

Written evidence submitted by CentreForum

About CentreForum

1. CentreForum is an independent think tank that develops evidence-based research to influence both national debate and policy making.
2. Our driving principle is to promote opportunity for all and particularly for those in disadvantaged circumstances. Our work over the next two years will focus on policies to deliver good educational outcomes for all young people; on action to improve support and opportunities for children and young people with mental health problems; and on raising the quality of prison education. We will do this through our evidence based and data-driven research and publications; consultation with commentators, government and those on the frontline; and hosting events that bring together leading thinkers and decision-makers in these important areas of public policy.

Summary of submission

3. We have focused on two areas of the inquiry:
 - The current MATs landscape, including the number, size, and geographical coverage of MATs; and
 - How the performance of MATs should be assessed.
4. We find that:
 - The Government's ambition of full academisation will require significant growth in the number, size and geographical reach of multi-academy trusts;
 - MATs are generally small in size and clustered within particular areas. Just under two-thirds of MATs are sponsor led, and larger MATs with a greater geographical spread tend to be linked with an academy sponsor;
 - Under CentreForum's assumptions of how multi-academy trusts might grow and expand, we estimate that there would be capacity for approximately two thirds of schools that are currently maintained by local authorities;
 - The DfE's proposed performance measures for MATs are a good basis on which to assess the performance of multi-academy trusts. We also consider that: DfE should also publish measures for local authorities; special schools should be incorporated; separate measures by pupil characteristics should be considered.

- Based on the limited data that DfE have published it is the spread between the very best and the very worst, whether that's academy chains or local authorities, that is most striking. More so than any differences between chains and local authorities. Tackling these differences is important regardless of how the academies system now expands. As well as for use in accountability these measures should be used to identify the best and worst performers and further research is needed on the drivers of their differing performance.

The current MATs landscape, including in terms of the number, size, and geographical coverage of MATs¹

5. The Government's White Paper "Education Excellence Everywhere" set out the intention that in a move to a fully academised system most schools will join multi-academy trusts² though they have more recently clarified that for many schools there will be no obligation to do so.³

The number and size of MATs

6. By March 2016 there were a total of 3,609 academies, free schools, UTCs and studio schools recorded as being within a multi-academy trust; this represents around two-thirds of the total of such schools. There were a total of 973 multi-academy trust arrangements recorded in Edubase meaning that the average MAT has 3.7 schools.
7. The white paper states that "*on average MATs can begin to fully develop the centralised systems and functions that will deliver [the benefits of being in a MAT] at a size of around 10-15 academies*".⁴ This estimation appears to be derived from an assumption around the number of pupils (and hence the level of funding) that are required in order to make a MAT sustainable.⁵ What is evident from the department's data is that, on average, MATs that are currently operating are much smaller than the suggested level at which the benefits of being in a MAT are evident.
8. Furthermore, this average masks the number of very small MATs. *Figure 1* groups MATs by the number of schools open within them. The vast majority of these are small in size and in many cases have just a single academy (the funding agreement often being set up as a MAT so as to allow additional

¹ This analysis is based on data held within Edubase, the Department for Education's register of educational establishments, as at 4 April 2016. Individual multi-academy trusts have been identified through the "Trusts (name)" field in this dataset. Analysis is restricted to academies that had opened by 31 March 2016 as it appeared that not all of the trust data for academies that opened in April had been updated at that point. This will not have any material impact on the conclusions of the analysis.

² DfE(2016) "Education Excellence Everywhere" p.58

³ <https://www.gov.uk/government/news/10-facts-you-need-to-know-about-academies>

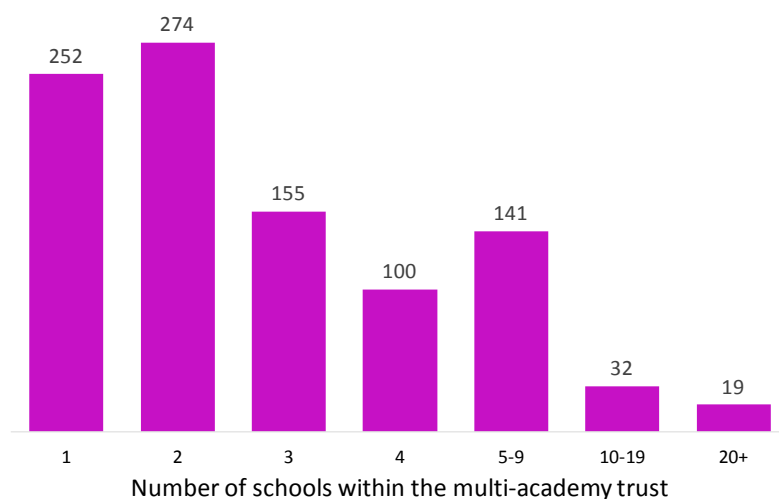
⁴ DfE(2016) "Education Excellence Everywhere" p.58

⁵ Presentation from Sir David Carter "The Leadership Challenges we face if we are to create a world class education system"

schools to be added at a later date.) Over half of multi-academy trusts (526) currently have just one or two schools.

9. There are relatively few large multi-academy trusts – which we define as having 10 or more schools– and almost all are linked with academy sponsors including the likes of ARK, Harris, Oasis and Ormiston. The largest trust that is not recorded as being linked with an academy sponsor has 11 schools.

Figure 1: Number of multi-academy trusts by size of trust



10. Full academisation would require the significant expansion of existing multi-academy trusts and the development of new ones. The white paper sets out some of the ways that the Department for Education believes this will be achieved including:

- the recruitment of new sponsors and incentivising existing sponsors to expand through the Sponsor Capacity Fund;
- providing support to MATs to enable them to grow sustainably such as through Future Leaders; and
- expecting some local authority staff to join existing MATs or set up new ones.⁶

11. The white paper does not quantify the additional capacity that will be introduced through each of these routes, indeed it would be very difficult to do so at this stage. What we do know is that full academisation will mean approximately 16,000 schools that are currently maintained by local authorities becoming academies. By considering the current profile of MATs by their size we estimate that:

- Developing *new* MATs that match the number and size of existing small MATs (fewer than 10 academies) would create an additional capacity for around 2,500 academies;

⁶ DfE(2016) "Education Excellence Everywhere" pp. 83-84

- Growing all existing MATs that are currently below DfE's suggested size of 10-15 schools up to that level would create an additional capacity for around 6,500 academies;
 - Doubling the number of large (more than 10 academies) MATs would create an additional capacity for 1,000 academies; and
 - Growing all existing large (more than 10 academies) MATs by 50% would create an additional capacity for around 500 academies.
12. All of these taken together would create an additional capacity in the system for around 10,500 academies. In other words, under what appear to be ambitious approaches to expansion of multi-academy trusts, there would still only be capacity for approximately two thirds of schools that are currently maintained by local authorities.

Geographical spread

13. The white paper puts forward a number of arguments about the benefits of MATs compared with the local authority model. Amongst these is the fact that multi-academy trusts need not be restricted by historic local authority boundaries:

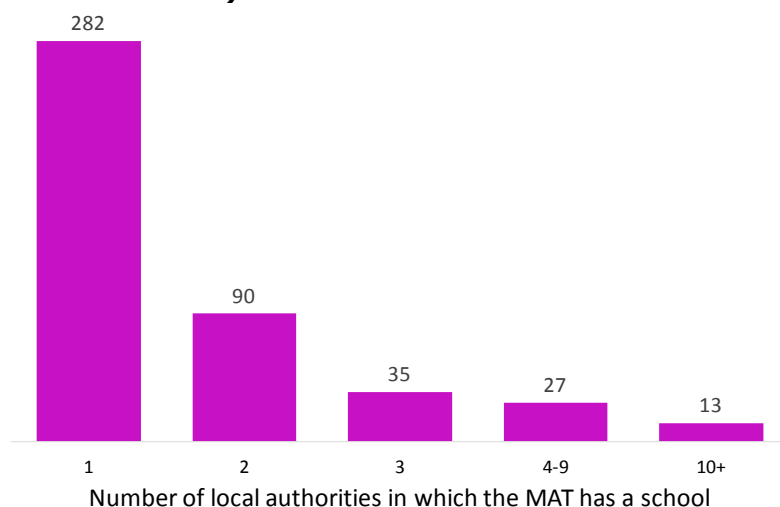
*"If performing well, MATs can scale their success nationwide, taking effective models from one part of the country to the toughest areas in a way that no high-performing local authority ever could."*⁷

14. CentreForum's recent Annual Report highlighted that there is considerable variation in the performance of schools between local authorities and between regions.⁸ By the end of secondary school there is a clear north-south divide with the highest performance found in London. The reasons for these disparities are complex and are likely to result from a combination of factors at pupil, school and area level. Nevertheless, within this context the flexibility to apply successful approaches from one area to another is to be welcomed.
15. Delivering this ambition would require a different approach to the current structure of multi-academy trusts. Most MATs are based on a relatively small geographical area, frequently contained within just one local authority. *Figure 2* groups MATs by the number of local authorities that their schools are based in, the analysis is restricted to those MATs with at least three schools (giving a total of 447 MATs). Nearly two thirds are currently operating within just one local authority area.

⁷ DfE(2016) "Education Excellence Everywhere" pp. 58-59

⁸ CentreForum(2016) "Education in England: Annual Report 2016"

Figure 2: Number of multi-academy trusts grouped by the number of local authorities in which they currently have schools (restricted to MATs with at least three schools)⁹



16. Given that oversight of academy sponsors and tackling underperformance rests with Regional Schools Commissioners it is also worth considering how MATs cut across RSC boundaries. Of the 447 MATs with at least three schools nearly 90% (391) are based entirely within one region. There are nine MATs which cut across four or more of the RSC regions. It is the larger, more established, academy sponsors that have greatest geographical reach. Oasis have schools in 19 local authorities, United Learning 24 and AET 25.

Composition of MATs

17. Given the way that the academies programme has grown the earliest MATs consisted entirely of secondary (or all through) sponsored academies. The majority of MATs now include at least some primary provision and similarly a majority include converter academies.
18. In a large number of MATs there is a limited mix of provision. *Figure 3* groups multi-academy trusts by the phases of schools that are contained within them. Of those MATs with at least three schools, 188 consist entirely of either primary, secondary or special.¹⁰ That is to say that in just over two fifths of MATs there is currently no phase mix. Similarly, *Figure 4* shows that there are 123 MATs with only converter academies and 23 that only serve sponsored academies

⁹ Source: Edubase, March 2016

¹⁰ For this analysis secondary includes all-through schools and the small number of 16+ provision. Special includes alternative provision.

Figure 3: Number of multi-academy trusts by the mix of provision that they currently have (restricted to MATs with at least three schools)¹¹

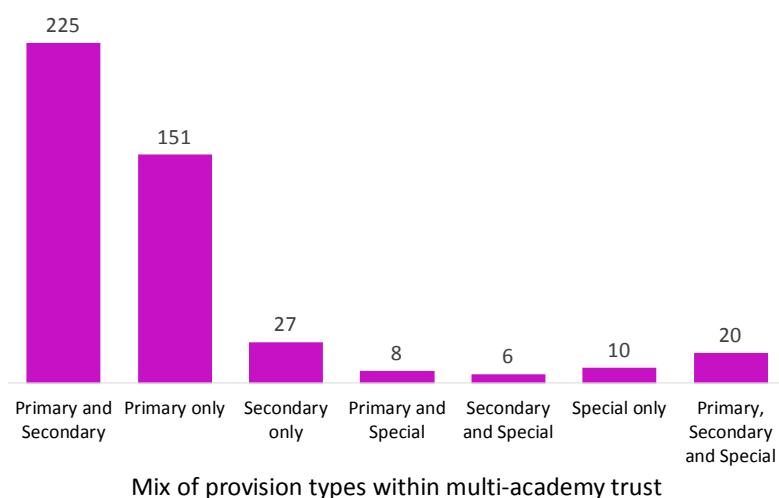
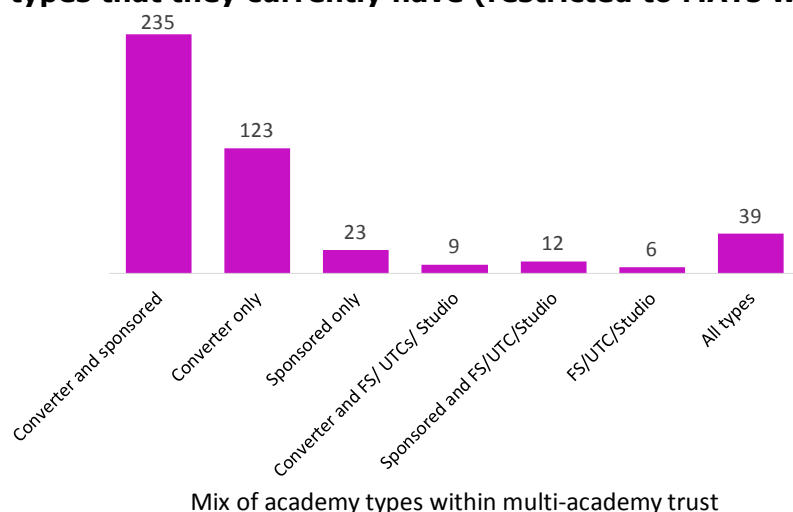


Figure 4: Number of multi-academy trusts grouped by the mix of academy types that they currently have (restricted to MATs with at least three schools)¹¹



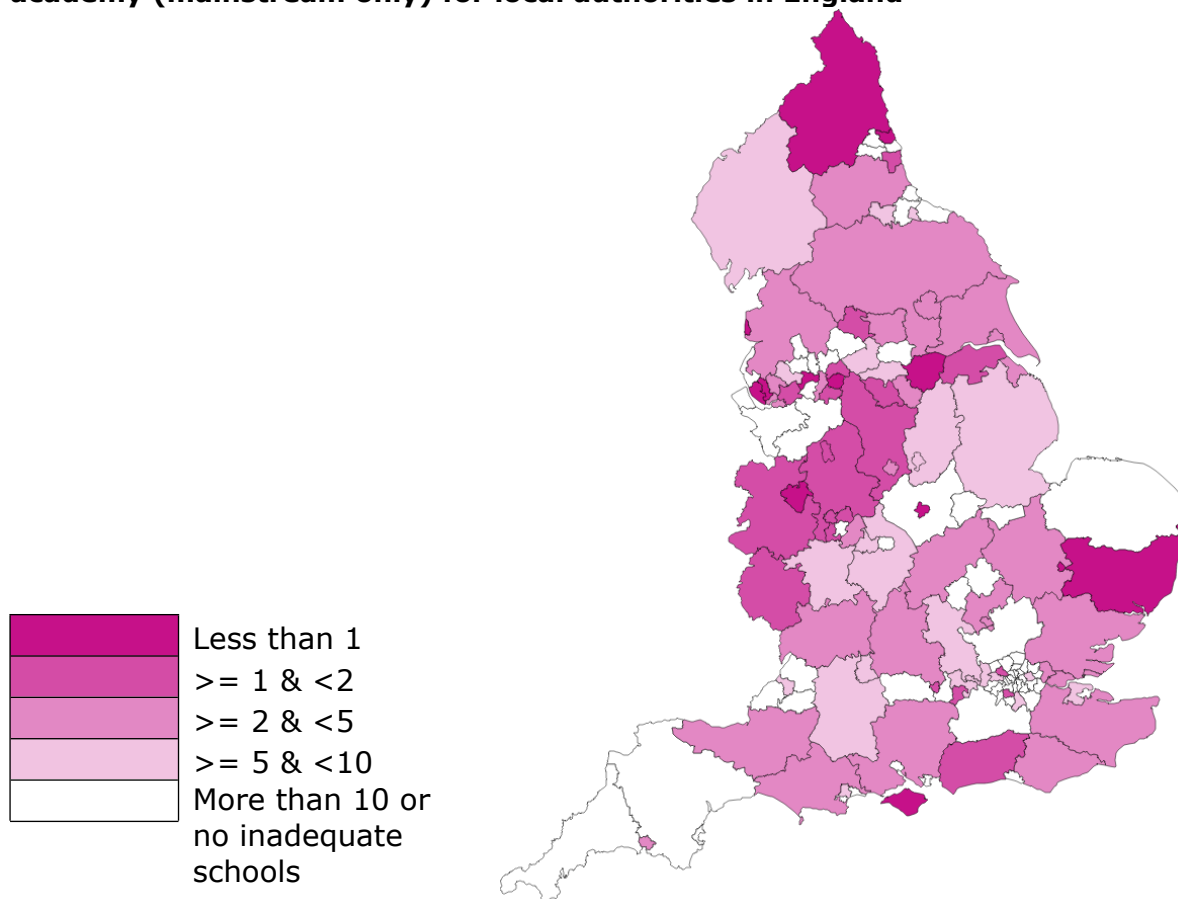
Where are MATs needed?

19. Chapter 5 of the white paper set out the ambition of “High quality sponsors, where they are needed.” We have no direct measures of sponsor capacity within a particular area. Whilst DfE publish a list of approved sponsors this does not provide information on the areas in which they are able to operate nor the capacity or appetite they have for taking on additional schools.
20. We can calculate a proxy measure based on the number of outstanding academies that are within an area and how that compares with the number of failing local authority maintained schools and academies. This shares some of the features of the approach adopted by DfE in their proposed measures for Achieving Excellence Areas.¹² In *Figure 5* we plot, for each local

¹¹ Source: Edubase, March 2016. For the purposes of this analysis ‘secondary’ includes secondary, all-through and 16-19 provision. ‘Special’ includes special schools and alternative provision.

authority area, the number of outstanding academies there are per failing maintained school or academy. Whilst this measure is fairly crude it does give an indication of the presence of high performing sponsors or high performing academies that may be able to partner with other schools. It illustrates the inconsistent coverage across the country and the need to develop high quality multi-academy trusts in areas where coverage is weakest.

Figure 5: Number of outstanding academies per failing maintained school or academy (mainstream only) for local authorities in England¹³



¹² DfE(2016) "Defining 'achieving excellence areas'" <https://www.gov.uk/government/publications/defining-achieving-excellence-areas-methodology>

¹³ Source: Ofsted monthly management information March 2016 <https://www.gov.uk/government/statistical-data-sets/monthly-management-information-ofsteds-school-inspections-outcomes>

How the performance of MATs should be assessed

22. The white paper includes a commitment to:

*"Launch new accountability measures for MATs, publishing MAT performance tables in addition to the continued publication of, and focus on, inspection and performance data at individual school level."*¹⁴

23. To date there has been no published assessment of performance at multi-academy trust level as part of the established school accountability frameworks (that is to say school performance tables or Ofsted.) This development is therefore to be welcomed particularly the commitment to continue to publish data for individual schools.
24. However, the white paper lacks any detail as to how the Department for Education will assess the performance of multi-academy trusts or the timescales for implementing this commitment. Indeed, it does not talk explicitly about what performance they are trying assess. Performance of a multi-academy trust is multifaceted covering dimensions including the outcomes that pupils achieve, the retention and development of staff, whether the trust represents value for money and how expansion is managed.
25. Given the link with performance data for individual schools it is reasonable to assume that the intention refers to the outcomes that pupils achieve and that is the focus of this response. That is not to dismiss those other aspects and the Department for Education should set out how those are being monitored as part of the oversight role of the Regional Schools Commissioners.

Principles of measurement

26. Summarising performance data at trust level in a meaningful way, and in a way that will be understood by users, is challenging. It is certainly preferable that results at chain level relate directly to the results of the individual schools. This creates a clear line where improvements at one level lead to improvements at the other, though it can be difficult to disentangle what effect is down to the performance of a school, and what is down to the MAT (the schools might perform well regardless of the actions of the MAT.)
27. Simple aggregations of school level attainment measures (for example the proportion of pupils that achieve five good GCSEs) risk introducing perverse incentives. This is because a chain would have a disincentive to take on a low performing school – since it would likely pull its average performance down – and an incentive to take on a high performing school. Headline measures should take account of a school's starting point and capture in some way the improvement that has occurred under the MAT.
28. Measures of attainment often reflect the intake of a school rather than the impact that the school is having. Historic 'Progress' measures based on expected progress, e.g. the proportion of pupils making two levels of progress between

¹⁴ DfE(2016) "Education Excellence Everywhere" p.104

key stage 1 and key stage 2, are often as flawed as attainment. The propensity to make expected progress is strongly correlated with prior attainment (e.g. those with high attainment on entry are far more likely to then make at least the expected progress.) Therefore, fairer measures consider pupil attainment in comparison to pupils with similar prior attainment such as value added, or progress 8 – though even these can be subject to bias.

29. Finally, if the measures are to be credible it is essential that they can be understood by a wide range of users including by schools, academy trusts, RSCs and the wider public. Therefore, whilst the underlying methodology may be complex, the outcome should have a 'real world' interpretation.
30. We therefore have four principles that we believe a performance measure for a MAT should reflect. We believe that:
- the performance measure for a chain should link directly to performance measures for the individual schools to ensure consistent priorities;
 - measures should not create disincentives to take on more challenging schools and should therefore reflect the improvement (or otherwise) seen in a school since joining the MAT; and
 - efforts should be made to account for the fact that there is variation between the intakes of schools; and
 - the final measure should have a 'real world' interpretation.
31. As well as measuring the performance of MATs across all pupils, measures would ideally be broken down by pupil characteristics in particular disadvantaged pupils. This would be consistent with traditional measures of performance and protect against gaps growing for particular groups.

DfE's proposed measures

32. Immediately prior to the 2015 General Election DfE published a statistical working paper with proposed measures of performance within academy chains and within local authorities.¹⁵ The working paper was intended to open a debate about how performance might be measured. The department has since made no further comment on these measures and whether they intend to use them for MAT accountability. DfE proposed two measures of performance:
- How well schools in a given chain or local authority are currently performing based on current value added scores; and
 - How that performance has changed over time by looking at improvement in value added.

¹⁵ DfE (2015) "Measuring the performance of schools within academy chains and local authorities"

The statistical working paper outlines the proposed measures in some detail. We briefly outline the underlying methodology here.

Underlying data - value added

33. DfE's proposed measures use school level value added scores. While there are a variety of approaches to measuring value added they all share the same principle that they measure a school's performance by comparing the outcomes that pupils achieve in comparison to pupils with similar prior attainment nationally, school scores are then the average of all pupil scores.
34. At key stage 4 a school's value added score is centred around 1000. A score above 1000 means that pupils achieved higher results than pupils with similar prior attainment nationally, a score of below 1000 means that pupils achieved lower results than pupils with similar prior attainment nationally. At key stage 2 measures are centred around 100 and are interpreted in a similar way.

Measure 1: How well schools in a given chain or local authority are currently performing based on current value added scores

35. DfE's first measure attempts to capture how well schools within a chain or local authority are currently performing. It is the average of the value added scores of schools within the chain or local authority, these are weighted by:
 - pupil numbers, so that a schools contribution to the overall measure is proportional to its size (i.e. larger schools carry more weight); and
 - length of time the school has been with the chain or local authority meaning that those that have been with the chain or LA the longest carry more weight than those that have recently joined. Schools open less than a year are excluded entirely.
36. These scores are centred around 1000 and interpreted in the same way as measures for individual schools.

Measure 2: How that performance has changed over time by looking at improvement in value added.

37. The second measure looks at how the value added scores for schools have changed over time and in particular how the performance of schools within academy chains has changed since they joined the chain.
38. There are some complexities in looking at changes over time. For example, it would be possible to look at a schools score in a baseline year and how much it has changed since that point. However, schools with the lowest starting point tend to see the biggest increases and those with the highest starting point the greatest falls. This is an effect that is known as 'regression to the mean'. This is countered in DfE's proposed approach by grouping schools by previous value

added outcomes and comparing improvement for a school to those who started from a similar position. Again the scores are weighted by the number of pupils and the length of time that a school has been with a chain or local authority.

CentreForum’s view of the proposed measures

39. We believe that the DfE’s proposed measures are a good basis on which to assess the performance of multi-academy trusts. They broadly meet our four principles for measurement:

Principle	Assessment of proposed DfE measure
The performance measure for a chain should link directly to performance measures for the individual schools to ensure consistent priorities	The measures are based on value added measures that are used at school level. However historically value added has not been seen as the key headline measure by schools (instead focussing on 5+A*-C including English and maths). See further discussion below.
Measures should not create disincentives to take on more challenging schools and should therefore reflect the improvement (or otherwise) seen in a school since joining the MAT	The improvement measure compares performance to schools with a similar starting point and so minimises disincentives to take on challenging schools. The average VA measure could potentially cause a disincentive if used in isolation.
Efforts should be made to account for the fact that there is variation between the intakes of schools	The use of value added controls for pupil prior attainment to reduce bias caused by different intakes.
The final measure should have a ‘real world’ interpretation	Whilst the methodology is complex the final measures can be interpreted in terms of GCSE outcomes.

40. There are however a number of areas that DfE should consider further if they are to use these measures for MAT accountability.

40.1 Comparison with local authorities:

The white paper makes no mention of performance measures for local authorities. The majority of state-funded schools remain maintained by local authorities and will do for some time. Crucially, the white paper proposes that in “local authorities which are underperforming... we will take new powers to ensure schools become academies to a faster time scale” but gives no indication of what underperformance for a local authority means within this context.¹⁶ Therefore, there is a need for an assessment of the performance of local authorities and we believe that it should be done in a way that is consistent with multi-academy trusts.

¹⁶ DfE(2016) “Education Excellence Everywhere” pg 55.

40.2 Primary schools and post-16 provision:

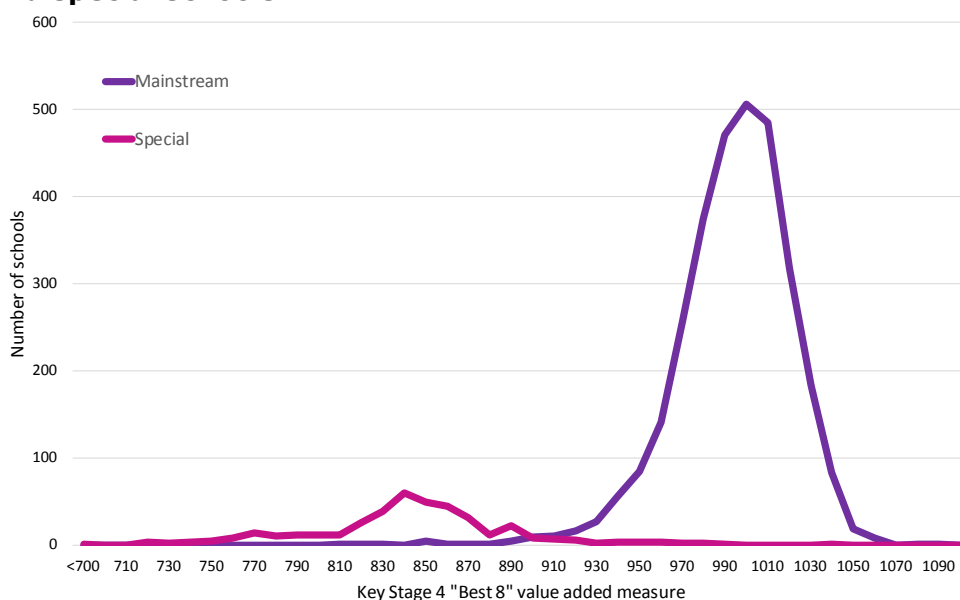
The working paper suggests that a similar methodology could be applied to primary schools once the principles are established. We believe that it is essential that primary academies are considered within performance measures. These should also be based on measures of value added. Whilst we recognise that there are issues with the reliability of key stage 1 as a baseline it is still better than looking at attainment alone (and will remain as a baseline for school level performance measures.) We also recognise the challenges in creating measures for performance at key stage 5 but the department should continue to consider how these might be incorporated.

Ultimately DfE should publish measures separately for each phase for each MAT and then an overall measure, this will allow users to identify the relative strengths of MATs for different phases whilst still having an overall assessment.

40.3 Special schools:

The statistical working paper does not set out an intention to capture the performance of special schools. There are now nearly 200 special academies and free schools, over half of which are in multi-academy trusts. Performance measures for these schools are published in the performance tables and so there is a precedent for comparing on a similar basis. However, even when comparing to pupils with similar prior attainment pupils in special schools generally make slower progress. This is demonstrated by looking at the distribution of key stage 4 value added scores for mainstream and special schools. Figure 6 plots the number of schools at each point on the value added distribution.

Figure 6: Distribution of 2015 key stage 4 value added scores for mainstream and special schools¹⁷



It shows that mainstream schools, as expected, are centred around the national average of 1000, special schools are centred around a score of around 850. In this instance, simple value added is a poorer assessment of school effectiveness and therefore MATs which include special schools are likely to see their overall

¹⁷ DfE(2016) Secondary School Performance Tables

value added come down with their inclusion even if those schools are actually performing well.

This is also an issue with the improvement in value added measure since the assumption that schools with similar value added have a similar level of effectiveness and hence capacity to improve is violated. A further complication is the relatively small cohorts that are seen in special schools, this means that value added measures are subject to wide confidence intervals and in many cases they do not even have a score published.

Despite these challenges it is important to capture the performance of special schools. The department may wish to consider a separate improvement model for them meaning that they are only compared with other special schools (though even this has issues as the type of SEN provision can vary considerably.) We would advise against including within the overall value added measure unless the underlying value added model is revised to better reflect the relative rates of progress of pupils in special schools.

40.4 Number of schools required within a MAT in order to be published:

The working paper restricted analysis to those academy chains with at least five schools with results. Our own analysis has shown that such a rule means a large majority of MATs would be excluded from performance measures. We believe that it would be appropriate to publish data where a MAT has three schools with results.

40.5 Choice of performance measure:

As set out above we believe that using value added as the basis for these measures is the fairest way that is currently available. It means that the performance of pupils is compared to that of similar pupils nationally. However, one of our criteria for a successful measure is that it links directly with measures of accountability at school level. Historically the key measures have been based on measures of attainment with some inclusion of progress – e.g. schools are below the floor standard at key stage 4 if less than 40 per cent of pupils achieve five good GCSEs including English and maths and progress is below average.

Far less emphasis is placed on measures of value added. This means that schools have had an incentive to behave in a particular way – for example, trying to push pupils at the D/C boundary above that level rather than moving pupils from an A to an A*. This effect is countered by the introduction of progress 8 which captures performance across a range of subjects after controlling for prior attainment and rewards success at each point of the attainment distribution. From 2016, floor standards will be based on Progress 8 for all schools. In time the Department should use progress 8 as the basis for its measures of MAT performance, this will ensure a clear link between performance at school and at MAT level. However, we would caution against applying progress 8 retrospectively as it would be unfair to judge schools using an accountability system that was not in place at the time. Schools would have been likely to make different curriculum choices if progress 8 had been in place.

40.6 Impact on disadvantaged pupils:

The working paper indicates that the department will consider whether there are further breakdowns of the measure that could be published (such as for disadvantaged pupils.) We believe doing so would be a positive step as single headline measures for all pupils risk masking considerable within MAT variation.

What the data from the published chain and local authority measure currently tells us

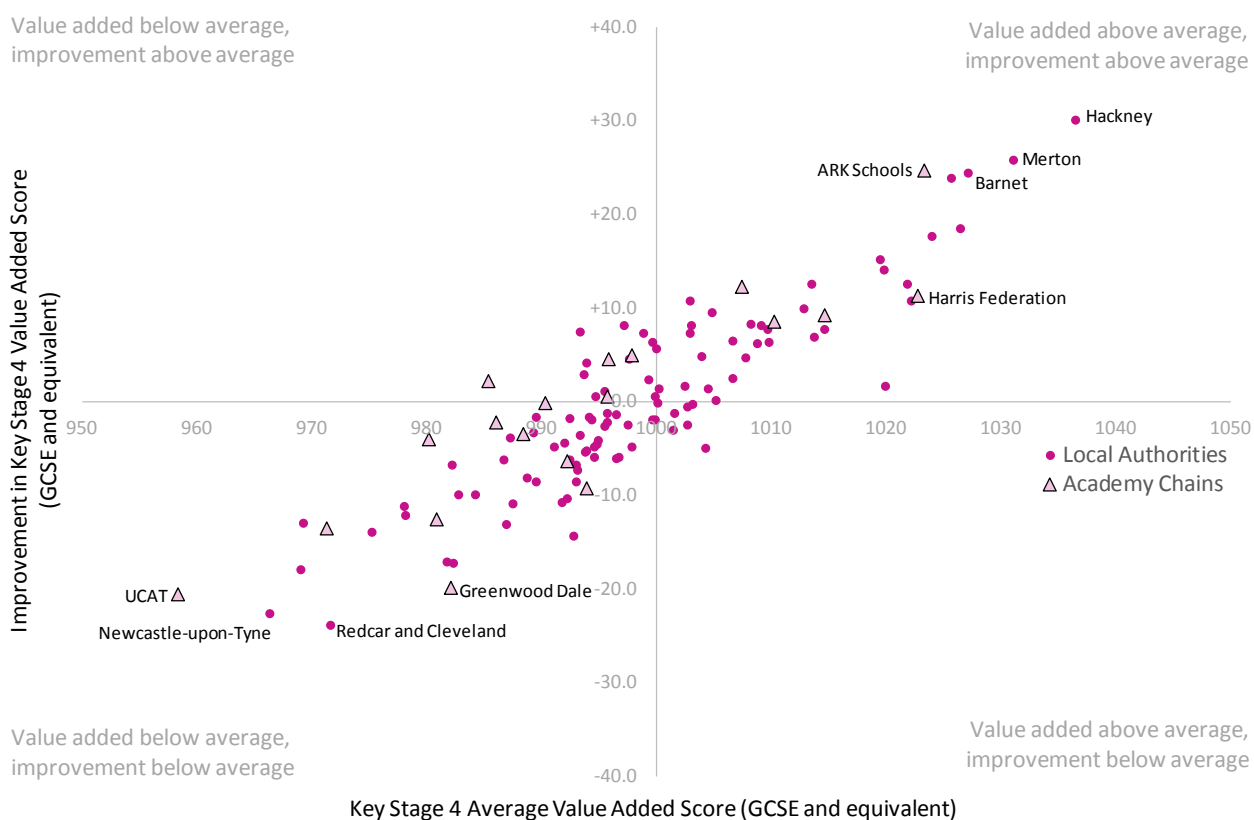
41. The DfE's statistical working paper provides data on academy chains and local authorities with at least five schools with results giving data for 20 chains (19 chains on improvement measures) and 100 local authorities, though it gave no commentary on the results.¹⁸
42. Given that this is a relatively small subset of academy chains and only covers results for one year, caution should be applied when interpreting these results but what the data does show is:
 - There are examples of both high and low performing chains and local authorities;
 - The variation within each group is considerable, the difference between the improvement rates in the highest and lowest performing chains is equivalent to just under a grade in every subject a pupil takes, the difference is slightly higher still for local authorities;
 - This variation is much more important than the variation between the two groups; and
 - The highest performing local authorities are in London, the highest performing chains have a strong presence in London.
43. *Figures 7 and 8* plot the improvement in value added against current value added on a GCSE and equivalent and GCSE only basis. Academy chains and local authorities that appear in the top right hand quadrant are above average on both current and improvement in value added. Figure 10 then groups academy chains and local authorities by the proportion of each that are in the top and bottom 25% of all academy chains and local authorities in the data.
44. Local authorities and chains are equally likely to be amongst the top performers in GCSE and equivalent value added but chains have a greater propensity to be amongst the lowest performers, this might reflect that those chains haven't yet been able to turn around poorly performing schools. Chains are under-represented amongst the top performing on the GCSE only value added measure but they are over-represented on the improvement measure on this basis.
45. It is the spread between the very best and the very worst, whether that's academy chains or local authorities, that is most striking. More so than any

¹⁸ DfE (2015) "Measuring the performance of schools within academy chains and local authorities" pp.48-57

differences between chains and local authorities. For example, the difference in improvement measure between ARK and UCAT amounts to just under a grade in every subject.

46. Tackling these differences is important regardless of how the academies system now expands but is particularly important if MATs are expected to grow quickly.¹⁹ As well as for use in accountability these measures should be used to identify the best and worst performers and further research is needed on the drivers of their differing performance.

Figure 8: Current performance and improvement of schools in academy chains and local authorities – GCSE and equivalent 2014



¹⁹ See for example the CentreForum article "Are we reaping the consequences of academy chains growing too fast?" <https://www.tes.com/news/school-news/breaking-views/are-we-reaping-consequences-academy-chains-growing-too-fast>

Figure 9: Current performance and improvement of schools in academy chains and local authorities – GCSE only 2014



Figure 10: Distribution of academy chain and local authority scores on the four measures proposed by DfE

