Value Added Measures of School GCSE Performance
Value Added Measures of School GCSE Performance

A Research Report produced by Professor David Jesson, York University.
FOREWORD

The Government is committed to working towards giving all children the first class education to which they are entitled. This means encouraging teachers and governors to work towards excellence and to recognise achievement. It is therefore essential that we give teachers the tools to enable them to improve standards and meet targets. We want to help all children reach their maximum potential to enable them to succeed in life and contribute to the economy of the country.

To help develop these tools my Department has sponsored a project, undertaken by Professor David Jesson of York University with the support of the Technology Colleges Trust, to look at value added data with the help of 31 secondary schools, including 12 CTCs. Besides testing the usefulness of KS3 assessments in estimating performance at GCSE the project also explored the potential for using this data as a benchmark to map progression from KS3 to GCSE. The project was carried out in parallel to SCAA's Value Added National Project, and bears out a number of the findings of the larger project. SCAA is now consulting on the establishment of a national value added system to cover all four key stages from 1998.

The information provided by this project will support the development of the national system by showing how value added data can be used as part of the essential self-evaluation by performance of the school. This will allow teachers to be more aware of the effect their work is having on their pupils and to address local problems which might impede pupils reaching the GCSE standard indicated at KS3 stage. The project should, of course, be seen in the wider context of work on school improvement. It is a significant step forward, and I look forward to seeing the results of its use.

Stephen Byers MP
Minister of State for Education and Employment
Introduction

This is the final Report of an investigation into the feasibility, the practice and the outcomes of assessing the effectiveness of schools in delivering GCSE qualifications to their pupils over the critical two year period from Key Stage 3 to the end of Year 11.

The Interim Report (HMSO; 1996) showed clearly that it was possible to collect, match and evaluate evidence from individual pupils whose relevant National Curriculum Key Stage 3 test and GCSE examination records were available. However, its findings were based on a relatively small opportunity sample of schools which happen to have the appropriate data available. Therefore, it could do little more than point to the potential for providing the value-added measures to which this Project's title refers.

The second phase was able to utilise the 1994 Key Stage 3 results of around 5000 pupils in over thirty schools, a much larger sample, although still not necessarily fully representative nationally. These KS3 'scores' were then matched with the pupils' GCSE results taken up to, and including, Summer 1996. The schools involved in this second phase comprised all the City Technology Colleges whose pupils took GCSE examinations in Year 11, together with a substantial number of other schools recruited through the agency of the Technology Colleges Trust. The involved institutions included inner-city, suburban and rural comprehensive schools, denominational and single sex schools along with a small number of selective grammar schools. Although not designed to be a 'representative' sample of Year 11 pupils; they do however, provide a broader frame of reference than those available for analysis in the Interim Report.

The Data

For each pupil a record was constructed which gave information in the following three categories:

A: Background

School attended;

Gender;

Date of birth.

B: Intake Measure

Key Stage 3 scores on tests in English, Mathematics and Science;

Key Stage 3 'levels' derived from test scores in each subject.

C: Outcome Measures
Number of entries to GCSE Examinations;

Number of A* to C Passes;

Number of A* to G passes;

Total GCSE 'points' achieved (A* = 8; A = 7 .... G = 1; Fail/Absence = 0);

Average number of GCSE 'points' per examination entry;

Average number of A* to C Passes per examination entry.

The data on pupil 'intake'

Value-added measures of school performance require an 'intake' measure for each pupil against which to assess their 'outcomes'. For this purpose we constructed a 'Key Stage 3 average score' by summing the (Finely differentiated) levels they achieved in English, Mathematics and Science and averaging this to obtain an 'overall Key Stage 3 level'. The resulting distribution of 'intake' scores is shown in Figure 1.

Schools varied substantially in their 'Key Stage 3 composition; not only was there wide variation in the average of their pupils' scores but also in the range of scores between different pupils within them. In Figure 2 we show this information for each of the Project schools.

From Figure 2 it is clear that some schools have intakes that are substantially 'lower' in Key Stage 3 scores than most, whilst others (mainly grammar schools) have pupils with relatively 'high' intake scores. The BoxPlot also gives some indication of the 'spread' of pupils in each school across the middle fifty percent of the range. This issue is illustrated additionally in Figure 3 which shows the cumulative distribution of 'intake' scores for two contrasting schools. School A has a much lower 'intake score' than School B, but the 'spread' of scores between the middle half of its pupils is twice that of School B. The issue of school composition and its impact on pupil performance is of considerable substantive importance in the current debate about 'selection' for schools; we shall return to this matter at a later stage of our analysis.

Data on pupils' outcomes

Schools differed considerably in the numbers of GCSE examinations they offered their pupils, and this gave rise to the need to differentiate between 'quantity' of outcome (as measured say, by the total number of A* to C passes which pupils achieved, or by their total points score) and its 'quality' as evidenced by the 'amount of outcome' for each GCSE examination entered.

'Quantity' measures of outcome are clearly likely to favour schools which present their pupils with opportunities for taking more, rather than less, GCSE examinations, unless, of course, there is a threshold beyond which a law of diminishing returns exists for the pupil concerned. On the other hand, the actual grades which pupils achieved in each examination for which they were entered, provide an important outcome in their own right, and should not be assimilated into a single figure which, by virtue of administrative differences between schools, conceals this
Distribution of KS3 Levels for Project Schools and Colleges
(based on 4600+ pupils in 31 Schools)

Figure 1
Distribution of Pupils' Key Stage 3 Levels in Project Schools and Colleges

**Figure 2**
Cumulative Percentage Distribution of KS3 Levels for two Project Schools

Note differences in 'Average (50%) Level' and 'Spread' (25% to 75%)
information. It seemed therefore important to keep both of these approaches to assessing pupils' and schools' outcomes so that, if there was evidence that they prompted different views about the same school's performance, we should be alerted to this. It is additionally important to keep the issue of the 'quality' of outcomes at the forefront given the increasing importance of non-GCSE qualifications, particularly in schools with strong relationships with their local commercial and industrial communities who are implementing GNVQ courses. One impact of taking examinations in these areas is to reduce the number of GCSE entries which pupils taking these courses make.

Under these circumstances we have created four outcome measures;

I Measures of 'Quantity'

(i) Total GCSE points score; and,

(ii) Total number of A* to C grades.

II Measures of 'Quality'

(iii) Average number of GCSE points per GCSE examination entered; and

(iv) Average number of A* to C passes per GCSE examination entered.

How well do Key Stage 3 'Levels' relate to GCSE Performance?

We present these outcomes in Figures 4a to 4d which show each of them in relation to pupils' Key Stage 3 levels. In each figure it is very noticeable that the 'outcome' measures cluster fairly tightly together at each level of pupils' intake scores. This is greater in the case of the 'quality' measures, since the additional variable of the number of examinations for which pupils entered does not so obviously affect the outcome measure used.

Outcomes at Year 11

The four measures of performance which we have identified above, are highly correlated with each other for individual pupils. This is entirely to be expected since pupils are grouped within schools where the differences in examination entry policy between pupils are very much less than they are between schools. Pupils' outcomes, therefore, represent four distinct but comparable measures of their performance. The correlations between these pupil outcome measures are shown in Table 1 below.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Total Points</th>
<th>Average Points</th>
<th>No of A* - C Passes</th>
<th>Average A* - C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Pts</td>
<td>1.00</td>
<td>0.93</td>
<td>0.93</td>
<td>0.87</td>
</tr>
<tr>
<td>Avge Pts</td>
<td>1.00</td>
<td>1.00</td>
<td>0.90</td>
<td>0.92</td>
</tr>
<tr>
<td>A* - C</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.97</td>
</tr>
<tr>
<td>Avge A* - C</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

At school level, however, the situation was rather different. Schools entered their pupils for
Pupils' GCSE Outcomes in relation to Key Stage 3 Average Levels
Total Number of GCSE Points achieved

Figure 4a
Pupils' GCSE Outcomes in relation to Key Stage 3 Average Levels

Average Points per Examination Entered

Figure 4b
Pupils' GCSE Outcomes in relation to Key Stage 3 Average Levels

Number of A* to C Grades achieved

Figure 4c
Pupils' GCSE Outcomes in relation to Key Stage 3 Average Levels
Number of A* to C Grades achieved per Entry

Figure 4d
examinations in a fairly consistent manner - but, the *average level* differed substantially between them. This imposes a different framework for the outcomes which 'similar' pupils achieved in different schools. For example, in a school where ten examination entries was the norm, a total points score of 45 would represent a pattern of individual results around the C/D level. The same total score in another school where only eight entries was 'normal', such a total points score would represent a possible combination of B and C grades; a very different situation in terms of the 'quality' of the outcomes achieved. The correlation between these school outcomes is shown below in Table 2 below.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Total Points</th>
<th>Average Points</th>
<th>No of A* - C Passes</th>
<th>Average A* - C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Pts</td>
<td>1.00</td>
<td>0.86</td>
<td>0.94</td>
<td>0.88</td>
</tr>
<tr>
<td>Avge Pts</td>
<td></td>
<td>1.00</td>
<td>0.93</td>
<td>0.99</td>
</tr>
<tr>
<td>A* - C</td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.95</td>
</tr>
<tr>
<td>Avge A* - C</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

It is clear from Table 2 that, whilst the 'quantity' and 'quality' measures at school level correlate with themselves very highly ($r = 0.94$, $r = 0.99$) the correlations between these types are lower. The greatest difference is in the correlations between Total Points and 'Average Points per Entry' which, at pupil level, was 0.93, and is reduced to 0.86 as a correlation between schools' outcomes on these two measures. Rather less marked considerations apply to the relationship between the total number of A* to C passes, and the average number achieved per examination entered. These observations give substance to our decision to report each of these outcomes independently for each school, since they do appear to measure different aspects of schools' performance.

The diagrams showing the relationships between individual pupil's outcome measures and their Key Stage 3 levels imply that the latter has a considerable influence on the former. In Table 3, below, we have classified pupils into five groups - the 'bottom' ten percent with 'low' Key Stage 3 scores; the next twenty percent with 'below average' Key Stage 3 scores; the middle forty percent with 'around average' Key Stage 3 scores; the next twenty percent with 'above average' scores and the 'top' ten percent with 'high scores' from their Key Stage 3 assessments. Table 3 also shows the average GCSE and other outcome scores achieved by girls and boys in each of these categories.
<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Below Average</th>
<th>Average</th>
<th>Above Average</th>
<th>High</th>
<th>Avge KS3 Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls (%)</td>
<td>10</td>
<td>20</td>
<td>40</td>
<td>20</td>
<td>10</td>
<td>5.08</td>
</tr>
<tr>
<td>Boys (%)</td>
<td>9</td>
<td>21</td>
<td>41</td>
<td>20</td>
<td>9</td>
<td>5.08</td>
</tr>
</tbody>
</table>

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Number of KS 3 Points</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall</td>
</tr>
<tr>
<td>Girls</td>
<td>17</td>
<td>29</td>
<td>43</td>
<td>56</td>
<td>66</td>
<td>42</td>
</tr>
<tr>
<td>Boys</td>
<td>15</td>
<td>26</td>
<td>40</td>
<td>53</td>
<td>66</td>
<td>39.5</td>
</tr>
</tbody>
</table>

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average KS 3 Points per Entry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>2.1</td>
<td>3.4</td>
<td>4.8</td>
<td>5.8</td>
<td>6.6</td>
<td>4.6</td>
</tr>
<tr>
<td>Boys</td>
<td>2.0</td>
<td>3.1</td>
<td>4.4</td>
<td>5.4</td>
<td>6.5</td>
<td>4.4</td>
</tr>
</tbody>
</table>

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em><em>Number of GCSE A</em> - C Passes</em>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>0.6</td>
<td>1.7</td>
<td>5.6</td>
<td>8.8</td>
<td>9.9</td>
<td>5.4</td>
</tr>
<tr>
<td>Boys</td>
<td>0.3</td>
<td>1.4</td>
<td>4.7</td>
<td>8.3</td>
<td>9.8</td>
<td>4.8</td>
</tr>
</tbody>
</table>

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em><em>Average GCSE A</em> - C Passes per Entry</em>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>0.1</td>
<td>0.2</td>
<td>0.6</td>
<td>0.9</td>
<td>1.0</td>
<td>0.57</td>
</tr>
<tr>
<td>Boys</td>
<td>0</td>
<td>0.1</td>
<td>0.5</td>
<td>0.8</td>
<td>1.0</td>
<td>0.51</td>
</tr>
</tbody>
</table>

For each group of pupils, apart from the very 'top', girls outperformed boys by a considerable margin in each of the four outcome measures. This effect is most marked in the 'total points' outcome, but is present in each of the other three outcomes.

**Ways of helping schools review their own performance**

In Figure 5 we have shown the range of GCSE scores achieved by the middle fifty percent of pupils at a number of different 'Key Stage 3' starting points. Extracting the 'trend' line from the top of the 'box' (the upper quartile); its middle point (the median), and the bottom (the lower quartile) we can construct Figure 6, which shows for 'any' Key Stage 3 starting point what was the middle range of GCSE scores achieved by 'similar' pupils two years later.

We have constructed Figure 6 for pupils' Total GCSE Points Score'. The method employed could clearly be extended to each of the outcomes which we are discussing; we will, however, restrict our exploration of the 'technique' to this one measure in the interests of brevity in reporting.

The three 'lines' on the diagram represent a 'reasonable expectation' for the range of GCSE scores within which any pupil from a given Key Stage 3 score might be expected to achieve. Any systematic tendency, in a particular school, for the results for some groups of pupils, or indeed for the many of them, to fall outside the 'tramlines' would be indicative of 'performance' which was characteristically 'different' from that of the majority of pupils 'like theirs' in other schools. If nothing else, this should alert that school to there being a 'performance' issue to be further explored. Of course, the situation where many pupils lie above the 'tramlines' - and thereby
Range of Total GCSE Points for Groups of Pupils (Groups defined by Pupils' Key Stage 3 Levels)

Figure 5
Pupils' GCSE Outcomes in relation to Key Stage 3 Average Levels

Range of Scores achieved by pupils with differing KS3 Levels

Figure 6
indicative of 'additional value' being provided for these pupils by this school - is fundamentally
different in the urgency of its messages than the corresponding one where many pupils lie below
these tramlines. Examples of three different schools' pupils' results are shown in the Appendix
as Figures A, B and C.

There is a strong case to be made for schools to be much more 'open' with performance
indicators such as these - because if there is evidence that particular schools consistently provide
'additional value' for their pupils - then schools which do not achieve these levels of outcome
will be very keen to learn how they might gain a similar dividend for their pupils. This is surely
one vital focus which school improvement programmes must not overlook.

There is, of course, no simple rule for translating the visual impressions gained from identifying
where a schools' pupils are when plotted on this kind diagram into a statistically appropriate
'indicator'. Indeed, there is a degree of arbitrariness in choosing to draw the lines where we have
done - but, at the very least, if schools can accept the discipline of looking carefully at their
pupils results in a framework which allows comparisons with hundreds (or possibly thousands)
of other pupils 'like' theirs it would certainly be a major step forward in utilising constructively
the large amounts of data which are now available to assist in evaluating performance. It would,
at a stroke, defuse much of the disinformation provided by 'performance' tables showing
outcome alone, as well as helping governors, parents and schools themselves to focus on
potential benchmarks for pupils' future performance.

The relevance of diagrams such as those which we have shown for tracking pupils' performance
in each of the outcomes we have discussed should now be obvious - however, the matter does
not even need to rest there.

To the extent that there is a 'common core' of subjects which pupils enter for GCSE
examinations, it may also be possible to track their relative performance in these subjects,
individually, alongside those explorations of 'overall school performance' which have been the
focus of this Report to date.

In Figure 7 we show the outcomes for the Mathematics examinations taken by all the pupils in
our sample. We have adopted a similar procedure in drawing the boxes covering the range of
mathematical outcomes achieved by the 'middle fifty percent' of pupils in each of our indicated
Key Stage 3 score groups. Extracting the 'tramlines' in a similar way to that described above, we
could provide a framework with which schools can check the performance of their pupils. Since
each entry on such a chart represents the outcomes of a known, individual pupil, such
identification may show up interesting and worthwhile indications of performance which almost
no other method can make so easily accessible. It might, for example, be that boys do better
than girls, or that pupils in a particular 'set' achieve well beyond what others might 'reasonably'
expect. And by using the information provided in the diagram along with the additional
interpretative framework which schools' own knowledge of their own pupils can provide - there
exists the possibility that valuable insights into pupil performance, and schools' contributions to
this may be made possible. We have provided in Appendix Figures D to F examples of the
'tramline' diagrams for pupils' performances in English, Mathematics and Science.

There should, however, be a statistical health warning attached to schools' use of such 'evidence'
- recognising the variability implicit in any form of measurement, and particularly so in the case
GCSE Subject Performance in relation to KS3 Average Levels

Range of Scores in Maths for pupils with differing KS3 scores

Figure 7
of so called 'measures' of educational performance. We have no brief for making fine
distinctions between pupils, and any attempt to do so will do a dis-service to the future
credibility of the actions which flow from such.

The evidence base provides an 'indication' of schools' relative performance, not a hard and fast
definition of it.

We have, so far, provided an interpretative tool for schools' use, and have identified some of its
limitations in practice. Each of our investigations has related to the outcomes of one specific
type - with little opportunity to address the question of whether schools were performing well
across the range of outcomes we have identified. It is possible to go further than this by going to
a stage of analysis which reflects something of the underlying statistical relationships between
pupils' backgrounds, school membership and their performance in GCSE examinations. To that
we now turn.

A statistical account of GCSE outcomes

Around five thousand pupils in more than thirty schools, each with an average Year 11 cohort
size of over 150, provides a good opportunity to utilise the statistical modelling framework
provided by 'multi-level modelling'. This is a refinement of 'ordinary least-squares regression'
which has been developed by Goldstein, and others, at the University of London's Institute of
Education. This method of analysing the 'hierarchical' data of pupils clustered by school has
been the focus of a number of investigations in the school effectiveness field (as well as in many
others)\textsuperscript{vi}. In effect, what these procedures do is to estimate a 'predicted' GCSE outcome score for
each pupil conditional not only on their own personal characteristics, but on the outcomes of
similar pupils from all the other schools. Additionally the method allows an estimate to be made
of the degree to which pupils' membership of a particular school 'adds to' or 'subtracts from'
the overall expectation of achievement for the generality of pupils.

The results from such an analysis provide a number of useful outputs which we will utilise in our
account of performance in the schools of our sample:

First, it provides a 'predictive' model linking pupils' background and Key Stage 3 characteristics
to the 'expected' outcome in each of the four outcomes chosen;

Second, it generates school 'residuals', and an estimate of the range of uncertainty in this, from
which an assessment of whether the school appears to be producing 'better' or 'worse' results than
others can be made;

Third, it allows investigation of the possibility that different groupings within schools perform
differentially; for example; girls vs boys, or 'high' vs 'low' prior attainers; and,

Fourth, it provides a measure of how well the 'model's predictions fit the outcome measure
being used. The size of this measure gives some indication of how much the inter-pupil and
inter-school viability is reduced by using the 'explanatory' factors in the model; for example,
using pupils Key Stage 3 'scores', much of the variability in GCSE outcomes can be 'associated'
with this. Similarly for gender and other factors which make a substantive contribution to our
understanding of what 'makes a difference' to pupils' outcome scores\textsuperscript{xvi}. 
Reporting the results of the analysis

The background factors which were found to make a significant difference in pupils' GCSE outcomes were:

- Gender;
- Key Stage 3 aggregated 'level'.

(N.B., pupils' date of birth, and specifically, whether they were 'summer-born', did not feature in this account. There was clear evidence that 'summer-born' pupils did less well in their Key Stage 3 scores than others, so a 'non-significant role for the 'summer-birth' factor here really means that these pupils suffered no 'additional' disadvantage between Years 9 and 11, rather than that no such disadvantage exists as a factor in their experience of school.)

In Table 4 below we show how to calculate 'estimated' performance for pupils with given gender and Key Stage 3 score. The results are quoted for a girl of 'average' Key Stage 3 score. To calculate the appropriate measure for different pupils proceed as indicated in the section following Table 4.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Predicted Female with 'average' KS3 score</th>
<th>Outcome if Male</th>
<th>Score if KS3 level = 5.0</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Points</td>
<td>41.5</td>
<td>-2.6</td>
<td>13.8</td>
<td>69%</td>
</tr>
<tr>
<td>Total Points/Entry</td>
<td>4.50</td>
<td>-0.22</td>
<td>1.23</td>
<td>74%</td>
</tr>
<tr>
<td>A* - C Passes</td>
<td>5.20</td>
<td>-0.6</td>
<td>2.91</td>
<td>68%</td>
</tr>
<tr>
<td>A* - C/Entry</td>
<td>0.55</td>
<td>-0.1</td>
<td>0.29</td>
<td>68%</td>
</tr>
</tbody>
</table>

To calculate the predicted total points score outcome for a male with Key Stage 3 score of 5.5, we proceed as follows:

Predicted total points = 41.5 - (2.6) + (0.5 x 13.8) = 45.8

Table 4 shows how strongly the factors of pupils' Key Stage 3 starting points and their gender, are associated with each outcome. 74 percent of the variability in pupils' average points score per GCSE entry is 'explained' by the two factors of Key Stage 3 score and gender - this is particularly noteworthy - and it implies a very strong relationship indeed between 'intake' and 'outcome'. The other values for R² are also, historically, amongst the highest recorded in this field. Figures 8a to 8d show the actual and predicted outcome scores achieved by pupils in each of the four outcome measures we have used.

How schools are evaluated by these models is, of course, an important issue - for we have seen that the different outcomes are not identical at school level. We have taken, therefore, the somewhat restrictive decision to report only those schools where there was evidence, across each of the outcomes, of 'better' or 'worse' performance. Schools themselves will be provided with the details of how well, or otherwise, they performed on each of the outcomes. For our
Pupils' GCSE Outcomes in relation to Key Stage 3 Average Levels

Predicted GCSE Points superimposed; Girls (Upper line), Boys (Lower).

Figure 8a
Pupils' GCSE Outcomes in relation to Key Stage 3 Average Levels
Predicted Average GCSE Points superimposed; Girls (Upper line), Boys (Lower).

Figure 8b
Pupils' GCSE Outcomes in relation to Key Stage 3 Average Levels
Predicted Number of A* to C Grades achieved superimposed (Girls, Upper line)
Pupils' GCSE Outcomes in relation to Key Stage 3 Average Levels

Predicted Number of A* to C Grades achieved per Entry (Girls, Upper line)

Figure 8d
purposes there is considerable value in restricting our attention to that minority of schools where there appeared to be consistent evidence of 'added' or 'subtracted' value.

Table 5th below shows six schools which were consistently evaluated as providing significantly 'better' or 'worse' performance across each of the four outcomes which we have analysed. It also shows for each outcome, the degree to which each school's performance was 'better' or 'worse' than expected.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Residual School</th>
<th>Total Points</th>
<th>Average Points</th>
<th>Number A*-C</th>
<th>Avge A*-C</th>
<th>Entry Average</th>
<th>Entry Spread</th>
<th>Average KS3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>+ 6.3</td>
<td>+ 0.3</td>
<td>+ 1.3</td>
<td>+ 0.1</td>
<td>9.7</td>
<td>1.2</td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>+ 3.8</td>
<td>+ 0.4</td>
<td>+ 0.8</td>
<td>+ 0.1</td>
<td>8.4</td>
<td>1.6</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>+ 6.3</td>
<td>+ 0.2</td>
<td>+ 1.3</td>
<td>+ 0.1</td>
<td>10.2</td>
<td>1.9</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>+ 4.4</td>
<td>+ 0.4</td>
<td>+ 1.2</td>
<td>+ 0.1</td>
<td>9.1</td>
<td>1.4</td>
<td>5.2</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>- 5.0</td>
<td>- 0.4</td>
<td>- 1.3</td>
<td>- 0.1</td>
<td>8.4</td>
<td>1.3</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>- 8.1</td>
<td>- 0.2</td>
<td>- 1.6</td>
<td>- 0.1</td>
<td>7.4</td>
<td>0.9</td>
<td>4.9</td>
<td></td>
</tr>
</tbody>
</table>

It is interesting to note that the schools which feature in Table 5 have Key Stage 3 scores fairly close to the average for the Project as a whole. We have seen, in Figure 2, that schools comprise very different pupil compositions in terms both of the average and of the spread of their Key Stage 3 scores. There appears to be no evidence from our analysis that neither schools which serve a narrow range of pupils, nor those which have relatively high Key Stage 3 starting points on average, do particularly well in the outcomes reported here. Outcomes for selective schools were 'as expected', given the characteristics of their pupils, and the performances of pupils similar to theirs in other schools. Figure 2 has indicated that many of the Project schools have pupils with Key Stage 3 scores which fall well within the range of those in the selective schools. These pupils, in the non-selective sector appear to have done 'as well as' or in some cases, as indicated in Table 5 above, 'better than' their counterparts in the selective schools.

This is not necessarily a general finding; our sample was not drawn to address this issue - but, in the spirit of looking towards 'indicators' of added value, it would seem that selective schools have no advantage over others in this respect.

At the extremes, the schools in Table 5 provided very contrasting educational outcomes for the pupils they comprised. The methods we have used imply that similar pupils educated in schools B and E would achieve, on average:

- almost 9 GCSE points difference in total points score;
- just under a grade difference for each examination entered;
- more than 2 additional A* to C grades in total; and
- 0.2 of an A* to C grade per entry.

These represent very substantial and important outcome differences for pupils attending these schools. Some schools appear to be systematically 'advantaging' their pupils, whilst others clearly do the opposite. What lies behind these contrasting sets of outcomes? Is there scope, for
example, for School E to 'learn from' the situation in School B in order to improve its own performance?

What is clear is that, without access to data of the kind made available for this Report, we would not now have this detailed comparison between schools, and hence there would be no catalyst for finding out how to 'improve' the relatively weak set of outcomes in School E.

Clearly there is a case for information of this type to be made available to each school so that each can evaluate the extent to which its performance compares to that which similar pupils achieve in a range of other schools.

Obviously, the matter should not stop there - because, if there is evidence of 'underperformance', schools need to investigate further to discover whether there are particular areas of 'weakness' - and even then, they will need to have access to 'exemplars' of practice where things have turned out rather 'better' and from which they might begin the process of 'turning around' their situation.

Comparing Schools' performance in 'core' subject areas

Using the same methods we have explored above in 'modelling' pupils' global outcomes across the range of subjects which they studied, we may also look at the four subject areas of English, Mathematics, Science and Modern Language.

A complication arises in the cases of Science and Modern Language, that pupils may obtain these qualifications by different routes. For example, the great majority of pupils take Double Science, but an important minority take a combination of the separate sciences whilst others take 'Single' science. In Modern Language - the majority of pupils who take any foreign language do so in French, but a range of others are available in different schools. Some pupils take a second modern language - which may or may not include French.

We have taken a simplifying decision in what follows: we have compared pupils' examination performance where there are reasonable grounds to believe that the performances are comparable. Thus, we have simply compared pupils' outcomes in 'Science' (pupils' best results in any forms of science has been included).

However, this option is not really open to us in Modern Languages, so we have taken our exemplar of a modern language to be French, and have simply compared modern language performance through the vehicle of this subject. Other decisions were of course possible, such as, for example, taking the 'best' grade in whatever language a pupil achieved. This may be a more appropriate way to proceed to ensure comprehensive comparative coverage of schools' outcomes - but for the purposes of this Report we have decided to proceed on the more simple set of assumptions indicated above.
Reporting the outcomes of 'subject' performance

We used each individual pupil's grade in the subject concerned as the outcome to be 'modelled'. And proceeded in exactly the same way as with the global measures of outcome reported above. The 'predicted' outcome scores were as indicated in Table 6 below:

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Predicted Female with 'average' KS3 score</th>
<th>Outcome if Male</th>
<th>Score if KS3 level = 5.0</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>4.8</td>
<td>-0.5</td>
<td>+1.2</td>
<td>62%</td>
</tr>
<tr>
<td>Mathematics</td>
<td>4.0</td>
<td>+0.3</td>
<td>+1.5</td>
<td>69%</td>
</tr>
<tr>
<td>Double Science</td>
<td>4.4</td>
<td>+0.2</td>
<td>+1.3</td>
<td>64%</td>
</tr>
<tr>
<td>French</td>
<td>4.3</td>
<td>-0.6</td>
<td>+1.3</td>
<td>54%</td>
</tr>
</tbody>
</table>

It is possible to calculate the 'expected' grade for any pupil in these subjects using the method outlined in the 'box' following Table 4. Schools were not uniformly effective across subject departments, the correlations between schools' subject outcomes suggested considerable variation within schools in their effectiveness in these subject areas. (See Endnote xi) Evaluating 'subject' performance within schools should form an important part of each school's review procedures. The Appendix provides a visual representation of the actual and 'modelled' outcomes in each subject area. (See Figures G, H, I and J). How well a school does in English can be seen by plotting its pupils results on Fig G.

Other issues

The most obvious distractions, viewed externally, to be drawn between schools relate, first, to the composition of their pupil body and second, to the examination entry regime in place.

In respect of the first we have found no evidence that schools which are 'selective' in intake produce 'better' outcomes than other schools. Although this was not an issue our sample of schools was chosen to address, the limited conclusion to be drawn is that, as far Project schools are concerned, 'selective status' did not feature as an indicator of better performance.

Entry policy is, however, a rather different matter, for two very different reasons. Figure 9 shows that many schools gave pupils the choice of around nine examination entries, but that some, for example, schools A and C in Figure 9, offered significantly different levels. Where lower numbers of GCSE entries occurred, these were often supplemented by GNVQ courses which have considerable provenance for the future qualification framework within which it is appropriate to evaluate schools' performance. It is unfortunate that it was not until quite late in the day that outline agreement was reached to incorporate achievements at GNVQ into the 16+ framework with 'equivalents' to GCSE passes. There is still a dearth of evidence relating GNVQ qualifications to Key Stage 3, or other, starting points, and there is also a lack of differentiation in the outcomes of these courses in any way comparable to the A* to G (and fail) grades awarded in GCSE. This is a vital area for further study, and in particular, in the context of this Report, that such qualifications should be recognised in any account of schools' performance.
Examination Entry Profile for 4600+ Pupils
Comparative Profiles for three Project Schools

Number of GCSE Examination Entries

Cumulative Percentage

A
B
C

Figure 9
The second issue relating to entry policy is rather more subtle. Some schools entered pupils for relatively large numbers of GCSE examinations - and we have sought to take account of this in our definition of 'quantity' and 'quality' of outcomes. However, whilst there is an obvious distinction between schools in the average number of GCSE entries their pupils make, there is another, rather less obvious difference - and that is the way in which pupils from the same school may have very different 'entry profiles'. We have attempted to capture this numerically by calculating the within school 'spread' of examination entries. In Figure 9 schools A and C have a similar 'spread' of examination entries, whilst school B provides a very much wider range for its pupils.

The spread of entries was, in some schools, relatively small - implying that little differentiation was evident between pupils' entry rates. However, in others, there appeared to be much greater 'selectivity' of examination entry rates for different pupils, and not always because they differed in Key Stage 3 starting points.

When we attempted to model this as a 'school-factor' it did not feature as a significant factor but, there were worrying indicators that where there was greater differentiation between pupils' entry rates in a given school, there appeared to be a distinct negative impact on outcomes.

If it were possible to take this further, it might be found that by choosing some pupils for whom a high examination profile was thought appropriate, and by contrast others were given rather lower targets, schools inadvertently contributed to a culture of lower expectations on the part of some. And, in the worst cases, even possibly aided a 'counter-academic' culture amongst a minority of pupils. Such an outcome could well then militate against the school being able to maximise the examination performance of its pupil body, and might by contrast, through peer pressure and other means, be responsible for lowering the achievements of all.

This is a speculative account of one feature which has been surprisingly on the edge of statistical significance in a number of recent investigations. It is suggested that this matter also could form a useful additional element of any further work undertaken in this value-added field.

Conclusions

This study has shown that it is possible to represent the comparative progress of pupils in different schools over the two final years of compulsory schooling - and that the National Curriculum assessments at Key Stage 3 can play an important part in this. Schools clearly differ in the effectiveness of their provision - and we have sought to provide two methods by which such information can be put to constructive use.

First, by the use of diagnostic 'quartile' plots we have enabled schools to obtain a visual account of how well the performance of their pupils compares with that of many others in other schools - and we have identified the ways that this can be done both for 'global' outcomes as well as for individual subject departments. These methods have a clear value and importance in the area of school self evaluation. No doubt, it is necessary for schools to learn to obtain quick and effective access to material of this kind - and a publication along the lines of the Department's GCSE to A/AS level might be helpful. However, it is likely that the evidence which we have collected here is capable of expansion to include a larger number of pupils and schools. It should also be updated regularly because if schools are being challenged now to improve their
performance year on year, we need evidence about current performance to make judgements, rather than relying on patterns which may have been appropriate in 1996, but might be less so in 1997 or 1998, for example.

This first approach, which has as its audience schools' own self evaluation, has been supplemented by a statistical account which allows for much greater precision. The outcomes of the second approach are in one sense more limited, but nonetheless important for that. The second approach embraces a more directly accountable framework - and one which has identified clear and consistent examples of both 'good' and 'bad' outcomes. The frameworks within which these evaluations are made are obviously external to the school, and hence provide some support for external investigation and audit. We have raised the question about 'why' such large and systematic differences as we have observed could have arisen between the four schools on the one hand and the two on the other. If we are serious about seeking to improve performance overall, then such indicators must be taken seriously.

The benefit of this might, at first, appear to be mainly negative - in that schools E and F really are not providing the outcomes which their pupils', their parents and school governors should expect. This situation, once known about, must be a priority for review and action, with further subsequent evaluation.

However, the positive side to all this is that schools may now have a number of exemplars of 'better' outcomes - in ways that challenge them not to rest with the, possibly, 'satisfactory' levels of achievement they may already have. This should stimulate them to continue the process of continual improvement to benefit not only their pupils, but ultimately their community and society as a whole. The desirable outcomes to which all this evaluative activity is directed is to help schools provide society with more better qualified young people - and who, themselves, respond positively to the challenge of building further on the success to which their school careers have so far led them.

Action points to ensure schools' access to comparative performance evaluation information

Schools in the Project sample voluntarily made their data accessible for the purposes which have been developed above; at present there are no arrangements in place whereby such analyses could be carried out at a local level.

There are at least two reasons for this:

- Schools are deemed to 'own' any data emerging from National Curriculum and GCSE and GCE assessments, there is no right of access to this by 'their' LEA or indeed any other body than the DfEE, and then only in aggregate form.

- LEAs do not have a duty to provide comparative performance information for their schools, and indeed, such has been the marginalisation of LEA activity (at least until recent date) that relatively few LEAs have taken the steps necessary to ensure that they can provide schools with appropriate and relevant performance comparisons.
Two changes in current administrative procedures could ensure that maximum use was now made of the kinds of performance data which we have shown to be essential in evaluating school performance:

- a right granted to LEAs to access the pupil-level assessment data on all pupils in 'their' schools (and in an agreed format); and,
- a duty placed on LEAs to utilise such information, on schools' behalf, in ways which help to inform each school's self-evaluation and 'improvement' agenda; helping also to challenge assumptions about 'low' expectations or performance.

The Year II rolls of the schools concerned comprised 5182 pupils; Key Stage 3 results were available for some 4921 pupils. The 'matching' process provided complete records for 4622 pupils; around ninety percent of those schools' Year II rolls.

There was considerable variation between schools in the proportions of pupils for whom a matched record could be created. In some schools this was almost 100 percent, in most it was above ninety percent but in a few schools there were very considerable shortfalls.

This is an issue which should be studied further, since if schools do 'recruit' or 'lose' substantial numbers of pupils between Key Stage 3 and GCSE, the matched record represents only a part of the impact of that school on its pupils. In extreme cases this could lead to a complete mis-specification of the situation in those schools.

The mean and standard deviation of the three Key Stage Subject 'levels' were very similar as indicated in the Table below:

<table>
<thead>
<tr>
<th>Subject</th>
<th>Average 'Level'</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>5.1</td>
<td>1.1</td>
</tr>
<tr>
<td>Mathematics</td>
<td>5.0</td>
<td>1.2</td>
</tr>
<tr>
<td>Science</td>
<td>5.1</td>
<td>1.1</td>
</tr>
</tbody>
</table>

In 1994 there were seven possible Key Stage 3 sets of papers which pupils could take: in English a single paper covered all possible levels available; in Science there were two 'tiers' covering Levels 3 to 6 and 4 to 7; whereas in Mathematics there were four tiers, covering Levels 3 to 5; levels 4 to 6; levels 5 to 7 and levels 6 to 8.

In each set of papers SCAA published a conversion table giving the 'level equivalents' of scores on each tier of the papers in each subject. We have modified this to provide an interpolated level which allows for finer differentiation of pupils' Key Stage 3 attainments. Thus we have available a set of 'finely differentiated' levels for each pupil in each of the three subject areas. 'Conversion graphs may be constructed to show how a given score on a given tier test allows interpolation of a finely differentiated level. We have assumed that, although integer levels allocated to test scores are constant across a range of scores, an underlying continuum exists which differentiates between pupils' performances at each 'end' of the range of scores allocated to the given level.

Figure 1 indicates that there is an 'upper-bound' on Key Stage 3 levels. Thus the Key Stage 3 aggregated level scores are only approximately normally distributed.

This method was helpfully displayed as visual examples in the DfEE's publication 'GCSE to GCE A/AS Value-Added Briefing for Schools and Colleges' (DfEE, 1995). It was also implicit in the Briefing Paper 'Value Added in Education' (DfEE, 1995).

It is possible to approach the 'modelling' of schools' outcomes in less sophisticated statistical ways, but the
crucial question which needs to be addressed is how the method used handles the uncertainties in the estimates of performance which it makes. The advantage of the multilevel modelling approach utilised here is that it provides clear, if conservative, estimates of this. Additionally, in the interests of furthering important research questions which may have an influence on future policy matters, the multilevel model approach allows the possibility of exploring whether there are particular 'group' effects within certain schools, and whether the general 'predictive' models need to be modified in particular schools to take account of 'unusual' features of performance.

It is perfectly possible for schools to differ, for example, in their effects on 'low' as opposed to 'high' prior attaining groups - and even though this may be a rare phenomenon, its existence should at least be recognised. For schools' use of general value added ideas we shall adopt an approach pioneered by DfEE's Analytical Services Branch in providing a framework for evaluating 'value-addedness' in the progress which schools and colleges provide their students with from GCSE to A/AS Level (DfEE; 1995).

The diagrams showing outcomes plotted against Key Stage 3 Level scores indicate that there is a relatively high correlation between them. This is even more pronounced for the 'quality' measures, and should be compared with more 'traditional' situations in which the prior attainment score was supplied on entry to secondary school rather than only two years before GCSE examinations as here. It has become a 'typical' expectation that around 50 percent of the variability in GCSE outcomes can be associated with differences in prior attainment scores. Using Key Stage 3 scores the measure has turned out to be closer to, and in some instances, greater than, 70 percent.

The form of equation specifying the relationship between GCSE score and Key Stage 3 aggregated level was chosen to be linear, although slightly higher correlations were recorded using, in addition, a squared and cubic form of the Key Stage 3 measure. A linear measure was preferred, in this Report, because of the greater simplicity conferred on the calculations involved in deriving Table 4. There were no substantive differences between the linear and other models' evaluations of schools' performance reported in Table 5.

The models were \( y_{ij} = (\beta_0 + u_{0i}) + (\beta_1 + u_{1i})x_{ij} + \beta_2z_{ij} + e_{ij} \)

\( (y_{ij} \) is the outcome variable, \( x_{ij} \) is the Key Stage 3 score for each pupil, and \( z_{ij} = 0 \) if pupil is female, \( z_{ij} = 1 \) if male) and where \( u_{0i}, u_{1i} \) are random variables with parameters:

\[
\begin{align*}
E(u_{0i}) &= E(u_{1i}) = 0 \\
\text{var}(u_{0i}) &= \sigma^2_u \quad \text{var}(u_{1i}) = \sigma^2_v \\
\text{cov}(u_{0i}, u_{1i}) &= \sigma_{uv}
\end{align*}
\]

In effect these models estimate an intercept and 'slope' (on the Key Stage 3 score) for each school, with fixed effects for gender and Key Stage 3 score.

Questions of 'differences in slope' - that is, investigations as to whether schools 'affected' their less able pupils differently from their more able peers - were pursued but there were no instances where schools were found to be 'differentially effective' across all four outcome measures. We have therefore not reported these estimates.

There was, however, evidence of 'differential effectiveness' for each of the four outcomes we have utilised. The numbers of schools so identified was much greater for the 'quantity' than the 'quality' measures. In the latter case the variance estimates indicated results on the margin of significance.
We present below the variance table for each of the four outcome measures:

<table>
<thead>
<tr>
<th>Outcome</th>
<th>GCSE Pts</th>
<th>Avge Pts per Entry</th>
<th>No of A* to C Passes</th>
<th>Avg A* to C per Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>25.9 (7.0)</td>
<td>0.044 (0.013)</td>
<td>0.714 (0.20)</td>
<td>0.0035 (0.001)</td>
</tr>
<tr>
<td>Covariance</td>
<td>4.6 (2.0)</td>
<td>-0.002 (0.005)</td>
<td>0.196 (0.07)</td>
<td>-0.00003 (0.03)</td>
</tr>
<tr>
<td>Slope on KS3</td>
<td>(3.1) (1.0)</td>
<td>0.008 (0.004)</td>
<td>0.124 (0.042)</td>
<td>0.0006 (0.0003)</td>
</tr>
<tr>
<td>Pupil Level</td>
<td>73.3 (1.6)</td>
<td>0.581 (0.013)</td>
<td>4.02 (0.087)</td>
<td>0.043 (0.0009)</td>
</tr>
</tbody>
</table>

The correlation between schools' estimates of performance as evaluated under the four performance measures we have used is shown in the Table below:

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Total Points</th>
<th>Average Points</th>
<th>No of A*-C Passes</th>
<th>Average A*-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Pts</td>
<td>1.00</td>
<td>0.86</td>
<td>0.98</td>
<td>0.88</td>
</tr>
<tr>
<td>Avge Pts</td>
<td>1.00</td>
<td>0.93</td>
<td></td>
<td>0.99</td>
</tr>
<tr>
<td>A*-C</td>
<td>1.00</td>
<td></td>
<td></td>
<td>0.95</td>
</tr>
<tr>
<td>Avge A*-C</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

This table indicates differing 'views' of schools' performance arising from the use of different performance measures. If nothing else, it indicates the importance of utilising frameworks for evaluation which comprehend significant parts of schools' contribution to pupils' examination results - and that excellence in all aspects should be a pre-requisite of any indicator of outstanding school performance. By the same token, for accountability purposes, a fourfold 'negative' evaluation of performance has much greater indicative strength than one based solely on one aspect of schools' outcomes.

In Table 5 we have shown only those schools where the 'residual estimates' were significantly different from zero in each of the four outcome measures. It would have been possible to quote the standard errors of each residual, however, given the principle on which they were chosen and avoiding additional complexity in the Tables this course was not pursued. We emphasise that each residual for each school was significantly different from zero.

The correlations between subject performances in the four 'core' subjects within schools are shown below. It should be noted that 'French' results were not available for two schools.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>English</th>
<th>Mathematics</th>
<th>Science</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>1.00</td>
<td>0.23</td>
<td>0.30</td>
<td>0.35</td>
</tr>
<tr>
<td>Mathematics</td>
<td></td>
<td>1.00</td>
<td>0.10</td>
<td>0.06</td>
</tr>
<tr>
<td>Science</td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.12</td>
</tr>
<tr>
<td>French</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

The small size of these correlations between different subjects suggests that departments may differ substantially in their effectiveness within schools. It would appear, therefore, that measures of a schools' overall effectiveness, as shown by the indicators we have utilised above, may not be wholly informative about pupils' outcomes, and that a 'departmental focus' is necessary in addition. It is possible, although the analysis has not yet been carried out, that it may be the 'mix' of 'effective' and 'less effective' departments which gives a school its distinctiveness, and that 'overall effectiveness' is simply the aggregated effect of these 'departmental effects'. Could it in some cases that one very well performing department compensates for rather mediocre performance in most others?
Considerations such as these should prompt schools to take very seriously indeed the evaluation of subject department's performance. Availability of data in appropriate format alongside diagrams such as those shown in Appendix Figures G, H, I and J should help schools pursue these goals with insight and sharper focus than has conventionally been available to them by more 'traditional' methods.
School A's GCSE Performance
Comparison with 'Average' Outcomes in all Schools

Appendix Figure A
School B's GCSE Performance
Comparison with 'Average' Outcomes in all Schools

Appendix Figure B
School C's GCSE Performance
Comparison with 'Average' Outcomes in all Schools

Appendix Figure C
Range of GCSE Scores in 'Core' Subjects for pupils with similar KS3 Levels

English

Appendix Figure D
Range of GCSE Scores in 'Core' Subjects for pupils with similar KS3 Levels

Mathematics

Key Stage Level 3 Group

Grade in Subject

Lower Quartile

Upper Quartile

Appendix Figure E
Range of GCSE Scores in 'Core' Subjects for pupils with similar KS3 Levels
Science

![Graph showing the range of GCSE scores in Science for pupils with similar KS3 levels. The graph includes upper and lower quartile lines, with Key Stage Level 3 Group on the x-axis and Grade in Subject on the y-axis. Appendix Figure F]
Subject Outcomes at GCSE in relation to Key Stage 3

English

Appendix Figure G
Subject Outcomes at GCSE in relation to Key Stage 3
Mathematics

Pupils' Key Stage 3 Average Level

Appendix Figure H
Subject Outcomes at GCSE in relation to Key Stage 3 Science (Best Grade in Double/Single/Separate Science)

Appendix Figure 1
Subject Outcomes at GCSE in relation to Key Stage 3
French

Appendix Figure J