CURRICULUM CORNER

Reasoning with Data – Time for a Rethink?

James Nicholson, Jim Ridgway and Sean McCusker
University of Durham, England.
e-mail: j.r.nicholson@durham.ac.uk
　　　 jim.ridgway@durham.ac.uk
　　　 sean.mccusker@durham.ac.uk

Summary
Reasoning with data is already pervasive in society, and its importance as a life skill is increasing. We argue that the current statistics curriculum in the United Kingdom at the secondary level does not prepare our young people adequately, and suggest ways in which it could be improved.

◆ INTRODUCTION ◆

The current provision for data handling and statistics in the UK focuses primarily on procedural knowledge, applied to one or two (linearly related) variables. We contend that this provision is inadequate for informed citizenship in the 21st century. In the UK, major government reports from Tomlinson (2004) and Smith (2004) have opened a discussion on the role and locus of reasoning with data in the curriculum. Internationally, there is a need to address the key issues of the nature and development of skills in reasoning with data, and how this development can be promoted, especially if development is supported appropriately by good graphic interfaces.

The English National Curriculum sets out programmes of study for every curriculum subject for students aged 5 to 16 years. Northern Ireland and Wales have structures, and a curriculum, which are very similar to England, and the curriculum in Scotland is also similar, although the structures differ slightly from those in England. We will use the term UK throughout this article, with the detail referring to the English National Curriculum at secondary school level unless otherwise stated.

Statistics occurs as a topic for study under Data Handling, one of four strands in Mathematics (the other strands are Using and Applying Mathematics; Number and Algebra; Shape, Space and Measure). There are rich opportunities across the curriculum to develop an understanding of complex issues with the help of appropriate statistical techniques (notably in Science; Geography; History; Personal, Social and Health Education; Business Studies; and Citizenship). Unfortunately, the National Curriculum was written subject by subject with no attempt to produce coherence across subjects; as a result, the same topic in statistics is to be presented to students of different ages in different subjects, and some important topics do not appear at all (see Holmes (2000) for a detailed analysis). Another consequence is that teachers with no background in statistics education are introducing statistical concepts to students, with little appreciation of the inherent difficulties associated with those concepts.

Recent developments at ages 11–14 years (Key Stage 3 in the UK), and in its assessment in the Statutory Assessment Tasks (SATs) at age 14 years, have tried to move data handling beyond rote learning of rules and techniques towards a better understanding of concepts such as mean and measures of spread. Nicholson and Darnton (2003), for example, explored some of the issues for

1.   Four boys and two girls received text messages.

> The mean number of messages received by the four boys was 20
>
> The mean number of messages received by the two girls was 26

Use the information in the box to decide if each statement below is True or False.

(a)   The person who received the most messages must have been a girl.

☐ True   ☐ False

Explain your answer.

(b)   The mean number of messages received by the six people was 23

☐ True   ☐ False

Explain your answer.

Fig. 1. *Text messages*: Changes to assessment 2003: Sample materials for Key Stage 3 Mathematics, QCA 2002.
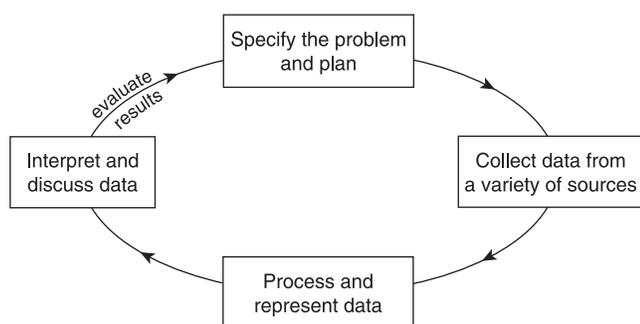
Fig. 2. The data handling cycle in the National Curriculum.

classroom teachers raised by items such as *Text messages*, shown in figure 1, which required pupils to go beyond knowing how to calculate the mean, and enabled those who could do so to demonstrate a deeper understanding of the concept.

While it is difficult in timed, written tests to assess the whole of the data handling cycle (see figure 2) adequately, the SATs have also tried to reflect the notion that statistics is about understanding data, and to move towards using real data, in context, rather than using stylized examples.

In high-stakes assessment at age 16 years (General Certificate of Secondary Education (GCSE)), the introduction of the compulsory data handling coursework project in Mathematics caused considerable angst among teachers and students. To score highly on it requires reasoning with complex data, for example by looking at trends or making comparisons within and across groups. However, this had not been part of the mathematical diet of either the students or, for the most part, their teachers, and good teaching and learning resources were scarce. Displays that facilitate the analysis

of complex data – such as the use of colour – are expensive to produce. To learn to deal with data properly, students need exposure to a range of contexts and scenarios where the relationships are of different strengths so that they can develop appropriate language to describe data. Moreover, while text-books traditionally provide answers to all numerical problems, it is common that no answers are provided for any questions requiring discursive responses, as these do.

As well as Mathematics, which is part of the statutory curriculum, two examination boards also offer a GCSE in Statistics where there is a greater emphasis on interpretation, but at present the take-up of these courses is limited – a total of 39,666 candidates in 2004 compared with an entry of 741,682 in Mathematics across England, Wales and Northern Ireland.

In high-stakes assessment at age 18 years (Advanced level (A-level)), our recent analysis of all the specimen papers for statistics modules offered within the six Mathematics specifications showed between 3% and 33% of the assessment credit was awarded for modelling and interpretation skills, and the remainder rewarded procedural competence and computational accuracy. Moreover, there is almost no reasoning from evidence: hypothesis testing and generating confidence intervals are included, but are assessed within standard formats revealing very little of the depth of understanding a candidate has of the nature of what can legitimately be inferred from the evidence available.

On a more positive note, there is a new A-level Statistics specification, offered by the Assessment and Qualifications Authority (AQA), which can be taken alongside Mathematics, in which greater emphasis is placed on modelling and interpretation skills, for which the proportions of marks available range from 25% to 52%, and the Mathematics for Education and Industry (MEI) project has recently announced an AS in Statistics with similar aims.

◆ IS THE CURRENT SITUATION ◆ SATISFACTORY?

We believe there are two issues relating to the curriculum here. One is that students do little in the way of reasoning with data even within the univariate and linear bivariate arena. The other is

Main reason for economic inactivity: by age and sex, spring 2003[1]

United Kingdom				Percentages
	16 to 24	25 to 34	35 to 49	50 to 59/64
Males				
Long term sick	4	40	60	42
Looking after family	1	11	16	3
Students	84	24	5	0
Retired	0	0	1	23
Other	11	26	18	32
Total	100	100	100	100
Females				
Long term sick	3	8	25	35
Looking after family	23	75	59	25
Students	65	8	4	1
Retired	0	0	0	13
Other	9	9	12	27
Total	100	100	100	100

What does this mean? Is it that 84% of all males aged 16 to 24 are students? Or that 84% of male students are aged 16 to 24? Or something else?

[1] *Data are not Census adjusted.*
Source: Labour Market Trends, October 2003 (see reference 4)

Fig. 3. Reasoning about data from a two-way table.

whether or not that arena is too limited for a 21st century curriculum, and it is this issue we deal with first.

◆ WHAT IS MISSING? ◆

Ridgway et al. (2004) point to the complexity of modelling in realistic settings. Almost everything around us is influenced by multiple factors, which have interactions and whose effects vary in size and on the timescale in which they take place.

There is a quantum leap between dealing with one variable, or two variable linear data, and working with multiple variables that include interactions and nonlinear relationships. Working with greater numbers of variables increases the complexity incrementally, but the framework for handling such situations can be learnt by working with three or four variables.

The Statistics curriculum today is dominated by the statistics developed in the pre-computer age. If a curriculum were to be devised from scratch in the light of current statistical knowledge, it would not be dominated by univariate parameterized distributions and linear bivariate models. It would include bootstrapping methods, nonlinear regression and multiple regression models, all of which are computationally intensive but conceptually accessible. It might include time series modelling with simple autocorrelations, to introduce the concept of feedback loops.

Very large data sets are now available to schools for analysis. Here, almost any way of partitioning the data set is likely to reveal statistically significant differences, which will not necessarily have any practical or conceptual importance; that is, analysis will often show very small 'effect sizes' with no real-world 'significance'. Schield (2004) observes that while chance (random error) dominates in small-sized, well-designed experiments, bias (systematic error) can dominate in poorly designed studies regardless of size, and confounding (the influence of a lurking variable) dominates in populations or large-scale well-designed observational studies.

Media accounts often make use of two-way tables. The new curriculum should place greater emphasis on the understanding and interpretation of such tables, and reasoning using conditional probability from them. Schield (2001) argues the case that students have more need of a sound understanding of the grammatical rules involved in describing rates and percentages, and in interpreting tables presenting such information, than of many other statistical ideas. Figure 3 provides an example from the National Statistics web site http://www.statistics.gov.uk/downloads/theme_compendia/fosi2004/SocialInequalities_full.pdf.

Students should deal explicitly with counter-intuitive realities arising from Simpson's paradox so that they become aware of the potential difficulties in trying to use causal reasoning based on observational studies. Simpson's paradox can arise when data from unequal-sized groups are combined into a single data set, and there are one or more

	Accepted	Rejected	% accepted
All boys	107	101	51.4
All girls	89	118	43.0
Arts boys	22	54	28.9
Arts girls	57	106	35.0
Science boys	85	47	64.4
Science girls	32	12	72.7

Fig. 4. Illustration of Simpson's paradox.

confounding variables present. The paradox arises when the direction of a relationship in each component part of a population is reversed when the population is considered as a single entity.

If the data for all applications to a university, shown in figure 4, are considered, it appears that boys are favoured as 51.4% of applications from boys have been successful, against only 43.0% of girls. However, if the success rates for applications to both the Science and Arts faculties are considered separately, the girls have fared better in both instances: the reason that overall the boys rate is higher is that a higher proportion of boys applied to Science which had a markedly higher success rate than had applications to the Arts faculty.

Crowe (2002) illuminates Simpson's paradox vividly through a data story told through the eyes of a decision maker in public health, showing the power of interesting and relevant contextualization, which can illustrate important statistical ideas.

Students also need to be aware of some important ideas about measurement. One example is the difference in the factors that influence measures of physical constructs (e.g. mass, temperature) and measures of social constructs (e.g. health, school success). For example, in the social domain, what is measured by a particular indicator can change if the measure changes its status from a low-stakes measure to a high-stakes measure. Examples can be taken from target setting in public services. The proportion of students achieving 5 or more grade A* to Cs at GCSE was chosen as a good, broad indicator of school performance. However, the status of this statistic in determining school league table positions, and in target setting, has resulted in schools being motivated to target scarce resources to support the group of students predicted to be close to the borderline on that measurement. As a result, what was once a good proxy measure of general school academic performance has lost its value as a general indicator. Increased performance on that measure can be viewed as demonstrating a rise in standards, but it could equally be the result of raising scores on a specific measure without a concomitant rise in general standards. The report of the Royal Statistical Society (RSS) Working Party on Performance Monitoring in the Public Services (2005) offers an extended discussion of the issues raised by performance indicators.

◆ COULD WE DO BETTER ON THE ◆ CURRENT CURRICULUM CONTENT?

We give students lots of practice in solving all levels of algebra problems, and the level of difficulty is steadily built up. We have a reasonably good understanding of which techniques are inherently more complex, or require more abstract thought, and also how changing the details of a certain type of equation can substantially alter the level of demand. So, for example, solving quadratics is harder than solving linear equations; and while $12x + 7 = 10x + 11$ and $x + 5 = 4x + 9$ are structurally similar equations (at first glance one might even think the first one is the harder, because it has larger numbers in it), the second will cause many more difficulties for pupils because the answer is both negative and not an integer.

The same is true for all areas of the mathematical curriculum – drill and practice of all techniques is extensive, and there is a pretty well understood hierarchical structure for all strands – except for reasoning with data. Peter Holmes in the Schools Council Project on Statistical Education (1980) provides a comprehensive ordering of the concepts and techniques for the contents of standard statistics courses, and Rangecroft (1991a, b) provides a suggested progression for graph work in the curriculum which was in place at that time, but neither framework includes reasoning with complex data.

There is an urgent need for curriculum materials that provide more in the way of experience for pupils in reasoning with data than they have done previously. For example, consider the two graphs shown in figures 5 and 6 (CCEA 2002). These show relationships in a data set available on the Council for the Curriculum, Examinations and Assessment (CCEA) coursework Web site (http://www.rewardinglearning.com/development/qualifications/gcse/gcse_math_internally_ass.html). The data set is useful as an example despite its limitations – the data set does not meet the minimum size (at least 20 variables with at least 100 data points
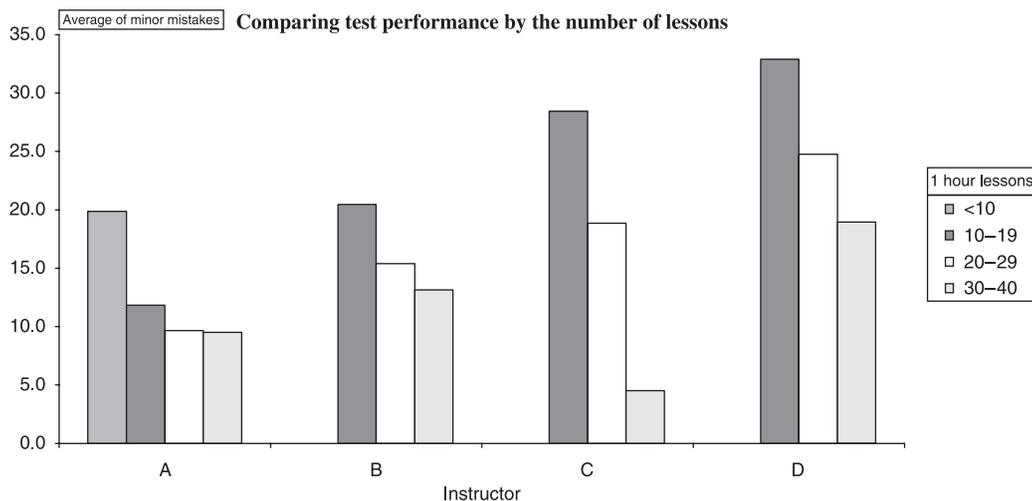
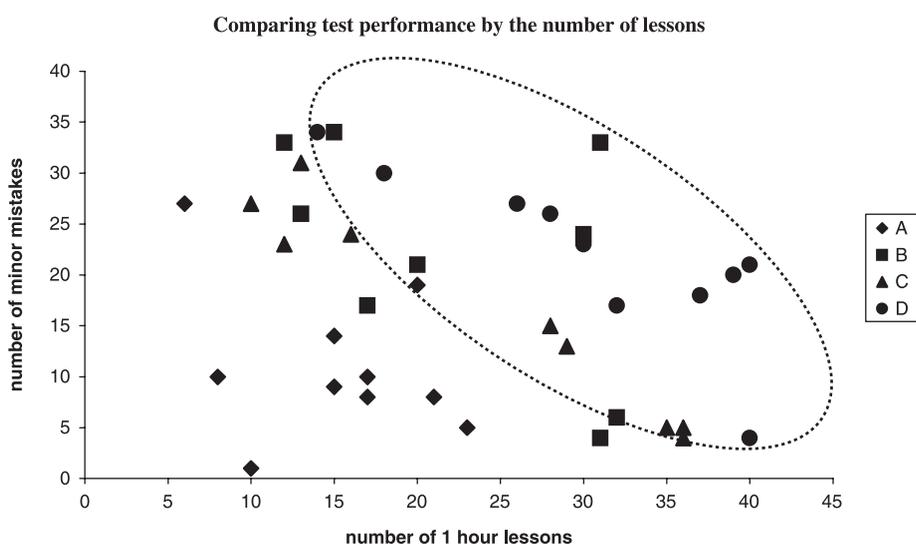**Fig. 5.** Comparative bar chart for driving test data.

Fig. 6. Comparative scattergraph for driving test data.

each) set out by the Qualifications and Curriculum Authority (QCA) (Browne, 2002).

One way this could be used would be as a classroom starter – display the comparative bar chart (figure 5) through a data projector, and ask the students to write down three things they think the data tell them. However, we need to appreciate that this is not a trivial task, and students need to be guided through the same sort of skills-building framework they experience within algebra or geometry or number work. In particular, for groups considering this sort of problem for the first time, or for younger, or weaker, groups, the above question is likely to be met with a blank response, because it is too open ended. It could be asked in a more closed manner, for example 'Which do you think is the best instructor? Which do you think is the worst? Explain your reasons.' This can lead to a discussion in the classroom about the different suggestions.

Getting all students to write something down – even though it will remain private to them unless it is contributed to the discussion stage – is more likely to result in them engaging with the process, and they are more likely to contribute ideas to the discussion if they already have put something down on paper about it.

The diagram in figure 6 displays related data in a different form – it actually has more detail but only of a sample, where the first diagram shows average statistics for the whole group. Showing the data initially just as the scatter diagram, and then with groups summarized, either by the sort of 'envelope' shown for instructor D in figure 6 or by the regression line for each instructor, helps the students to learn the sorts of questions to ask themselves when looking at complex data – in this case, 'are the outcomes shown for the different instructors really different or not?'
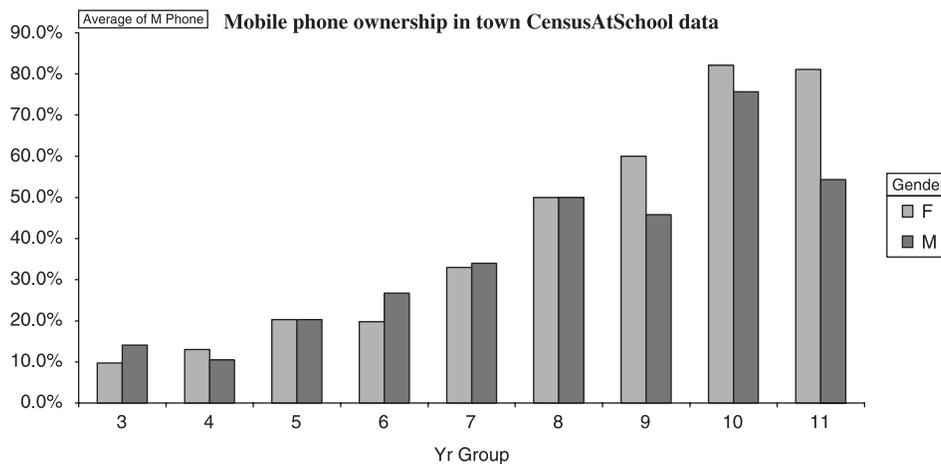
Fig. 7. Comparative bar chart for CensusAtSchool data on mobile phone ownership.

After gaining some experience at dealing with individual graphical displays, we would then want the students to consider questions such as: Do the two diagrams tell similar stories? Is there a difference in the information that is accessible through the two diagrams? We want them to explicitly consider such things so that they develop the metacognitive skills to make decisions as to what the most effective graphical representation is to illuminate a particular relationship, and to be aware that often there is more than one representation available.

Presenting a variety of data sets, with various graphical displays, also allows us to cover a range of situations, and helps develop students' ability to make judgements about what factors are more or less important. For example, figure 7 presents data from CensusAtSchool (downloaded from http://censusatschool.ntu.ac.uk/index.html) on the proportions of mobile phone ownership, in terms of the age and sex of the respondents. This set of data refers to pupils at schools situated in towns. There are differences between boys and girls, but they are not as large as the difference between age groups.

Classroom activities where students are presented with complex graphical representations that are discussed in qualitative and semi-quantitative terms can build confidence in students in their ability to say something meaningful about multivariate data sets, and can increase their confidence in exploring a variety of graphical representations which might illuminate a data set that they are investigating for themselves.

There is a need for a bank of appropriate resources to be accessible for teachers – technology in the form of data projection is increasingly available in schools with a minimum of disruption to the rhythm of the teaching day. However, good data sets are not always easy to come by, and certainly at the moment teachers need guidance as to how to structure the tasks, and how to mediate the discussion.

◆ RELATIONSHIPS WITH OTHER ◆ SUBJECTS

Many other subjects depend implicitly, if not always explicitly at the school level, on reasoning from evidence. While we believe that the core responsibility for teaching and learning statistics lies within the mathematics curriculum, it is vitally important that statistics is not seen as a set of abstract mathematical techniques.

There is already pressure on curriculum time, and there are demands for greater fluency in other important areas of mathematics such as algebra, so something has to give. We believe there is scope for substantially reducing the amount of time spent on repetitive, routine tasks such as calculations of summary statistics and graph drawing, which are now automated in virtually every working environment, and replacing it in mathematics by the core skills of reasoning with complex data, supported by the co-ordinated use of those skills in other subjects to encourage transferability. This should make the mathematics curriculum more relevant, and also create some extra time within mathematics to improve other key activities such as algebra, while not reducing the total amount of time students spend on working with data in the curriculum.

Transferability of key skills has always been an important educational aim, but has proved elusive. We believe there is now an opportunity to use technology to integrate reasoning with data into other curriculum areas; this would help students to see the relevance and pervasive nature of working with data in the modern world. Reinforcement of key ideas repeatedly in different contexts will enhance transferability.

However, cross-subject collaboration is not a trivial undertaking: most curriculum areas feel their content coverage is threatened by the limited time they have available. A major challenge is to develop data-rich materials that can be used to teach core curriculum concepts effectively.

Another possibility lies in the extended project that was proposed for the diploma structure in the Tomlinson Final Report (Tomlinson 2004). While the government has not accepted its proposals en bloc, and the use of identifiable GCSE subjects as core components of assessment at 16 is likely to remain, there is still a possibility that coursework across many subjects might be subsumed into a single extended project. Learning the processes of gathering evidence and using technology powerfully are important skills that could be harnessed in such an extended project. Some useful sources of ideas and materials have been developed, though far more needs to be done.

The UK Data Archive project (2004) on teaching Political Science using crime data in the Archive had two main aims. The first was to produce materials such as teaching and learning modules, and the second was to track the development process of creating and using e-learning data-rich resources. There are now Web-based resources hosted at the UK Data Archive Web site; these include printable and reproducible hard copies of the modules and a teacher's guide to accompany the resources.

The Data Archive has a very large collection of rich data sets, and the NSDSstat Data Analysis Program used in the Political Science materials provides a powerful interface with the data. Excel's pivot tables can be extremely powerful for analysing certain data structures, and are relatively easy to learn to use for this purpose. Other technologies are becoming available – the World Class Arena Problem Solving Tests used a number of new interfaces with data, which allowed pupils at ages 9 and 13 to explore relationships between multiple variables (Ridgway and McCusker 2003). A key requirement of those

interfaces was that they had to be transparent – in an assessment task, it was important that the students were not distracted by learning to drive the interface before engaging with the data. An example where even younger children work with multivariate displays is provided by du Feu (2005). Six-year-old students created a 3-dimensional representation using Lego blocks to aid their understanding and interpretation of data collected on a field trip.

It is extremely important that collaborations between statistics educators, subject specialists from other curriculum areas, and software developers get off the ground soon. For these sorts of materials to find their way into mainstream curriculum requirements, even on the Tomlinson timescales, there needs to be evidence that they can be successfully developed and used effectively, and manageably, by representative classroom teachers working in ordinary classrooms. The difficulties of ensuring that teachers in mathematics, and in other curriculum areas, can use even high-quality data-rich resources effectively in the classroom should not be underestimated, and there will be a very substantial need for professional development support for teachers.

◆ CONCLUSIONS ◆

The current curriculum serves students poorly. Interpreting evidence is a critical skill for personal well-being in terms of the decisions one makes, and more generally in terms of one's ability to engage meaningfully in debates about complex issues. Currently, students are posed problems involving just one or two variables, and most of the questions they are asked are designed to develop statistical technique that could be applied to any data set, rather than better conceptual understanding of the situation from which the data derive. The current curriculum ignores many of the recent developments in statistical thinking, and the power of computers to present multivariate data in accessible ways. There is a need for curriculum and assessment materials that use complex data sets, and which require analyses that cover a broad spectrum, from qualitative description through to quantitative analysis. Some examples of appropriate work are given here. Current developments in UK education offer an opportunity for greater collaboration across subject boundaries, and the inclusion of more extended projects.

References

Browne, R. (2002). Statistics in the Revised National Curriculum for England. *Teaching Statistics*, **24**(1), 2–4.

CCEA (2002). *Driving Tests*. GCSE coursework data set available at http://www.rewarding-learning.com/development/qualifications/gcse/gcse_math_internally_ass.html.

Crowe (2002). *Simpson's Paradox*. http://www.cawtech.freesereve.co.uk/simpsons.2.html.

du Feu, C. (2005). Bluebells and bias, stitchwort and statistics. *Teaching Statistics*, **27**(2), 34–6.

Holmes, P. (2000). *Statistics Across the English National Curriculum*. Royal Statistical Society Centre for Statistical Education, http://www.rsscse.org.uk/resources/natcur.html.

Nicholson, J.R. and Darnton, C. (2003). Mathematics teachers teaching statistics: what are the challenges for the classroom teacher? *Proceedings of the 54th Session of the International Statistical Institute.* Berlin, Germany.

QCA (2002). Text messages. In: *Changes to Assessment 2003: Sample Materials for Key Stage 3 Mathematics*. London: Qualifications and Curriculum Authority.

Rangecroft, M. (1991a). Graphwork – developing a progression. Part 1 – the early stages. *Teaching Statistics*, **13**(2), 44–6.

Rangecroft, M. (1991b). Graphwork – developing a progression. Part 2 – a diversity of graphs. *Teaching Statistics*, **13**(3), 90–2.

Ridgway, J. and McCusker, S. (2003). Using computers to assess new educational goals. *Assessment in Education*, **10**(3), 309–28.

Ridgway, J., McCusker, S. and Nicholson, J. (2004). Uncovering student statistical competences via new interfaces. *IASE Roundtable Proceedings.* Available at http://www.statlit.org/pdf/2004RidgwayNicholsonIASE.pdf.

Schield, M. (2001). *Describing Rates and Percentages in Tables.* 2001 Business of Communications Conference. http://web.augsburg.edu/~schield/MiloPapers/2001BusOfComm.pdf

Schield, M. (2004). Statistical literacy curriculum design. *IASE Roundtable Proceedings*. Available at http://www.statlit.org/pdf/2004SchieldIASE.pdf

Schools Council Project on Statistical Education (1980). *Teaching Statistics 11–16*. Slough: Foulsham Educational.

Smith, A. (2004). *Making Mathematics Count: The Report of Professor Adrian Smith's Inquiry into Post-14 Mathematics Education*. London: DFES. http://www.mathsinquiry.org.uk/report/index.html

Tomlinson, M. (2004). *14–19 Curriculum Qualifications Reform: Final Report of the Working Group on 14–19 Reform*. London: DFES. http://www.14-19reform.gov.uk/

UK Data Archive (2004). *Survey Data in Teaching Resource*: *Enhancing Critical Thinking and Data Literacy.* UK Data Archive. http://x4l.data-archive.ac.uk/

Working Party on Performance Monitoring in the Public Services (2005). Performance indicators: good, bad, and ugly. *Journal of the Royal Statistical Society A*, **168**(1), 1–27.

BURSARY IN STATISTICAL EDUCATION

A central aim of the **Teaching Statistics Trust** is to encourage and support excellent classroom practice in statistics education. The Trustees are aware that it is difficult for practising teachers of statistics to find the time to step back from their classroom duties and reflect on wider learning and teaching issues. It is also the case that much good practice at local level fails to be disseminated to the wider community.

To help remedy this we invite applications from practising teachers (in primary or secondary schools) for a bursary to support a project in statistics education. For example, you might wish to:

- carry out a classroom-based investigation,
- produce innovative teaching materials,
- explore better use of ICT in statistics teaching.

The project should relate to statistical learning at the school level (ages up to 19 years). The project must provide an outcome suitable for wider dissemination and therefore applicants will need to demonstrate that their project *will benefit other teachers*.

A sum of up to £3000 is available from the **Teaching Statistics Trust** to cover staff replacement, travel and other expenses. This will enable the successful teacher to be released from teaching duties for regular days during a school year. It may also be used to provide essential equipment, subject to approval. The work paid for by this bursary will be supervised. The Trustees will liaise with the successful applicant to arrange for a suitable supervisor.

For an application form, please write to:

Alan Graham, Centre for Mathematics Education, The Open University, Milton Keynes, MK7 6AA
e-mail: a.t.graham@open.ac.uk