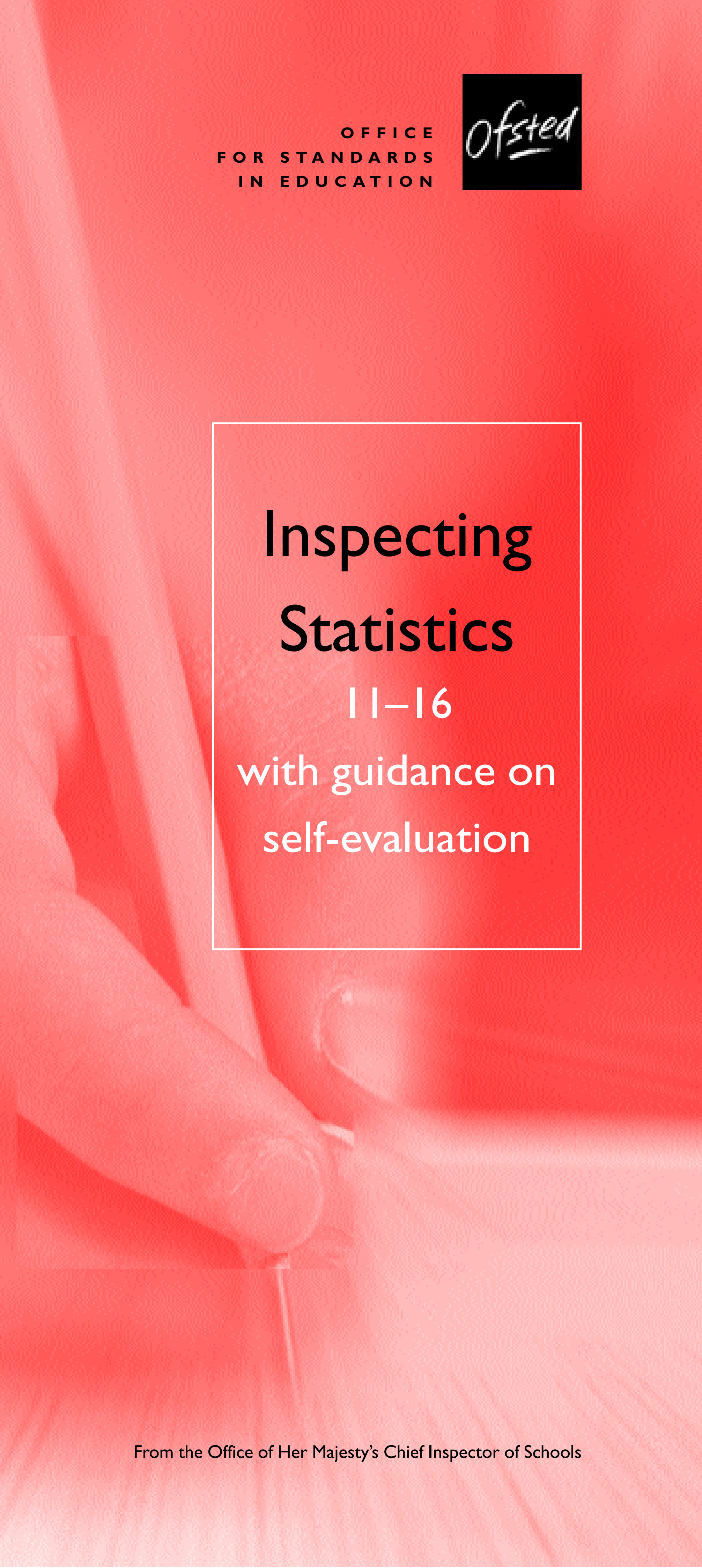
OFFICE
FOR STANDARDS
IN EDUCATION

Ofsted



Inspecting Statistics 11–16

with guidance on self-evaluation

From the Office of Her Majesty's Chief Inspector of Schools

Inspecting Statistics

11–16

with guidance on self-evaluation

Office for Standards in Education
Alexandra House
33 Kingsway
London
WC2B 6SE

Telephone 020 7421 6800

Web site: www.ofsted.gov.uk

Reference no. HMI 758

CONTENTS

	INTRODUCTION	5
	Context of statistics to age 16	5
	OFSTED inspection of statistics	5
	COMMON REQUIREMENTS	7
1	STANDARDS AND ACHIEVEMENT	9
1.1	Performance data	9
1.2	Analysis of pupils' work	11
1.3	Talking with pupils	12
1.4	Evidence from lessons	14
2	TEACHING AND LEARNING	15
2.1	What to look for in lessons	15
2.2	Observing lessons	16
2.3	Other evidence on teaching and learning	19
3	OTHER FACTORS AFFECTING QUALITY	21
4	SUMMARY	23

INTRODUCTION

This booklet is intended to help school inspectors, headteachers and staff to evaluate education in statistics for pupils aged 11–16. It complements OFSTED's inspection handbooks and is part of a suite of guidance booklets for inspecting subjects and aspects 11–16. The suite includes a booklet on mathematics.

The booklet gives guidance on evaluating standards and achievement, teaching and learning, and other factors that have a bearing on what is achieved. You will find advice on how to interpret performance data; what to look for in pupils' work and the questions to ask them about it; the subject-specific points to observe when in lessons; and how to bring your evaluations together to form a coherent view of the subject.

Examples of evidence and evaluations are *italicised* and presented in boxes, sometimes with a commentary to give further explanation. These examples show a range of ways in which evidence and findings can be recorded and reported. They are not meant to endorse any particular method or approach.

School inspectors on short inspections and senior staff in schools are likely to evaluate a range of subjects and may need to refer to more than one booklet. To facilitate this, all the subject guidance booklets can be downloaded from OFSTED's web site (*www.ofsted.gov.uk*).

Our School Inspection Helpline team, on 020 7421 6680, will respond to any queries (or you can e-mail *schoolinspection@ofsted.gov.uk*).

Context of statistics to age 16

Statistics is part of the National Curriculum for mathematics in Key Stages 3 and 4. However, this booklet is concerned entirely with statistics as a General Certificate of Secondary Education (GCSE) subject. There are just over 20,000 entries for GCSE in statistics each year, of which about 17,500 are in Year 11 in maintained schools. They include about 9,300 boys and 8,200 girls, with approximately 16,000 in comprehensive schools, 800 in selective schools and 750 in modern schools. From 2000 to 2001, the number of Year 11 maintained school entries rose by 3,500 and the number is expected to continue to increase with the growth of mathematics 'specialist schools'. Provision includes, for example, several classes with a full timetable allocation (2½ hours a week), double entry for mathematics GCSE and statistics (in a similar way to English language and literature, and often for more able pupils), and various forms of curtailed timetable allocation.

The general advice in this booklet is illustrated using a case study of the hypothetical Marion Richardson School. The school has one large mixed-ability group taking GCSE statistics in each of Years 10 and 11. All pupils take at least a short GCSE course in religious education (RE), with an allocation of two 50-minute lessons. Some opt to make this up to a full GCSE in RE (a further timetabled 50-minute lesson). Others use the 50 minutes in the timetable, together with a lesson formally timetabled at 3.30 pm to prepare for GCSE in statistics or one of several other subjects such as dance or Spanish. This timetabling provides 100 minutes a week rather than the normal 150 minutes. It is an arrangement that extends the range of the Key Stage 4 curriculum offered by increasing the length of some timetabled days; staff involved have timetable remission at other times in the week.

OFSTED inspection of statistics

Lessons formally timetabled before or after 'normal' hours are, for OFSTED purposes, just as much part of the formal curriculum as any other. When a subject is taken by only one or two groups in each of Years 10 and 11, lessons will not occur many times during an OFSTED inspection. So inspectors should make every effort to avail themselves of most of this limited available evidence. Where statistics has a GCSE entry of at least 20, it will feature in table S5 in the

Performance and Assessment (PANDA) report; its evaluation then merits a discrete section in Part D of the published report. Where it is a less significant part of the school's work, it is sufficient to add a paragraph to the section for mathematics.

COMMON REQUIREMENTS

Inspectors or evaluators in schools should have a good understanding of the key characteristics of the school and its pupils. The achievement of individuals and the different groups of pupils in the school must be evaluated to judge how effectively their needs are met. Credit should be given where, against the odds, pupils achieve well even though they may not have reached the levels of attainment expected for their age.

As an evaluator, you must be thoroughly familiar with the specific requirements for statistics in the National Curriculum. You will need to consider how successfully the subject contributes to pupils' spiritual, moral, social and cultural development, and how effectively it helps to prepare pupils for adult life in a culturally and ethnically diverse society.

When evaluating statistics, you should consider how well planning and teaching take account of the following principles of inclusion:

- setting suitable learning challenges;
- responding to pupils' diverse learning needs;
- overcoming potential barriers to learning and assessment for individuals and groups of pupils;
- promoting racial equality.

You need to be informed about the responsibilities and duties of schools regarding equal opportunities, particularly in respect of discrimination on grounds of gender, race and disability. These are covered by the Sex Discrimination Act 1975, the Race Relations Act 1976, the Race Relations (Amendment) Act 2000, and the Special Educational Needs and Disability Act 2001, and their respective Codes of Practice. These Acts underpin national policies on inclusion, on raising achievement and on the important role schools have in fostering better personal, community and race relations, and in addressing and preventing racism.

In many schools you will find additional resources and initiatives aimed at promoting educational inclusion. You must know about any nationally funded or local initiatives in which the school is involved so that you can assess their effectiveness. [1] There is guidance on this in the OFSTED publication *Inspecting New Developments in the Secondary Curriculum 11–16 with guidance on self-evaluation* (published 2001).

[1] Notably, Education Action Zones, Excellence in Cities, Ethnic Minority Achievement Grant (EMAG), and other programmes funded through the Standards Fund, the Single Regeneration Budget, the New Opportunities Fund and New Deal for Communities.

1 STANDARDS AND ACHIEVEMENT

If the subject was part of the curriculum when the school was last inspected, read the previous inspection report carefully to gather information about standards and achievement at that time. This provides the reference against which you can gauge improvement.

Interpret what past performance data say about standards in statistics. Identify the characteristic strengths and weaknesses in pupils' current work from direct observation. Judge the standards of 16 year olds using the same scale as for test and examination results: 'above average', 'average', 'below average', and so on. In the interests of clarity, you can report standards using terms such as 'high', 'average' or 'low'. You should explain convincingly any differences between past results and the standards you observe.

Judge pupils' achievement by considering whether standards in statistics are high enough, taking into account the pupils' starting points or capabilities, the progress they make over time and the demands made on them. Look for and pursue significant differences in achievement between groups of pupils, such as boys and girls, pupils of different ethnic heritage, those learning English as an additional language (EAL), Traveller children, those with special educational needs (SEN), and gifted and talented pupils.

1.1 Performance data

Each year, schools receive a PANDA report. Use the data and analyses it presents to gain an overall picture of the standards and achievement of pupils entered for GCSE examinations in statistics, and note areas for further exploration. Be cautious about reading too much into the data when the number of pupils involved is small.

Add to this picture by taking account of any other relevant data or analyses the school can provide, such as:

- GCSE results over time for foundation tier (grades C to G) and higher tier (grades A* to D), but remember that changes from year to year may well reflect differences in the capabilities of the group of pupils taking the subject, particularly where individual entry numbers are small;
- data from monitoring the relative performance of pupils of different ethnic heritage and other groups, but take care not to over-interpret statistics for a small number of individuals;
- value-added analyses, or comparison with the same pupils' attainment in mathematics as a useful indication of achievement.

Table S5 in the PANDA report shows details of the most recent GCSE results where at least 20 candidates have been entered. Consider, for instance, whether the school's distribution of pupils to grades looks significantly different from the national distribution, and look with particular care at the key indicators given in the table.

In interpreting GCSE results, take account of examination entry policy. For example, above average results might have been obtained because the entry was restricted to more able pupils. Always identify how many pupils this involves and investigate the reasons for their non-entry.

Use tables S1 and S3 in the PANDA report to check girls' and boys' results separately, to see if there are significant differences, keeping in mind the national picture. Remember that differences may reflect the differing capabilities of the girls and boys who take the subject. Discuss differences with the department and pursue the issue further in your first-hand observations. In many schools, where statistics GCSE features in table S5, the number of girls and boys entered may be insufficient for inclusion in tables S1 and S3. However, the department should have this information and may have collected it on a three-yearly basis to facilitate more robust interpretation.

Table S5 for Marion Richardson School shows the school's most recent results (S) compared with the national picture (N), when percentages are rounded.

Subject	S/N	% A*-A	%A*-C	%A*-G	Average points score (APS)	Entry as % of year group	%A*-C significant
Statistics	S	13.8	62.0	96.5	4.9	33.3	No
	N	16.6	67.1	96.8	5.0	2.7	
Mathematics	S	24.1	82.7	99.9	5.5	100.0	Yes
	N	9.3	47.1	96.7	4.2	94.8	

(As fewer than 20 boys were entered, table S1 gives no information about statistics. With fewer than 20 girls, table S3 also gives no information about statistics. As with mathematics, national results in statistics are very similar for boys and girls, the girls' results being very slightly higher.)

Table S6 for Marion Richardson School shows the following.

Subject	Subject APS	Average in all other subjects	School difference	National difference	Relative performance indicator	RPI significant
Statistics	4.93	5.78	-0.85	-0.84	-0.01	No
Mathematics	5.51	5.51	0.00	-0.39	0.39	Yes

Other relevant information is that:

- in the PANDA report, the school's overall performance (average total points per pupil) in GCSE/GNVQ (General National Vocational Qualification) was graded as very high (grade A*) against all other schools and well above average (grade A) against similar schools, and girls obtained better results than boys;
- the last inspection report (four years ago) noted that standards in statistics education were 'lower than in other subjects'.

Example ST1: notes on GCSE results

The profile of results looks broadly similar to that found nationally. (However, the average points score for statistics nationally is 5.01 compared with 4.48 for all subjects. Moreover, there is a national difference of -0.84: those entered for statistics have 5.01 + 0.84 = 5.85 average points in their other subjects.)

Overall, pupils do not do as well in statistics as in their other subjects; they do markedly less well than the year group does in mathematics. However, their performance in statistics relative to other subjects is on a par with that found nationally.

The school's average total points per pupil at GCSE compare very favourably with other schools (very high against all schools and well above average for similar schools). The high average total points per pupil are, in part, due to entry for more subjects than is usually possible, facilitated by the extended timetable. However, it is safe to say that the mathematics results, with +0.39 relative performance indicator, represent high achievement. The lower achievement in statistics may nevertheless be satisfactory. (And this would be consistent with the relative performance indicator at near zero.)

The department's three-year data analysis suggests that on the higher-tier paper, statistics results are almost a grade lower than those in mathematics, but that on the foundation tier, statistics results are not much weaker than those for mathematics.

Though girls did better than boys in GCSE overall, in statistics girls' and boys' results averaged over three years are very similar.

The department feels that, despite a small amount of overlap with the GCSE mathematics specification, there is insufficient time for the course. The problem is exacerbated by having a mixed group of foundation-tier and higher-tier pupils.

Hypotheses

Attainment is at least average overall, probably above average.

Pupils do not achieve as well in statistics as in their other subjects, but nonetheless achievement may be at least satisfactory overall; there may be better achievement for the lower attainers than for the higher attainers.

Boys and girls attain equally well, but boys may get more encouragement relative to girls than is the norm elsewhere in the school.

Mixed ability teaching and shortage of time may restrict attainment.

There are some caveats that should be mentioned. The GCSE Year 11 candidature is at present increasing by about 25 per cent from year to year: grade distributions and average points scores are likely to fluctuate from year to year. However, the national difference, like that for German, for example, is generally strongly negative, and there may be curricular reasons for this. Where entry is less than 20, a school difference could be worked out (and hence a relative performance indicator) but care is needed in interpreting such data.

Always discuss with the department the picture of standards and achievement you have formed from the performance data, and take into account their views and any other data they can supply.

Finally, remember that the performance data are important in supporting your evaluation but be sure not to give them undue weight or to report them at length. Report the salient points from the data in a way that reveals their significance, and give prominence to your own findings, to make sense of the total picture of standards and achievement.

1.2 Analysis of pupils' work

The main purpose of this activity is to collect and evaluate evidence of pupils' standards and achievement. Looking at pupils' work gives you a particularly good opportunity to assess whether the pupils' progress is sufficient over the course and whether the demands made on them are consistently high.

You can also collect evidence about pupils' attitudes by considering, for instance, whether their written work shows care and interest. Begin to make links with the quality of the teaching and how well work is marked or more formally assessed.

Analyse samples of work by gender, or focus on other groups within the school, where significant differences in achievement may be at issue, or where 'inclusion' may be precarious, for example, because of poor attendance, high mobility or lack of familiarity with the English language. You might examine the work of a gifted pupil or of a pupil who has not been in the country for long. Where any groups of pupils do not thrive as well as they should, examine whether the teaching and curriculum take sufficient account of their needs.

When evaluating standards of work in statistics (whether during analysis of work, talking to pupils or when observing lessons), important questions to consider include whether pupils:

- demonstrate a secure understanding of key statistical terms (such as 'crude' and 'standardised' rates) and concepts (like the distinction between 'correlation' and 'causality');
- apply concepts and theories to case studies and problems in real situations (for instance in constructing survey samples);
- interpret and evaluate the information they collect from various sources, or whether they accumulate information without using it effectively;
- perform to a progressively higher standard, when activities (such as the piloting of a questionnaire) are repeated during the course;
- understand how different topics (like Venn diagrams and conditional probability) integrate with one another;
- show understanding of the limitations (due for example to invalid measurements or non-representative samples) involved in decision making in statistics;
- know why estimates of probability tend to improve with sample size.

Example ST2: evidence and evaluation from analysing Year 11 pupils' work (including their work from Year 10)

Foundation tier (GCSE grades C–G)

The pupils can reliably construct percentage bar graphs and pie charts and translate from one form to another using information and communication technology and manually (with a calculator). Own data have been encouraged, for example, a Traveller pupil has presented the range of employment in his Traveller community.

Extensive crossing out suggests insecurity in constructing two-way contingency tables (for example, for number of sisters and *number of brothers); much more reliable in interpretation of such tables.*

With simple data, they can deal with median, mode and mean of discrete frequency distribution; however (for example), for number of goals (0–4)/number of games, a common error is to operate on number of goals without considering frequency. Slips in median are often due to failure to generate cumulative frequencies or to list data in order.

Work on median of grouped frequency distribution (in the description for grade C) is mostly very poor; many have incomplete work on this; about half can find the median class; many get the lower class boundary incorrect; some confusion between using $^1/_2n$ (appropriate for estimating the median from a sample of continuous data) and using $^1/_2(n+1)$ (which is used with discrete data); even when correctly computed, some go to the integer above.

About half draw cumulative frequency polygons correctly; the most common error is to plot the mid-points rather than the upper class boundaries.

Pupils seem generally aware of random and stratified sampling and purpose of piloting surveys. EAL pupil treated 'pilot' as synonymous with 'small', but this has been corrected in marking.

Higher tier (GCSE grades A–D)*

Secure on all foundation work, though there are some careless slips in handling grouped frequency distributions. Some correct extension work, for example, comparative pie charts with area proportional to total frequency.

All can use frequency polygons to obtain deciles and are aware of inter-decile range as a means of excluding extreme data, for example, range of male and female salaries in different professions.

All know how the shape of a normal distribution relates to its standard deviation. Most can use a formula to calculate standard deviation (in mathematics GCSE specification). Calculation of standard deviation for a frequency distribution is often hampered by use of the more cumbersome formula, and with grouped distributions the mid-interval value is often incorrect. Unfinished and incorrect work common on computation of standard scores. Complete muddle about variance and standard deviation, and with notion that to halve the variability in an estimate, four times the sample size is required. Marking has done little to clarify pupils' confusion. Gifted pupils have used statistical calculators to convert raw school test marks to T scores (with unit of one-tenth of a standard deviation).

Secure on tree diagrams in interesting contexts for conditional probability, for example, cervical smear test outcomes for those with and without malignant cells.

Work on histograms is secure except for open-ended classes where the usual rule of 'twice the previous class interval' has not been followed. (Seem to take the lowest class to the lowest value possible not below zero, and 'reflect' to the upper end.)

Pascal's triangle for 10 cases appears in the work and there is, for example, a binomial expansion for $(^1/_2+^1/_2)^2$ and $(^1/_2+^1/_2)^3$, but there is no evidence of this leading anywhere.

Overall

Foundation pupils show uniformly strong attainment in relation to the foundation specification for all but the most challenging work. Their books are well presented and work is largely complete: they are achieving well. Higher pupils' work shows very variable attainment. It is strong where topics overlap with mathematics GCSE, but there is confusion on important and more challenging topics unique to the statistics specification. There are suggestions that some such topics have been 'skipped over' in the teaching. Some underachievement is apparent, with attainment not as high as normal for higher tier pupils. Some interesting contexts have been chosen, thus indicating the wide scope of statistics.

[Attainment average (4)]

Commentary

This analysis of work is beginning to explain the picture which emerged from consideration of GCSE data. In the limited time available, foundation material is being covered effectively and these pupils achieve well. Higher work is of very patchy quality. It is particularly weak in the harder topics covered in Year 11. Do teachers have insufficient time in mixed-ability lessons in Year 11 to develop topics properly? We note here that there are strengths and weaknesses in marking. Pupils from different groups, for example, Traveller or gifted pupils, seem to be satisfactorily catered for.

1.3 Talking with pupils

Evidence of standards and achievement also comes from talking with pupils, either individually or collectively, both during lessons and in formally arranged meetings. Such meetings can also provide evidence of pupils' attitudes and personal development, together with indications of the quality of their teaching. When formal meetings are arranged, choose a few pupils to include any ability levels or groups of pupils which you particularly need to investigate. Use your earlier evidence to identify which areas of statistics studies you want to focus upon and which hypotheses you wish to test further.

Your questions need to take account of the ground the pupils have covered by the time you talk to them, but useful lines of enquiry include questions about:

- the difference between discrete and continuous data (give examples);
- the difference between a bar chart and a histogram (explore how they think a bar chart showing 50 people aged 31–40 could be altered to make a histogram);
- the difference between the arithmetic mean and the geometric mean (enquire whether they know how to find the geometric mean when inflation is 3%, 5%, 4%, 4%, 7% over five years, and whether it is the better of the two);
- how they would find the average temperature from a data-logged graph of temperature over 24 hours;
- how they would make an estimate of the number of cars crossing a bridge per minute twice as accurate;
- the interpretation of results in a mathematics test showing median 60 and mean 50, while for the same pupils results in an English test have a median 40 and mean 50.

It is particularly useful to include questions about recent work when the relevant written work has already been analysed. Alternatively, you could look at the written work together with the pupil(s).

Example ST3: notes from talking with a sample of Year 11 pupils

Questions about pupils' work confirm the provisional conclusion reached in the work analysis. Pupils have a fairly firm understanding of topics in the foundation-tier specification, except for foundation pupils on grade C material.

For example, a pupil identified as having moderate learning difficulties (SEN Code of Practice 'school action') can explain a simple aggregative price index as comparison of the price of a shopping basket of commodities with the price of the same at an earlier time. Higher-tier pupils know that, to be useful, such an index must be weighted, so that a loaf of bread and a pair of shoes feature in proportion to relative spending on the goods, not the price of each. These pupils understand the need for chain index numbers to take account of changes over time in the family shopping basket (for example, CDs rather than tapes).

Higher-tier pupils say that they are not clear about some topics on their specification. For example, the binomial distribution has been introduced (Bernoulli trials) with a suggestion that they use it for coursework, but pupils are unclear how to proceed. For example, the evaluator ascertains that they have done Pascal's triangle to n=10. However, pupils cannot suggest how they might sample 10 individuals (classed as male/female) going into a greengrocer's or bookmaker's and compare with null-hypothesis probabilities.

More able pupils can speculate about purposeful and demanding coursework projects. For example, a gifted pupil, as work both for statistics and for science, plans examination of half-a-dozen logs recently cut from different pine trees in a plantation. He aims to measure growth ring width for mean and standard deviation and express ring width as standard scores, thus establishing similar scores for each year irrespective of tree position. Hence he would deduce seasonal variation (rainfall differences summer by summer) and the trend (due to changing pollution) by a three-year moving average method.

Pupils say that statistics supports their work in mathematics, for instance, on probability and histograms. They also say that it has helped with work in other subjects. A talented swimmer has been monitoring his speed in training by semi-averages (dividing a time series into two parts and fitting a trendline to the two averages). A girl planning to study sociology after GCSE can see the need to eliminate extraneous common elements in correlating, on a town-by-town basis, the number of weddings and the number of churches (a common element being the population town by town). Pupils can give many instances of interesting contexts in statistics lessons which have given insight into the wide application of the subject.

[Attainment average (4)]

Commentary

In the main, this gives the impression of attainment which is above that which would be expected from a sample of pupils from an 'average' population. A gifted and an SEN pupil appear to be catered for reasonably, though able students in general are hampered by incomplete syllabus coverage. The course has connected well with other curricular areas and has engaged pupils' interest. Certain higher-tier topics may not have been developed to the point where pupils have mastery of them. The tentative judgement on achievement remains 'at least satisfactory'.

1.4 Evidence from lessons

Grade attainment by considering the difficulty of the work and how well the pupils are succeeding with it, bearing in mind the age of the class. Grade 4 signifies an average or typical standard of work for the age concerned; grades 1–3 reflect high standards that are increasingly above average; grades 5-7 signify low standards that are progressively below average.

Example ST4: Year 11 mixed-ability class of 17 boys and 13 girls; 50-minute lesson in main timetable

Objective: addition and multiplication rules in tree diagrams (sampling without replacement) and, for higher-tier pupils, security of the Lincoln Index for estimating population size

Starting with 10 white and 40 green peas in a bag. All have first tier of branches correct. Only half the pupils have second tier of branches entirely correct: of these, half have treated as sampling with replacement (denominator 50), and others have made careless slips in numerator. Teacher picks up systematic and random errors by individuals and gives general reminder. Half correctly multiply along branches and add between branches; others show variety of errors – taking second-tier probability only, multiplying or adding everything, adding when they should multiply, or adding before multiplying. Of the half who are getting it wrong, the evaluator finds that one-third are making slips but actually understand, while the others are muddled. The task is at least National Curriculum level 8 (for mathematics); about half are operating at this level but there is a 'long tail' of insecurity. On this task, about half the class are operating at GCSE grade B or above. Only one-third of the class are confident with set notation. Higher-tier pupils have a clear picture of the scope of the mark-release-recapture method for estimating the population of birds and insects, for example. Most can see that if the number recaptured is about 20, whether it is 20 or 21 makes a 5% difference to the estimated population. Most also realise that the probability of getting 'correct' population estimates, or various particular incorrect estimates, may be deduced from the tree diagrams. Only half-a-dozen can follow this through with any rigour to propose the number of second captures necessary to give specified probability that the estimate will lie within certain limits of uncertainty. However, as an unstructured task, this would be very challenging even as an Advanced Subsidiary GCE exercise. These pupils are probably operating at about A GCSE on this particular task. A few weaker higher-tier pupils are unsure whether the marked individuals are included in the final population estimate.*

[Attainment above average (3)]

Commentary

Foundation tier requires the addition law for mutually exclusive events and the multiplication law for independent events. Higher tier requires general application of the laws. However, the work is common with GCSE mathematics (and forms part of the description for grade A). There is some provision to suit the work to different abilities (extension to evaluation of the Lincoln Index), but all pupils are engaged on the conditional probability work. Where a topic is contiguous with mathematics GCSE, the pupils are fairly secure. It is not surprising that foundation pupils, who are gaining consolidation on this work, are struggling with the conditional element. However, the overall impression is of a spread of competence concentrated in upper levels (negative skew): this indicates above average attainment. This is not to say that such attainment can be found across a wider range of topics. Where statistics is reinforcing mathematics GCSE, there seems to be high attainment, particularly evident here among higher-tier pupils. On other topics, where coverage is skimped, attainment is notably lower. This suggests variable achievement from one topic to another. In the main, higher-tier pupils seem to underachieve because of time constraints on 'new' work.

General advice

Bring together all your evidence on standards to come to overall judgements. Note important changes since the last inspection. Compare the picture of standards from direct observation with that from past results. Remember that you will need to explain any differences convincingly.

From your detailed evidence, you should identify the relative strengths and weaknesses which characterise what the pupils are achieving, so that you can give the department a clear picture of standards and how they can be improved. It is also essential to explain why standards and achievement are as they are. Some links between teaching and achievement have already begun to be made. These will be developed further in the next section.

2 TEACHING AND LEARNING

Read the previous inspection report carefully to gain a picture of the strengths and weaknesses of statistics teaching and learning at that time. This provides a reference against which you can gauge improvement.

2.1 What to look for in lessons

Evaluate how effectively pupils acquire knowledge, skills and understanding in statistics. Focus on evaluating whether pupils are learning to the depth and breadth and with the speed and motivation you would expect, given what you know about them. Interpret the general criteria in the Handbook in the context of statistics education, by asking questions such as whether pupils:

- show understanding of what they are learning, for instance, by successfully applying the knowledge and skills they have been taught, to analyse topical statistics issues, drawn from newspapers, the media and other sources such as the Annual Abstract of Statistics and Monthly Digest of Statistics;
- understand the benefits of co-operative working and use this opportunity to learn more, for example, when carrying out a survey;
- are interested and motivated by the subject so that, for instance, they produce pieces of sustained research or make the most of using modern information and communication technology (ICT) – for example, statistical packages, simulation software, databases, data logging software, spreadsheets/graph drawing software – when completing assignments.

Judge the quality of teaching by how well it promotes effective learning in statistics. Identify the main strengths and weaknesses in the teaching which best explain how well pupils are learning, and make the links between the two clear.

Keep in mind that good statistics teaching, which leads to effective learning, is rooted in: good understanding of the subject and its examination specification; high expectations; and methods of teaching which cater well for all pupils in the class.

As well as the characteristics of all good teaching, effective statistics teaching is likely to include features such as the use of:

- topical and local issues (such as canvassing views on a town by-pass), to stimulate pupils' interest;
- a good blend of classroom teaching and support for pupils' individual work on assignments;
- case studies and simulations (such as mini-enterprises) so that pupils have opportunities to apply their learning and to learn to make decisions;
- 'real-world' activities (such as stock movement and sales accounts), to extend pupils' knowledge and experience of statistics studies within a vocational context;
- ICT and other up-to-date resources, so that pupils learn how to use modern equipment and software;
- an 'exploratory' approach by teachers, to serve as good role models for their pupils.

Look beyond the surface when evaluating teaching that has positive features such as good relationships or firm discipline. It may lack rigour or depth, or a command of the subject, or be badly matched to the pupils' needs, with the result that learning is not as good as it should be. This might happen where, for instance, pupils:

- accumulate information on methods and techniques without reflecting adequately on their application;
- are left to work too much on their own when completing assignments;
- repeat activities (such as stratified sampling) without the challenge being raised;
- use worksheets which rarely demand extended answers;
- are given too much credit for well-presented work which masks shallow thinking;
- do not have their work marked until the end of assignments.

Teaching is judged by the impact it has on pupils' learning and, as a consequence, teaching and learning will usually be of the same quality. Where this is not the case, you should carefully note and explain why the judgements are different.

2.2 Observing lessons

In principle, the choice of lessons to observe should be determined partly by what you need to follow up from your analysis of pupils' books and other sources, and partly by the need to sample the work of different teachers. However, there is likely to be a very limited number of lessons each week and evaluation should make the most of all available evidence. To maximise the small number of opportunities, observation of whole lessons is likely to be worthwhile. All potentially significant evidence should be recorded (which may necessitate lengthy notes): events that may seem inconsequential at the time may subsequently prove the key to interpretation of overall patterns. For example, at Marion Richardson School, an apparently minor omission in lesson ST5 (that follows) significantly affects the subsequent lesson (ST6) with the same pupils. Hasty delivery and apparently harmless banter in ST7 illustrate general weaknesses that explain why standards are as they are.

Always evaluate the effectiveness of any additional support provided in lessons for particular groups of pupils, such as pupils with SEN or developing bilingual pupils.

Examples ST5, ST6 and ST7 for Marion Richardson School illustrate different styles for recording evidence in lessons.

Example ST5: Year 10 mixed-ability class of 15 boys and 12 girls; 50-minute lesson (3.30–4.20 pm)

Objective: to extend previous knowledge of stem and leaf diagrams, median and quartiles to back-to-back diagram of this kind and use of it for determining median and quartiles

[Instances of good/very good teaching T+/T++; instances of good/very good learning L+/L++]

Assumed previous attainment: pupils can go from raw data (for example, on train delays) to an unordered stem and leaf diagram and thereby to an ordered one; they can find the median and quartiles when they lie on or between scores.

Teacher is very well prepared and purposeful. She hands out sheets printed on squared paper, with a right-to-left leaf and stem diagram for ages at death in a mining village in South Wales in 1820, with spaces for a left-to-right plot. She gives out raw data for a rural village at the same time and explains that the pupils are to establish a comparison. She checks that they understand the right-to-left plot, which is new to them, by asking: 'How many died at the age of 35?', 'How many people died in their 30s?', 'How many died under the age of 20?' (T++). This shows that pupils are confident in applying previous learning to the new forms of presentation (L++). After 15 minutes, three-quarters of the class have correctly processed data for 70 deaths in the second village, recording it in a matching left-right plot (L++). The teacher picks up slips such as failing to transcribe repeated data, reminding pupils to cross off data in the unordered diagram once it has been transcribed, in order to speed up scanning the lines (T++). Though some have not finished, to move things on, she has a completed overhead projector slide of the final pattern. It is apparent that people tended to die at an earlier age in the mining village. Pupils appreciate the importance of squared paper to ensure that leaves are set out at regular distances from the central stems; otherwise the pattern would be corrupted (T+, L+). Pupils speculate about the reliability of this picture: boy asks if there could have been an epidemic in the mining village. Teacher asks how such an epidemic might affect the shape and size of a standard leaf diagram (T++). Girl makes points about increased deaths (leaves) overall and possibly some age ranges (very young and very old) more affected than others (L++). Boy says that diagrams for successive years would indicate abnormal data that might be due to an epidemic (L++). Pupils who have not completed the diagram are given a sheet to take away with the completed pattern (T+). Homework is to find Q1, Q2 and Q3 for the data in each diagram; pupils are advised to total the number of leaves on each stem first (T+).

Summary

The teacher has chosen the task well: it uses interesting data and allows the value of back-to-back stem and leaf diagrams to be demonstrated clearly. She is assiduous in checking on understanding as the lesson unfolds. She has taken care of the possibility of irregular plotting (squared paper) and has anticipated that not all pupils will complete the task at the same time. She makes very good use of pupils' observations about the patterns to further their understanding. She gives a helpful hint to get them started on the homework. There is evidence of very good subject knowledge, very good rapport, the maintenance of a good pace of working and suitable challenge to pupils across the ability range. Very effective learning for boys and girls. If the pace had been just a little faster, it would have been possible to demonstrate finding the quartiles for one set of data, to minimise homework difficulties.

[Teaching and learning very good (2)]

Commentary

Teaching has been well thought out and things that could go wrong have been anticipated through good planning: it is a very successful lesson. This evidence is consistent with the evidence from talking with pupils and looking at their work, which indicated effective progress in Year 10. It is not excellent because, while the pace has been good, a little more could have been done to prepare pupils for the homework task. Teaching and learning in a lesson can only be judged on the concrete evidence apparent in the lesson. Example ST6 for the subsequent lesson shows that the evaluator's suspicions about the homework are borne out and pupils' insecure earlier learning hampers the lesson.

Example ST6: Year 10 mixed-ability class of 15 boys and 12 girls; 50-minute lesson in main timetable

Objective: to build on previous lesson, attending to homework problems, moving on to box and whisker plots and notion of outliers for higher-tier pupils

Previous attainment: this follows the lesson in ST5.

[S = strengths; W = weaknesses]

W Marked work shows that about half the pupils have made mistakes in finding quartiles Q1, Q2, Q3 on back-to-back stem and leaf diagrams (see ST5). The most common error is working from the left in the reversed diagram in moving from deaths in the teens to those in the 1920s and 1930s. For example, the pupils have worked down from 39, not up from 30. A few have made similar slips in the normal diagram, possibly by counting up from the bottom of the diagram (high ages) and working along leaves from left to right (low to higher). Out of 67 pieces of data, a few have jumped to the conclusion that the median age at death is 34 years. With 70 pieces of data, some have taken the 35th leaf as Q2 rather than halfway between the 35th and 36th. About half the Q1 and Q3 values are wrong in some way.

S The teacher identifies likely reasons for incorrect answers and is painstaking in annotating books accordingly.

S She has an overhead projection slide with the back-to-back diagram with totals for the number of leaves on each line (for example, 19 deaths among people in their 30s). She presents a systematic procedure for finding the median in the reversed diagram, explaining slips that can occur. She then gets pupils to talk through finding a first quartile. They do it successfully. Pupils are to do corrections as part of homework. She has sensibly avoided the confusing notation of a 'depth column' that is sometimes used to home in on the median in a stem and leaf diagram (that is, running totals from both top and bottom of distribution). All this takes about 15 minutes.

W The evaluator finds that some uncertainty about lower and upper quartiles persists: pupils calculate Q1 for 67 pieces of data by finding 67/4 rather than (67+1)/4; consequently they are not sure why Q1 is the 17th piece of data rather than half or three-quarters of the way between the 16th and 17th. Likewise, why is Q3 of 70 pieces of data not the midpoint between items 52 and 53?

S The teacher clearly and succinctly explains a general diagram showing how Q1, Q2, Q3 and the largest and smallest values can be represented in a box and whisker plot with a horizontal axis. Pupils give good attention.

W Pupil known to be gifted makes comparison with kite diagrams in biology to show, for example, variation in population density of molluscs between high and low water on a beach. Teacher seems unfamiliar with this notation and does not encourage this distraction. EAL pupil asks why it is called a 'whisper'.

S/W Pupils are to continue with the 67 items of data that have been considered in going over the homework, to find Q3, the smallest and largest values, and to present it as a box and whisker plot. They set about this speedily. The uncertainty with Q3 becomes apparent, pupils doing various things between the 50th and 51st value, but there are no other slips in dealing with the stem and leaf diagram as such. All pupils draw the box and whisker plot accurately from the five pieces of data they have.

S The lesson now proceeds with appropriately differentiated work. Foundation-tier pupils are to make a box and whisker plot for the other stem and leaf diagram: a suitable consolidation task. With higher-tier pupils, teacher clearly explains outliers; pupils are to identify outliers for the first stem and leaf diagram (in the event, at the lower end the criterion falls below zero, but several outliers at the upper end illustrate a point well). Teacher monitors both groups as they work independently. Brief consolidating plenary with the upper group, who are then moved on to completion of tasks on the second diagram.

S To diagnose problems with median and quartiles, six simple tasks are set where $\frac{1}{2}(n+1)$ *is/is not an integer and similarly for* $\frac{1}{4}n/\frac{3}{4}n$*. (Higher tier also to complete classwork.) This is likely to identify exactly what they have difficulty with.*

This is a well-conceived lesson, which is hampered by pupils' insecure understanding of procedures for finding medians and quartiles. This reduces the pace. The new work has been learnt well. The teacher is aware that further work needs to be done on medians and quartiles, and she shows good subject knowledge and flexibility in setting homework likely to pinpoint the difficulties requiring remedial work.

[Teaching and learning satisfactory (4)]

Commentary

The evidence from the earlier part of this lesson may look like an evaluation of attainment. However, it is very pertinent to the effectiveness of teaching and learning. It also reflects on the effectiveness of learning over time and the teacher's failure to prepare pupils for the homework in the previous lesson. Overall, the strengths of the lesson outweigh the weaknesses: it is coherent, follows through well on the earlier lesson, and provides for differences in ability. The homework has been closely marked, there is sufficient attention to correction of it in the lesson, and this is followed up with diagnostic homework.

Example ST7: *Year 11 mixed-ability class of 17 boys and 13 girls: 50-minute lesson (3.30–4.20 pm)*

Objective: to consolidate scatter diagram work with foundation-tier pupils and introduce Spearman's rank correlation coefficient for higher-tier pupils. (NB Spearman is in higher syllabus specification only.) About one third of the class arrive up to 10 minutes late, but there is almost full attendance

Previous attainment: scatter diagrams (in motivating contexts, for example, goals scored and league position, annual unemployment levels and average pay rises); appraisal of correlation by computing mean x and y and thus relocating the origin.

Teaching

Recapitulation on notion of correlation: seeks examples, encouraging precision about likely strength of correlations. Apparent that the teacher is 'padding this out' until latecomers are settled. This uses up potential productive time.

Data given of blood pressure, age and weight for nine men. Foundation candidates (15) are to plot scatter diagrams of blood pressure against each variable and apply the 'mean x/mean y' displaced origin method. Higher candidates are presented with Spearman's rank correlation coefficient formula: and overhead projection slide with columns completed for blood pressure (b), weight (w), rank b, rank w, d and d2. An interesting but demanding example to start with: there are three tied scores for blood pressure at ranks 1, 2 and 3, all of which are ascribed rank 2 and this is mentioned only in passing. Pupils are to compute the coefficient and repeat for blood pressure and age. Teacher sets out to monitor work of both groups. He is making a reasonable fist of balancing the needs of foundation and higher pupils by differentiated tasks that harmonise and inform one another.

The teacher keeps pupils working fairly speedily. He notices and deals appropriately with reverse ranking but does not spot that some final ranks differ from 9.

Plenary in which results of the two methods are compared. Good encouragement to less able boys to participate. However, girls are relatively passive and the teacher does little to get them to contribute ideas.

Homework: sheet with data about population density and distance from a town centre; pupils of each group to repeat the task. (Looking for negative correlation; no tied scores – hence it

Learning

Pupils are forthcoming in suggesting factors which might correlate positively, negatively or not at all. Some banter, in which boys eagerly suggest correlations between punctuality to the lesson and pupils' test scores, merit grades, leg length and other personal characteristics. (It turns out that latecomers have been at the staffroom fixing appointments for parents' evening.)

Foundation pupils are dexterous, working at a good pace to make accurate plots. Most can relocate the origin correctly. Both scatter diagrams have two plots in the new second quadrant: the pupils say there is a weak positive correlation and a few think it is a little stronger for weight. This is worthwhile consolidation, but pupils are largely honing existing skills rather than extending them. Higher-tier pupils have no difficulty using calculators to obtain a coefficient of 0.35 for weight. Pupils correctly set up a table for computation of correlation with age, copying the ranks for blood pressure. However, they are unable to explain to the evaluator the general procedure for what rank to assign when there is an even number of scores tied at the same point. As there are no tied scores for age, the problem is not revealed in the work. One or two have final ranks for age of 8 or 10, suggesting lack of systematic procedure; two pupils have ranked ages from lowest to highest (contrary to blood pressure). Differences are correct, though there is generally unnecessary preoccupation with signs. Only three pupils work through to a correct correlation of 0.20 in the time available. The task has not really revealed whether pupils understand or not.

Pupils can see that the coefficients give a useful quantitative measure consistent with the impression from the diagrams. Foundation boy observes that diagrams show consistent high blood pressure with high weight and advanced age and that this cannot be deduced from the coefficient alone. Pupils of all abilities can use their science and physical education knowledge to relate blood pressure to more precise factors: physical activity, anger, constriction of blood vessels, the tone of heart muscle and 'pills'.

More able pupil asks if she can try reciprocals and is encouraged to do so (but of what – both?)

will not become apparent whether this matter has been clinched.)

Teacher is keen to use the final five minutes to put correlation in a historical context and point the way for coursework projects. Quick reference to Sir Francis Galton's work in the 19th century on heights of parents and offspring and his association with Karl Pearson to develop a correlation coefficient (about 0.5). Mentions Galton's term 'eugenics', inherited intelligence, genius concentrated in the British, eugenics public policy in Singapore with financial incentives for graduates to have babies and for non-graduates to be sterilised after their second child. Why not repeat Galton's experiment as coursework? All rather garbled through pressure of time: important and sensitive issues about the use of statistics are dealt with clumsily. Lesson runs a few minutes overtime.

A few pupils have to leave to catch the bus. Majority attentive and interested. Boy captures the issue, albeit flippantly, 'Can we do that to those on foundation, sir?' Banter about measuring granddad's toes, nose etc when they get home.

Teaching is enthusiastic, but rushed because of late arrivals, insufficient time overall for the course and being behind schedule in covering the syllabus specification. There are difficulties in having to provide for pupils on the two different specifications. However, very interesting contexts are chosen and tasks are appropriate to needs. The finer points of the Spearman procedure are not explained thoroughly and neither classwork nor homework allows the teacher to find out whether pupils understand about tied scores. Some slips in ranking are not identified and therefore remain undealt with. There is very effective evaluation of the two approaches, which brings the two groups together in discussion. The final brief exposition is rushed.

[Teaching satisfactory (4); learning unsatisfactory (5)]

Commentary

The teacher is under even more pressure of time than usual because of the late arrivals, which are beyond his control: the teaching is graded higher than the learning. There has been a chance for foundation-tier pupils to consolidate their existing learning. Higher-tier pupils have begun to learn about working out the Spearman coefficient, but some opportunities to secure and check their understanding and to attend to slips have been missed. Very good evaluation of the different methods has been promoted. The final exposition was ill-conceived and possibly did more harm than good. The teacher acquiesces in a 'laddish' culture and does not always draw out participation from the girls. However, he has excellent subject knowledge, conveys enthusiasm, covers a lot of ground in the time available and generally maintains good attention and commitment among the pupils. So teaching is just about satisfactory, though learning is unsatisfactory on account of the unforeseeable problems at the start. This observation corroborates the impression, from talking with pupils and looking at their work, that teaching in Year 11, and for the higher tier in particular, is too rushed and, in consequence, these pupils underachieve.

2.3 Other evidence on teaching and learning

Overall judgements on teaching and learning must be more than a simple averaging of the grades from the individual lessons observed, though these are important findings. Valuable evidence also comes from the analysis of pupils' work, talking with pupils and the quality of departmental planning. Earlier examples (ST2 and ST3) show examples of this.

Weigh carefully all the evidence available to reach a balanced overall judgement. Make the links between teaching and pupils' standards and achievements crystal clear. Above all, you will need to ask yourself:

- if the judgements on teaching explain why standards and achievement are as they are;
- whether there are other aspects of provision which help to explain that picture.

If your judgements about teaching and what is achieved appear inconsistent, you must identify which other factors have a significant impact.

3 OTHER FACTORS AFFECTING QUALITY

The quality and impact of the curriculum, staffing, resources, accommodation, and leadership and management in statistics should be evaluated, and reported if the impact has a significant bearing on what is achieved.

Factors to consider in provision for statistics include:

- the time allocation over the two years (and any use made of time in Key Stage 3, for instance, after national tests);
- the arrangements for teaching the different syllabus specifications to foundation-tier and higher-tier pupils;
- opportunities for pupils to get a grasp of what they are opting for when they make subject choices in Year 9;
- the management and leadership of statistics as a subject in its own right (for example, whether it is taught by two or more teachers with overlaps and omissions resulting from inefficient co-ordination).

4 SUMMARY

Bring together your various judgements to form a coherent view of the subject. You should provide a convincing summary which explains why standards and achievement are as they are. The summary should refer centrally to the quality and impact of teaching, and include any other important factors that are necessary to account for what is achieved. Show how much the subject has improved since the last inspection and give a clear indication of the action needed to improve it further.

Below are some points which may be included in summary evaluation of statistics at Marion Richardson School. (These points also involve additional evidence not found in examples ST1-ST7.)

Example ST8: outline evaluation of statistics

The standards reached by 16 year olds are above those that might be expected from pupils of an average range of ability.

Pupils on foundation tier achieve well; those on higher tier achieve satisfactorily (though not as well as in other subjects). Pupils who follow the foundation specification show strong achievement on most of the required topics, but are insecure with finer points such as finding quartiles where the data are not straightforward. Those on the higher specification have a basic knowledge of more advanced topics, but have limited understanding of applications such as standard scores.

On the whole, teaching is good. Teachers generally cope well in providing appropriately different activities for foundation- and higher-tier pupils. However, time over the whole course is too short (by about 25 minutes a week) and consequently some of the more demanding topics in the higher specification are underdeveloped.

The department's evaluation of its work is robust and reliable.

Pupils are well motivated by work in contexts which relate to real life and a wide range of other subjects of the curriculum. They feel that statistics enhances their work in mathematics and other subjects and can see its importance for further studies. Pupils at the extremes of the ability range (and others such as Travellers and EAL pupils) are suitably catered for. Girls are not always sufficiently engaged in class discussion and they do not outperform boys to the extent that they do in most other subjects.

The subject is a valuable element in the curriculum. The school is entrepreneurial in facilitating it. It helps pupils' progress in other subjects and adds directly to pupils' total GCSE points score and therefore the school's average total points per pupil. Pupils' weaker grades relative to their other subjects reflect a national pattern, contributing to a lowering of some pupils' (and the school's) average GCSE points per entry.

Point for improvement

The school should find ways of providing additional teaching time for pupils working for the higher specification, in order to secure competence in the application of more advanced topics.

Draw on your detailed evidence base to amplify and illustrate the points made in such an outline summary

Office for Standards in Education
33 Kingsway, London WC2B 6SE

Telephone 020 7421 6800
www.ofsted.gov.uk

Designed and printed by
Matrix Print Consultants Ltd
Rothwell, Northants 01536 713811