for each school. However, there is a degree of uncertainty around value added scores in both the input and output measures and so this score represents the central point of a range of values in which the true value is likely to fall. This is addressed in the confidence intervals section.

Translation to GCSE outcomes

Value added scores are centred round 1000 at KS4, and differences from this are measured in GCSE (and equivalent) points. For example, a score of 1006 at KS4 means that, on average, pupils in the school achieved one grade higher in one GCSE than similar pupils nationally. Similarly, a simple year-on-year change is also measured in GCSE points.

The value added improvement score (as derived from the improvement methodology) is known as a ‘standardised score’; this enables comparison between different groups and different time periods on a consistent basis. However, unlike value added scores, it does not have associated units of measurement and hence cannot be interpreted directly in terms of pupil attainment.

It is, however, important that the measure can be interpreted in such a way so as to have a ‘real world’ meaning. Without this, there is an increased risk of the score being seen as a statistical ‘black-box’ which will reduce its credibility and make interpretation difficult.

The proposed approach is to remove the standardisation by treating the MAT as if it were a school that started as broadly average on value added (i.e. it had a value added of around 1000 at KS4.) In effect, the proposed approach answers the question

“If this MAT was a school that had started with an average value added, what relative improvement in one year would have given it this score on the new measure?”

Standardisation was carried out by dividing the relative improvement by the spread (standard deviation) of results in each comparison group. Therefore, to remove the standardisation, we multiply the overall improvement score by the standard deviation of improvements over one year of schools that had previously been broadly average (i.e. which were in the middle grouping of value added).

As the spread of scores varies within each value added group, applying one value to all schools may result in a score which, on the face of it, is different from what the MAT achieved. This would be particularly true if the majority of a MAT’s schools were in one of the lower value added groups where the spread of scores tends to be wider.

However, the resulting comparison with other MATs is fairer than the simple change in value added and this issue would in fact have very little impact on the majority of scores as they arise from schools that are not in the extreme groups.

This leads to a measure that is in GCSE and equivalent points and is phrased in the style:

“The performance in this MAT means that its schools have on average improved progress by three points, or half a GCSE grade, in a year compared to schools with a similar starting point”

The MAT improvement in value added measure at KS2 has a confidence interval of:

$$\text{MAT improvement in value added measure at KS2} \pm 1.96 \times \sqrt{\frac{\mathrm{W}_1 \times \mathrm{B}_1}{\mathrm{nB}_1 + \mathrm{W}_1} + \frac{\mathrm{W}_2 \times \mathrm{B}_2}{\mathrm{nB}_2 + \mathrm{W}_2}}$$

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).

The MAT improvement measure at KS4 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).

Department for Education

About this publication:

enquiries www.education.gov.uk/contactus

download www.gov.uk/government/publications

Reference: [SFR02/2017]

Follow us on Twitter: @educationgovuk

Like us on Facebook: facebook.com/educationgovuk