GUIDANCE PAPER

ASCL Association of School and College Leaders

Understanding and Interpreting the 2016 Key Stage 2 Results

Guidance at a glance

On 5 July 2016 the Department of Education (DfE) published a **statistical first release**[1] with provisional headline results of this year's Key Stage 2 assessments. This year's assessments (commonly known as SATs) have changed significantly from previous years.

This paper is intended to help school leaders understand these results, the context in which they exist and what they might mean for both primary and secondary schools.

What is tested, and in what way?

Reading, maths and grammar, punctuation and spelling (GPS) are all assessed through externally set and marked tests. In 2016, there was one reading paper, two GPS papers (one on spelling and one on punctuation and grammar), and three maths papers (one on arithmetic and two on reasoning). This year's tests are now all available **online**[2].

There are no longer separate tests for higher attainers; all children sat the same papers which were designed to assess the full content range from the equivalent of the old Level 3 to the old Level 6.

Writing is entirely teacher-assessed, using a **framework**[3] produced by the Standards and Testing Agency.

This guidance looks in more detail at the following:

1 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/534573/SFR30_2016_text.pdf

2 https://www.gov.uk/government/collections/key-stage-2-tests-past-papers

3 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/473675/Interim_teacher_assessment_frameworks_at_the_end_of_key_stage_2_PDFA_V3.pdf

1 How have the assessments changed?

Firstly, these are the first set of assessments designed to assess how well children have done against the new primary curriculum, introduced in September 2014. The new curriculum is significantly harder than the old, with some content taught one, two or even three years earlier than previously. The new assessments were therefore deliberately designed to be more challenging than their predecessors. The new curriculum also prioritises different knowledge and skills (such as a greater focus on grammar).

Secondly, the scoring system used for the tests has changed. Rather than being given a level, children were awarded a **raw score** (the number of marks achieved), which was then converted into a **scaled score** (designed to ensure parity from one year to another). The expected standard on each test was set at a scaled score of 100.

Thirdly, the approach to teacher assessment (on which the writing results were entirely based) was also different. Levels have gone, and replaced with a set of statements against which teachers had to assess children as working 'towards', 'at' or 'at greater depth within' the expected standard. In addition, the conceptual approach behind teacher assessment moved from a 'best fit' model (in which teachers could award a particular level if children had achieved most of the requirements) to a 'secure fit' model (in which children could only be assessed as having met a particular standard if they could demonstrate consistent achievement of all the requirements within the standard).

What information was released?

The information released on 5 July included:

- The **range**[4] of scaled scores children could achieve (80 to 120 on each test), and confirmation that this will stay the same in future years.
- A set of **conversion tables**[5] showing how raw scores can be translated into scaled scores. These showed the 'pass mark' for each subject (21 out of 50 for reading, 60 out of 110 for maths and 43 out of 70 for GPS). They also made clear the extent to which marks 'bunched' at the top and bottom of the scales (a raw score of 3, 4, 5, 6, or 7 in maths would all result in the lowest possible scaled score of 80, for example, while a raw score of 44, 45, 46, 47, 48, 49 or 50 in reading would all get you the highest possible scaled score of 120).
- The **statistical first release**[6], containing provisional national headline figures.
- An **information leaflet**[7] to help parents understand their child's results.
- Every primary school could also access its individual pupils' results via the NCA tools **website**[8]. Secondary school can use the same site to find out the results of their incoming Year 7s.

4 https://www.gov.uk/guidance/scaled-scores-at-key-stage-2#range-of-scaled-scores

5 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/534106/2016_KS2_scaled_scores.pdf

6 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/534573/SFR30_2016_text.pdf

7 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/534658/Information_for_parents_-_2016_NCT_results_at_the_end_of_key_stage_2.pdf

8 https://ncatools.education.gov.uk/Home.aspx

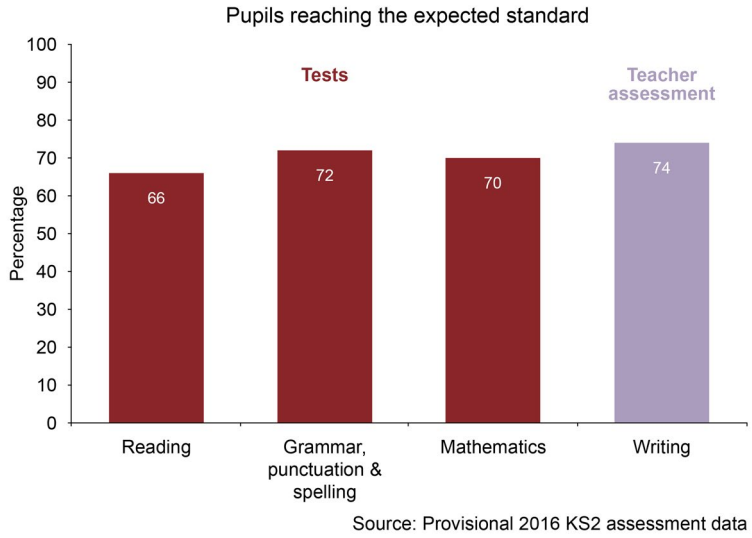
What information wasn't released?

This set of documents does not include:

- any indication of the distribution of results across the scale
- any information on how these results will be used to inform the new primary progress measure

More information on these is expected at the beginning of September 2016.

2 Headline results 2016



Source: Provisional 2016 KS2 assessment data

The average scaled score was 103 in reading and maths, and 104 in GPS.

And, the headline 'headline' result, 53% of pupils reached the expected standard in reading, writing **and** maths.

How do these results compare with last year?

The DfE has made it clear that, for all the reasons outlined above, this year's results should not be compared with previous years. However, they have also indicated that the new expected standard would be pegged at broadly equivalent to an old Level 4b (higher then last year's expected standard, which was Level 4).

Comparing this year's results with the number of pupils who achieved a Level 4 or 4b last year, looks like this:

Subject	% achieving Level 4 in 2015	% achieving Level 4b in 2015	% achieving expected standard in 2016
Reading	89%	80%	66%
GPS	80%	73%	72%
Maths	87%	77%	70%
Writing	87%	Not broken down into sub-levels	74%
Reading, writing and maths	80%	69% (based on Level 4 in writing)	53%

Clearly the percentage achieving the expected standard in reading, writing and maths can never be higher than the lowest percentage in any individual subject (and it can be much lower if there is not a good correlation between marks in different tests), so it is particularly hit by the low percentage of children meeting the expected standard in reading.

What can we learn from these results?

In terms of how this year's cohort compares to last year's, very little. While the new expected standard may have been nominally aligned with an old Level 4b, the difference in what was taught and assessed is sufficiently great to render this comparison almost meaningless. It seems highly unlikely, for example, that standards in reading between the class of 2015 and the class of 2016 have really plummeted in the way the chart above suggests.

The process used by DfE for setting the **expected standard**[9] is, in fact, deliberately designed to break the link between this year and last. It involved groups of teachers looking at the test questions, comparing them against the performance descriptor for the expected standard, and deciding whether, in their professional opinion, two-thirds of pupils working at just the expected standard would get each question right.

We should also be very cautious about comparing this year's results with subsequent years, at least for the next two or three years. As a general rule[10], when new forms of assessment are introduced, results dip as teachers and pupils get used to the change. As the assessments become embedded, results naturally rise without necessarily indicating a change in teaching quality or learner understanding. This common pattern, known as the 'sawtooth effect', is likely to be exacerbated in this case by the fact that this year's cohort of children have only been following the new national curriculum for two years (rather than the full four years of Key Stage 2, or even the full six years of formal primary education). The likely 'improvement' in results over the next few years should therefore be interpreted in this context.

In terms of how schools compare with each other, the results may serve slightly more purpose. However, caution should still be exercised. The small size of many Year 6 cohorts means that basing any judgement on the quality of a school on a single year's attainment is extremely dangerous. In addition, the writing results this year are likely to be particularly unreliable. There is substantial anecdotal evidence that both teachers and moderators interpreted the new teacher assessment framework in different ways. This rather worrying **survey**[11], for example, reveals stark differences of opinion among teachers about the extent to which the guidelines permitted them to support and scaffold children's writing.

What does this mean for primary schools?

There are **new primary floor standards**[12] for 2016, which these results will feed into. A school will be above the floor if at least 65% of pupils meet the expected standard in reading, writing and maths or the school achieves sufficient progress scores in all of these subjects.

We don't yet know how many schools will be below the attainment part of the floor standard; based on these headline figures, it is likely to be a substantial number.

9 https://www.gov.uk/guidance/key-stage-2-tests-standard-setting

10 See Linn, Graue and Sanders, 1990; and Linn, 1998

11 https://michaelt1979.wordpress.com/2016/06/15/consistency-in-teacher-assessment/

12 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/496158/Primary_school_accountability_in_2016.pdf year-2016-to-2017

Nor do we know how many schools will be below the progress part of the floor standard, as the DfE has not yet confirmed where this bar will be set. We do know, however, because the Secretary of State has made this **commitment in writing**[13], that no more than one percentage point more schools than last year (ie 6%) will be classed as below the floor overall. The progress bar, therefore, will have to be set at a level that enables this commitment to be met.

That commitment will be scant comfort, of course, to those schools that find themselves in the bottom 6%, and that information won't be available until September 2016 at the earliest.

What does this mean for secondary schools?

These results will impact secondary schools in a number of ways. To clarify straightaway one effect they won't have: the proposed Year 7 resit for children who haven't met the expected standard in the Year 6 tests is not happening in the 2016-17 academic year. We don't yet know whether this proposal will make it into legislation, and if so what the details will be, but it definitely won't apply to children starting Year 7 in 2016.

The changes to both the curriculum and the assessment regime in Key Stage 2 will, of course, affect the knowledge and skills with which children are leaving primary school. Most secondary schools are starting to consider the impact this might have on their Year 7 curriculum, but few have really got to grips with how different the curriculum really is. We would strongly encourage secondary school leaders (and heads of Year 7) who haven't yet looked properly at the **primary curriculum**[14] and this year's **tests**[15] to do so now, and to talk to colleagues in their partner primary schools about how they can build on the different knowledge and skills this year's children will have.

The Key Stage 2 results also form the input measure for Progress 8, and contribute to the way in which most secondary schools set targets for their students. Please see the appendix in section 4 of this paper for a technical analysis of this aspect of the Key Stage 2 figures.

It is, of course, important to remember that even with the Key Stage 2 starting point it is impossible to predict a pupil's Progress 8 score or set targets in the same way as was possible with expected progress measures. Future Attainment 8 estimates will change as a result of differences in national subject entry patterns and performance. Current Attainment 8 estimates do not, therefore, provide an accurate indication of future Attainment 8 estimates. Trying to predict Progress 8 scores based on Key Stage 2 results and previous years' Attainment 8 estimates is likely to be time-consuming and inaccurate. As now, schools should work to ensure that all pupils achieve the best possible grades based on the school's knowledge of their starting point (of which Key Stage 2 results are just one piece of data) and individual needs and interests.

Finally, these results have implications for the Year 7 literacy and numeracy catch-up premium funding given to secondary schools to support pupils who did not achieve the expected standard in reading or maths at the end of Key Stage 2. Unfortunately for schools hoping for a bonanza as a result of these figures, the government has **confirmed**[16] that the 2016-17 catch-up funding will be pegged to the 2015-16 amount, with schools receiving the same amount as they received last year, adjusted to reflect the percentage change in the size of their overall Year 7 cohort.

13 http://www.ascl.org.uk/resources/nicky-morgan-letter-to-all-primary-schools-reg-schools-below-the-floor-standard-6jun-16.html

14 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/425601/PRIMARY_national_curriculum.pdf

15 https://www.gov.uk/government/collections/key-stage-2-tests-past-papers

16 https://www.gov.uk/guidance/year-7-literacy-and-numeracy-catch-up-premium-guide-for-schools

3 What should school leaders do now?

These results, and the fact that they seem so different from last year, are bound to cause worry and uncertainty. The fact that we only have part of the picture so far, with just the headline figures released, will compound this.

There are, however, a number of actions both primary and secondary school leaders can take now to help minimise concerns and smooth the passage of the children affected by these changes as they embark on the next stage of their education.

Primary school leaders:

1 **Don't panic**: as explained above, while a large number of schools will undoubtedly find themselves below the attainment floor standard, most will be 'saved' by the new progress measure. Ofsted and the Regional Schools Commissioners should consider attainment data in the context of 2016 national averages and will not be making judgements on overall school performance until progress data is available.

2 **Reassure children and parents** that these results don't mean standards at the school have dropped, or that the children have not done well. It will help to set the results both of individual children and of the school as a whole in the context of the national figures, and to discourage direct comparisons with last year.

3 **Understand as well as you possibly can** why your school has achieved the results it has. If your results in particular subjects differ markedly from the national average, or from those of similar schools, can you explain why? Be prepared to resist attempts from any source to mis- or over-interpret these results.

Secondary leaders:

1 Continue to see the SATs results as just one piece of data, and to augment them with other assessments.

2 Look at the new primary curriculum and this year's tests, and consider what changes you might need to make to your Year 7 curriculum to ensure it builds effectively on pupils' prior learning.

3 Look at the question level **analysis**[17] for the 2016 tests in RAISEonline (due to be released by the end of July) to get a clearer sense of how your incoming Year 7s performed against specific questions and assessment strands. Reports are available for the whole cohort, pupil groups (eg disadvantaged) and individual pupils.

4 Use the analysis in the appendix in section 4 to think about how the new scaled scores might function as the input to progress measures.

Primary and secondary leaders

1 **Talk to each other!** However flawed levels were (and they were), they at least provided to some extent a common language, honed over many years, between primary and secondary schools. Secondary schools are understandably at a loss as to what a child with a score of 92 can really do, and how that might differ from a child with a score of 107. Now more than ever it is essential that primary and secondary teachers find ways to share what they know about children as they move to their new school, and so that their new teachers can build effectively on their prior learning. The appendix to this paper provides a starting point, which will be updated in September when we have a better sense of comparison.

2 Look out for further information from DfE in September, including on the distribution of results across the scale and ways in which these results might meaningfully be compared with previous years.

17 https://www.raiseonline.org/datamanagement/QLATemplateSelectionBlank.aspx

4 Appendix: Using KS2 data to predict KS4 results

This appendix is a first attempt to consider what the link between this year's Key Stage 2 results and the same cohort's Key Stage 4 results might look like. Please be aware that the analysis here is extremely speculative, and is based on historical data. This will be revised when more information is received about the distribution curve of this year's results.

The data from 2009+2014 has been used rather than 2010+2015 because 2010 was the year of the Key Stage 2 boycott. This data has been used rather than more recent Key Stage 2 data because it is the latest for which there is a link between Key Stage 2 input and Key Stage 4 output. We know that 'comparable outcomes' operates for Key Stage 4 results so there will be a kind of reverse comparison.

Analysis of 2009 graphs

Graphs 1 and 5: Bar graphs using the Key Stage 2 average Eng and Maths fine score (to 1 d.p. as per Progress 8 & Attainment 8) on the x-axis; graph 1) = numbers, graph 5) = cumulative percentage.

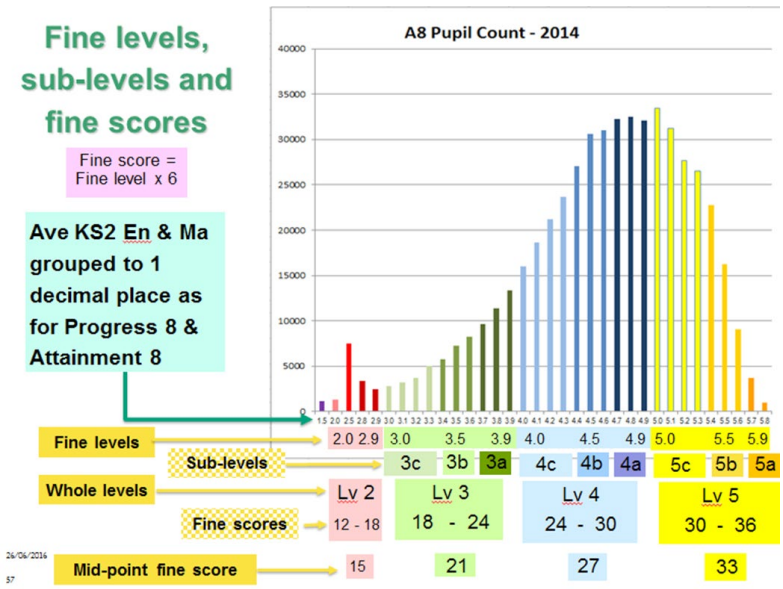
Graphs 2, 3 and 4: Bar graphs by Key Stage 2 sub-levels using the average of three Key Stage 2 scores (English and Maths and Science); graph 2) = %, graph 3) = cumulative percentage, graph 4 = numbers.

From **Graph 3**, in 2009 at Key Stage 2 around 27% of pupils were below Level 4b [fine score 4.33] (approx 'expected standard' = 100), and therefore a broadly similar percentage may be likely, implying around 70% meeting 'expected standard'.

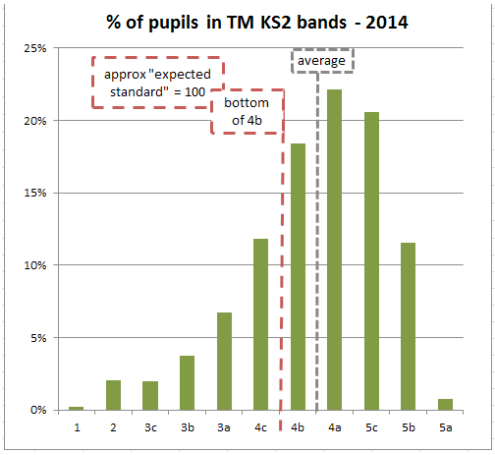
The national average in 2009 was around 4.6, therefore at the top end of Level 4b [fine score 4.66]. We also know that around 17% of pupils attained Level 4b, so there is a reasonable consistency amongst the figures (noting the Key Stage 2 distribution has a long lower tail which means there is a difference between mean and median).

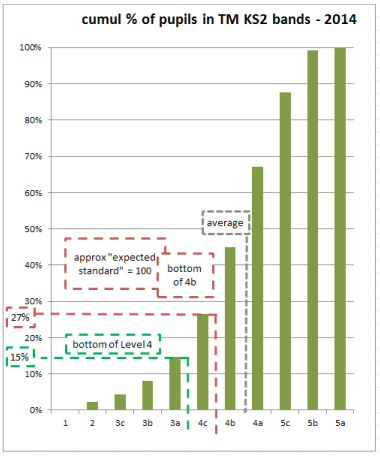
Note that there is a small and subtle difference between the distribution of Key Stage 2 sub-levels (the average of three KS2 scores (English and Maths and Science), and that of the Progress 8 input measure (calculated to the nearest 1 decimal point as average of Key Stage 2 Eng + Maths). There is a degree of overlap between the sub-level bands (where, for example, 4.33 is the divide) and the Progress 8 input (eg 4.3 is between 4.25 and 4.35), as well as subtle differences between averaging two distributions and three distributions.

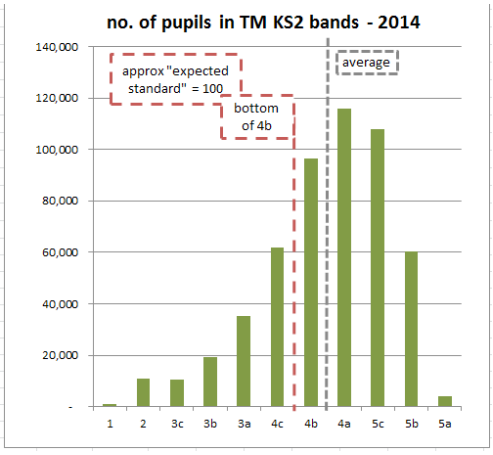
Graph 1 linking Progress 8 KS2 scores with KS2 levels, sub-levels and fine scores (numbers)



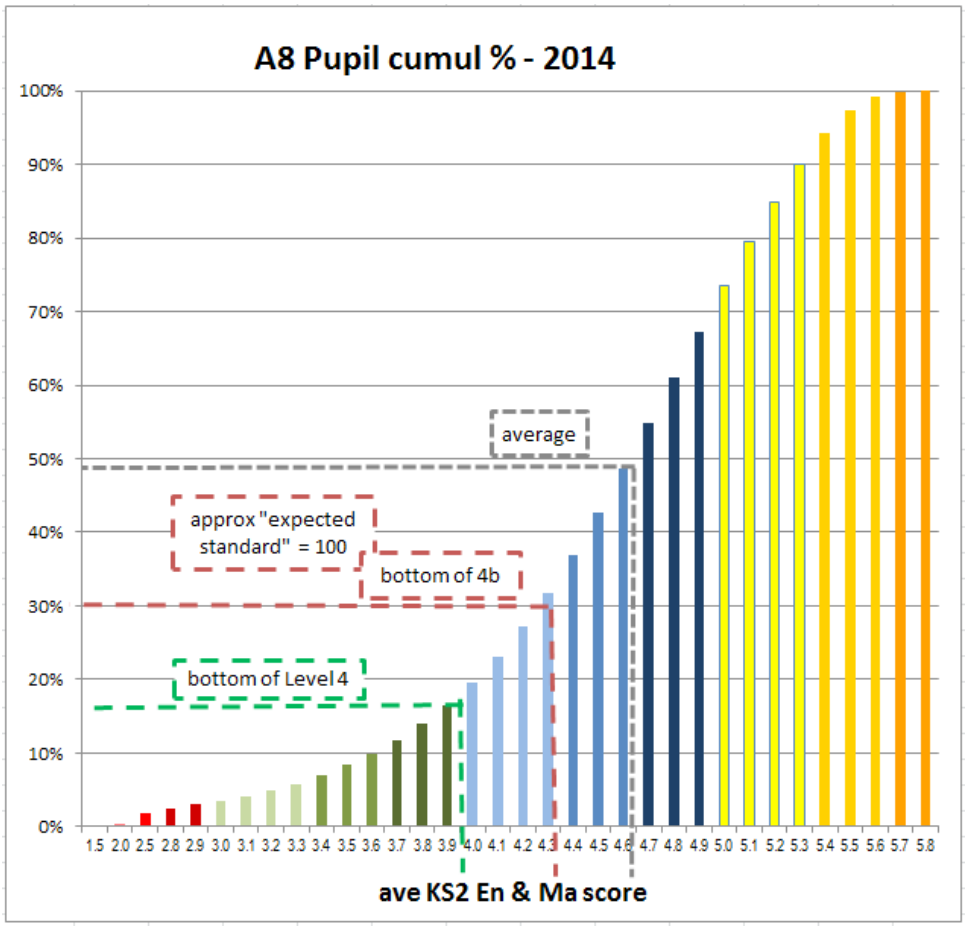
Graphs 2, 3 and 4 by KS2 sub-levels using 2014 KS4 transition matrices with 2009 KS2 input







Graph 5 linking Progress 8 KS2 scores with KS2 levels, sub-levels and fine scores (cumulative percentage)



Data for secondary schools

Secondary schools will be able to go onto the **NCA Tools website**[18] to submit UPN numbers (obtained through the CTF files from feeder schools) and get the scaled scores for the students they know are coming to their school.

Format of files:

1 A CTF file which can be imported into a school's MIS system. The data in this file comes from the ACOMP (https://www.gov.uk/government/publications/assessment-component-files-2016).

2 A comma delimited file which can be imported directly into a spreadsheet such as Excel.

Codes in the datafiles

A = Absent

B = Working below the level of the test

L = Left

M = Missing

T = Unable to access test

F = Pupil will take test in the future

P = Pupil has taken test in the past

H = Pupil cheating

Q = Maladministration

AS = Achieved Standard

NS = Not achieved Standard

CA = Ability to represent their actual ability in the test affected

CN = Ability to represent their actual ability in the test not affected

BLW = Below the standard of the pre-key stage*

HNM = Has not met the standard

PKF = Pre-key stage foundation

PKE = Pre-key stage early development

PKG = Pre-key stage growing development

EXS = Working at the expected standard

A = Absent

D = Disapplied

L = Left

F = Pupil will take test in the future

P = Pupil has taken test in the past

* to be reported with P-scales or NOTSEN as appropriate

5 Further information

How ASCL can help

ASCL Professional Development programmes events to assist with the dissemination and application of data at your school or college.

Getting to Grips with Accountability Measures: Leadership of Data Autumn Conferences **Monday 19 September 2016, London** and **Wednesday 28 September 2016, Manchester**.

Understanding the National Picture and its Implications for you: Leadership of Data Spring Conferences **Tuesday 7 February 2016, London** and **Tuesday 28 February 2017, Manchester**.

Visit **www.ascl.org.uk/conferences** or contact ASCL Professional Development on **0116 299 1122** for more information.

By ASCL Primary and Governance Specialist Julie McCulloch and David Blow, ASCL member and Head of The Ashcombe School in Surrey

Association of School and College Leaders
130 Regent Road, Leicester LE1 7PG
T: 0116 299 1122 E: info@ascl.org.uk W: www.ascl.org.uk

Find us on

July 2016
£4