Journal of Statistics Education

Teaching, Learning and Assessing Statistical Problem Solving

John Marriott, Neville Davies and Liz Gibson
Nottingham Trent University, UK

Journal of Statistics Education Volume 17, Number 1 (2009), www.amstat.org/v17n1/marriott.html

---

Key Words: Problem solving; Teaching; Learning; Assessment.

Abstract

In this paper we report the results from a major UK government-funded project, started in 2005, to review statistics and handling data within the school mathematics curriculum for students up to age 16. As a result of a survey of teachers we developed new teaching materials that explicitly use a problem-solving approach for the teaching and learning of statistics through real contexts. We also report the development of a corresponding assessment regime and how this works in the classroom.

Controversially, in September 2006 the UK government announced that coursework was to be dropped for mathematics exams sat by 16-year-olds. A consequence of this decision is that areas of the curriculum previously only assessed via this method will no longer be assessed. These include the stages of design, collection of data, analysis and reporting which are essential components of a statistical investigation. The mechanism outlined here could provide some new and useful ways of coupling new teaching methods with learning and doing assessment – in short, they could go some way towards making up for the educational loss of *not* doing coursework. Also, our findings have implications for teaching, learning and assessing statistics for students of the subject at all ages.

1. Introduction

The state of mathematics education in the United Kingdom has been an issue of concern for some years. Following a commission by the UK government, Smith (2004) published a report into post-14 mathematics provision. His report made wide-ranging recommendations for improving mathematics education in schools in England. Based on one of the recommendations, in 2005 the Qualifications and Curriculum Authority (QCA) commissioned the Royal Statistical Society Centre for Statistical Education (RSSCSE) to review the position of the *teaching* of statistics and handling data (S&HD) in the curriculum. (The three authors of this paper comprised the RSSCSE/QCA review team.) Indeed, Smith had suggested that S&HD education might be improved by teaching it through other subjects, such as science and geography, rather than from its current position in the mathematics curriculum.

In autumn 2005 the RSSCSE/QCA Review carried out a national survey of heads of mathematics, geography and science to determine their views, needs and capabilities within the S&HD area. We were surprised to find that teachers of science and geography appeared more confident than those of mathematics when teaching school students to understand and interact with statistical concepts and ideas - skills that are fundamental for getting students to develop their statistical literacy.

For statutory curriculum reasons at this time it was not possible for the RSSCSE/QCA Review project to recommend the teaching of S&HD *solely* through other subjects and so the Smith suggestion cited in the first paragraph could not be implemented. However, examining the approach to the teaching and learning of S&HD within the science, geography and other curricular it was noted from the survey that the key differences in teaching were in the *application* of S&HD to real contexts and problems arising from within

each subject. This corresponds with the growing body of opinion that has been suggesting that statistics is best taught through problem solving. These research findings led the project team to develop a set of learning and teaching resources for use in mathematics lessons by mathematics teachers that draw on real problems in real contexts. Alongside the teaching materials a new assessment regime was developed that is designed specifically to test the effectiveness of teaching and learning in a problem solving setting. In this paper we describe the strategy we used to produce the resources and assessment. The full report can be obtained by emailing the first named author (john.marriott@ntu.ac.uk).

In the next section we review the way teachers of statistics have gradually proposed changing the way statistics is taught to make it more relevant, with some making the case for using a problem solving approach entirely. As part of the RSSCSE/QCA Review we devised a portfolio of problems through which some topics in statistics can be taught and we report the development of resources to support this approach in section 2. In section 3 we describe one of the resources that we produced in some detail and also briefly describe how they were trialed in schools. In section 4 we report the development of methods for assessing this approach and present an analysis of the results obtained from trialing the new assessment. In section 5 we present some conclusions and suggest that our findings could be used for the development of teaching, learning and assessment of statistics at all levels.

2. Learning and Teaching Through Problem Solving

A simple paradigm for solving problems using statistics is summarised in the English National Curriculum using four activities: specify the problem and plan; collect data from a variety of suitable sources; process and represent the data; and interpret and discuss the results. The activities are cyclic (illustrated in Figure 1) because it may be necessary to refine the initial approach to solving a problem and repeat the process over again. Throughout this paper we refer to this four stage cycle as the (statistical) problem solving approach (PSA). It is also known as 'the handling data cycle'. There are variations of this cycle that are currently being used – for example Wild and Pfannkuch (1999) proposed a five-stage procedure that is closely allied with the cycle illustrated in Figure 1.

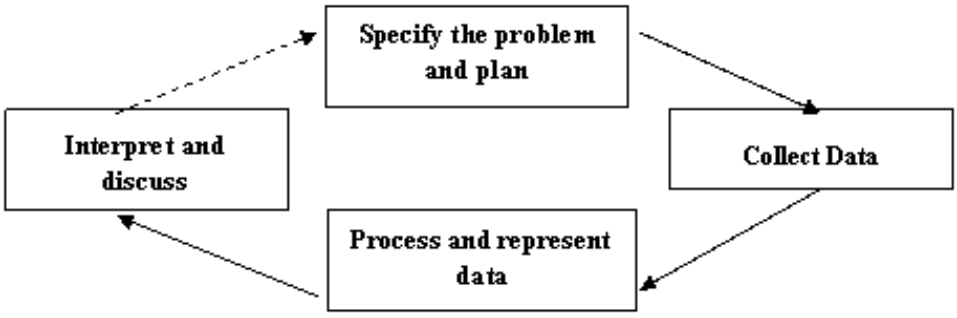

Figure 1. The statistical problem solving approach

It was from the early to mid 1990s that literature began to emerge which was explicit in advocating the use of the PSA (see, for example, Chatfield 1995 and 2002) for the teaching of statistics. Stuart (1995 and 2003) discusses the dominance of mathematical thinking in statistics education and suggests the PSA as being a good counter-measure to this. Garfield (1995) and Garfield and Ben-Zvi (2007) summarise educational research views on statistical learning which suggest that teaching statistics through solving problems is considered to improve students' skills, particularly as they interact with real data, see also Cobb (1992) and Cobb and Moore (1997).

More recently Franklin and Mewborn (2006) reported that the American Statistical Association has endorsed the reports from the Guidelines for Assessment and Instruction in Statistics Education (GAISE) project which advocate the active learning of statistics using real data and a problem solving approach. Rossman et al (2006) found that using the problem solving approach in the teaching of statistics is of great benefit to both teachers and learners, a view also supported by Groth (2006).

The PSA has been included in the English National Curriculum since 2000. However, the survey of heads of mathematics carried out in 2005 by the RSSCSE/QCA Review project suggested that the good intentions implied by this curriculum specification were never really carried through into the taught, learned or assessed versions of the same curriculum. The Review found compelling evidence

that even heads of mathematics departments in secondary schools were not confident about teaching the PSA.

Following the decision to abolish coursework in September 2006, in October the UK curriculum development organisation *Mathematics in Education and Industry* (MEI, 2006) produced an evaluation of the role of this method of assessment in mathematics. It provides evidence that the intended outcome of coursework was not achieved. The full report 'Coursework in Mathematics' is available from their website. It notes that many factors contributed to the negative opinion that teachers had formed of the handling data part of the coursework, with the pressure of time being one of the most important. The RSSCSE/QCA Review found that many teachers did not fully understand the importance of the PSA, or indeed how it worked in practice. In addition to this the demands and nature of assessment needed for the coursework meant that the full cyclical nature of the approach was neither taught nor assessed. It appears that for many teachers the PSA will not have formed a key part of statistics courses they will have studied or attended. The RSSCSE/QCA Review project came to an early conclusion that there was a pressing need to both help teachers develop professionally in this area and also to produce materials for use in the classroom.

Carrying out the problem solving cycle depicted in Figure 1, using real data in real-world problem contexts, requires a number of different cognitive skills. Thus, in devising a teaching, learning materials and assessment regime that can be used to grade students' problem solving skills, these *cognitive* skills need to be identified. Different forms of learning also need to be recognized. Bloom et al. (1956) published a taxonomy of educational objectives which was later revised by Anderson and Krathwol (2001). The six categories of Anderson's revised taxonomy are: remembering; understanding; applying; analysing; evaluating; and creating. These categories are considered to be a hierarchy of skills – although there is some educational debate as to whether they are also progressive.

In considering the development of the teaching, learning and assessment resources that use the PSA we completed a mapping from the handling data specification within the English national curriculum for mathematics onto the categories of the revised taxonomy. Table 1 presents this mapping and shows that *each* stage in the cycle demands the use of at least four levels of the taxonomy. For example, even at the first stage of the PSA, the only category that is not used to any extent is 'evaluating'.

Anderson and Krathwohl (2001) also introduced a second attribute/category which they refer to as the *knowledge* dimension. The categories of this dimension, representing the outcomes of the thinking process, are factual, conceptual, procedural and metacognitive. The classification in Table 1 was used, together with a two way table that combines the cognitive process dimension with the knowledge dimension, to produce a mapping of the learning objectives of the statistical PSA onto the resulting two way classification. This table is not reproduced here, but can be viewed at www.rsscse.org.uk/qca/doc/PSAtwowaymap.pdf. This process then naturally identified our starting point for the development of both the teaching and learning materials and the assessment and its associated questions.

Table 1: Revised Taxonomy and the Problem Solving Approach		
Stage of problem solving approach	**Descriptor (from the QCA specification)**	**Level in taxonomy (from the revised taxonomy)**
1. Specify the problem and plan	formulate questions in terms of the data needed, and consider what inferences can be drawn from the data; decide what data to collect (including sample size and data format) and what statistical analysis is needed	Remember: **recognising; recalling** Understand: **interpreting; exemplifying; explaining** Apply: **executing** Analyse: **differentiating; organising** Create: **planning; producing; generating**
2. Collect data	collect data from a variety of suitable sources, including experiments and surveys, and primary and secondary sources	Remember: **recognising; recalling** Understand: **classifying; comparing** Apply: **executing; implementing** Analyse: **organising** Create: **planning**
3. Process and represent the data	turn the raw data into usable information that gives insight into the problem	Remember: **recognising; recalling** Understand: **interpreting; exemplifying; classifying; summarising** Apply: **executing; implementing** Analyse: **differentiating; organising**
4. Interpret and discuss the data	answer the initial question by drawing conclusions from the data	Remember: **recalling** Understand: **interpreting; exemplifying; summarising; inferring; comparing; explaining** Analyse: **differentiating; organising; attributing** Evaluate: **checking; critiquing** Create: **generating; producing**

From January to September 2006 the RSSCSE/QCA Review project produced eight exemplar resources for teaching statistical topics through the PSA. These were designed, written, trialed and refined using input from practicing teachers in secondary schools in England. The resources are freely available from the web site www.rsscse.org.uk/qca. These materials are designed to support teachers in delivering statistical concepts in a holistic manner. A simple and appealing version of the handling data cycle diagram was used (see Figure 2) which is reinforced and repeated throughout the delivery and presentation of the materials.

These resources differ from what has previously been used in schools in two important respects. The first, and arguably the most important, part of the PSA occurs at the *planning* stage of the cycle. These resources lead students into a detailed discussion of the problem and come to a decision that *data* can assist them in addressing the problem and attempt to seek a solution. The students then decide what data they think would be the most helpful and they discuss and arrive at a decision as to how best these data might be collected. This experience draws the students into the problem and establishes *their* ownership of the procedures that are to follow and, therefore, of the solutions that are eventually formulated. The second important feature of the materials is that there is regular reinforcement of the cycle involved in the PSA with constant reminders of the current stage of the process as the investigation proceeds. These two features serve to establish and reinforce the statistical PSA as a natural and powerful evidence-based and logical approach to solving problems.

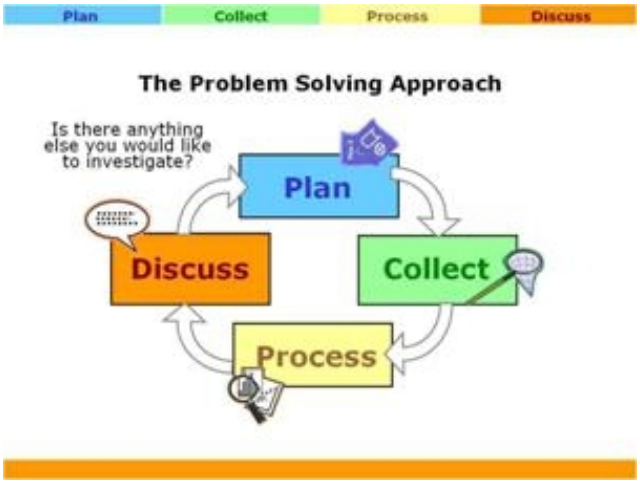

Figure 2. Schematic diagram of the problem solving approach

The teaching materials and supporting documentation are designed for formative and summative assessment. Teachers are provided with notes which they can draw on as much or as little as they need to. These contain suggestions for discussion and so support teachers in their use of questioning for formative assessment. There are also worksheets which allow teachers to get quick insights into learners' understanding throughout the lesson. Summative approaches to assessing the materials enabled us to judge the effectiveness of the materials in terms of the learning in the classroom.

3. New Resources for Teaching and Learning - an Example

In order to fully illustrate the manner in which the PSA is used in the resources we have developed, we describe in detail one of them, called "*How safe is your area?*" On the website for this resource (www.rsscse.org.uk/qca/Resource1.htm) the teacher can find a teacher's overview of the resource, access to both PowerPoint and over head projector (OHP) versions of the resource and a pupil worksheet and pupil questionnaire. We discuss the PowerPoint version in this section and PowerPoint notes that advise the teacher on the use of the presentation are also available to download. These duplicate in large part the notes provided within the PowerPoint file itself.

3.1. The resource

The presentation opens with the slide in Figure 3 and a statement of the lesson objectives

Objectives

Children should learn:

- to relate individual statistical techniques to a wider problem;
- to think analytically about a statistical problem;
- to apply a variety of techniques to solve a problem.

The four stages of the PSA are clearly marked at the top of each slide and are colour coded throughout the presentation to provide visual cues to the pupils. Additional visual cues are provided by associating different icons with each stage in the cycle. During any stage this icon is displayed together with an enlarged colour coded box to allow pupils to associate their current activities with the stage of the cycle they are currently working on. These points are illustrated for the PLAN stage in Figure 3 and Figure 4.

Figure 3 Introductory slide in "*How safe is your area?"*

Figure 4 Initial slide for the PLAN stage

As the first slide for each stage of the PSA appears, the associated notes for the teacher show the objectives. For the PLAN stage these are

Objectives

Children should learn:

- the context of the problem;
- to ask questions about presented information;
- to hypothesise.

The teacher is then advised to begin with a quick discussion of the media and reporting of crime – some example questions are provided:

- What was the most recent crime that made the headlines?
- What kinds of crimes tend to make the headlines?
- Why do they make headlines?

These questions are presented on a slide that is then followed by slides that enable pupils to learn the context of the problem by discussing questions that relate to the incidence of crime in different regions, the perceived relative 'importance' of different types of crime and how they are recorded by the authorities.

Figure 5 Different approaches to the problem

The class could address the "How safe is your area?" problem by investigating the perception of crime. This will require them to conduct their own survey (primary data), or they could study actual crime figures (secondary data) from government websites to answer the same question (with less processing of data at the initial stage). Once the appropriate choice has been made (see Figure 5)

the PowerPoint presentation branches off down one of two different routes of investigation. At the end of the first route the PowerPoint presentation offers the second alternative route as an extension activity. If neither button is selected, the PowerPoint presentation will, by default, follow the secondary data route. Here we follow the primary data route as an illustration. Figure 6 shows the first slide for this route.

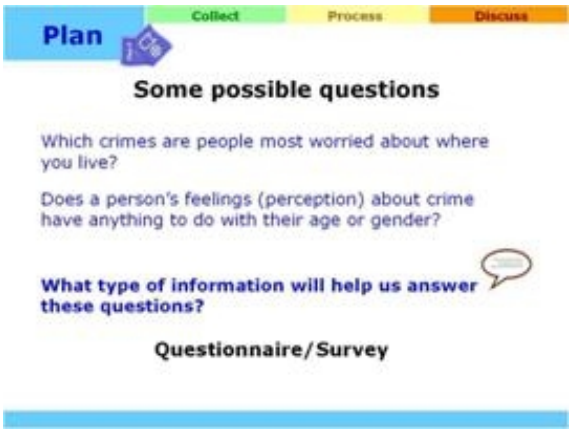

Figure 6 Formulation of the PLAN

Figure 7 Consideration of data sources

The final part of the PLAN stage can now take place in which the pupils discuss how they can collect suitable data. Clearly there is an overlap between the PLAN and COLLECT stages of the cycle and this is emphasized by highlighting both at this point of the presentation as is shown in Figure 7.

At the transition point between stages of the cycle the use of the PSA is emphasized by the use of slides that remind the pupils where they currently are in the cycle. Figure 8 shows the first of these for this example.

Figure 8 The use of the PSA being reinforced

The fifth phase (2004 – 2005 UK academic year) of the *CensusAtSchool* project (www.censusatschool.org) contained four questions on pupils' perception of crime. Pupils were asked "how worried are you about…": having property stolen; being mugged; being attacked; and being insulted. The range of responses was limited to 'Very', 'Fairly' 'Not very', and 'Not at all'. These alternatives are the same as in a recent European survey of crime. The RSSCSE/QCA website contains random samples of data from learners' responses from each region of the UK, each sample containing anonymous answers for 12-16 yr olds. PDF versions of these data can be downloaded from the right hand side of the "How Safe Is Your Area?" webpage using drop-down menus.

How Safe Is Your Area?

Age ☐ Gender ☐

How worried are you of being a victim of...

	Not at all	Not very	Fairly	Very
Stolen	o	o	o	o
Attacked	o	o	o	o
Mugged	o	o	o	o
Insulted	o	o	o	o
	o	o	o	o

Figure 9 The questionnaire

Figure 10 The summary sheet

Alternatively, pupils could be encouraged to record their own perceptions and Figure 9 and Figure 10 show the pupil-centered resources that are available for the collection of primary data for this example. Figure 9 shows the single question questionnaire, these are also available from the website and can be distributed around the class. Figure 10 shows the summary sheet and each pupil is encouraged to enter their data as a line in this sheet. The resulting sheets can then be processed.

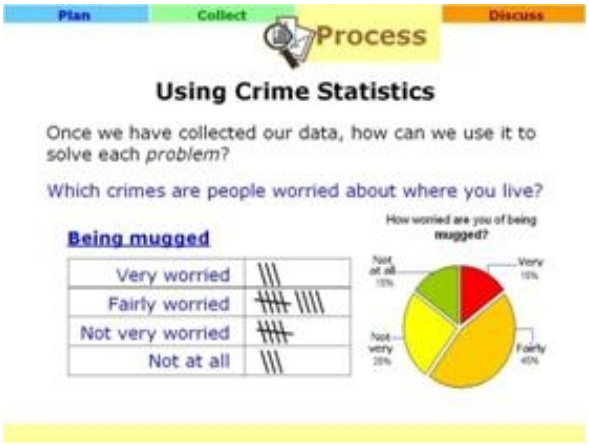

Figure 11 Processing the data

At the PROCESS stage of the PSA, illustrated in Figure 11, the pupils are prompted to discuss how the data can be assembled into a table. A discussion of whether the use of percentages or totals is more useful is then followed by the associated calculations. The discussion then turns to how data might be presented graphically.

The final, DISCUSS, stage can start with a look at some of the results from a random sample taken from the *CensusAtSchool* web site and how the responses there compare with the perceptions from the class. This discussion can then be extended to other questions from the *CensusAtSchool* survey. The slide shown in Figure 12 illustrates a discussion around the perceptions of two different crimes and Figure 13 shows questions that return the discussion to the original PLAN stage and help to summarise the pupils findings.

Figure 12 Comparisons

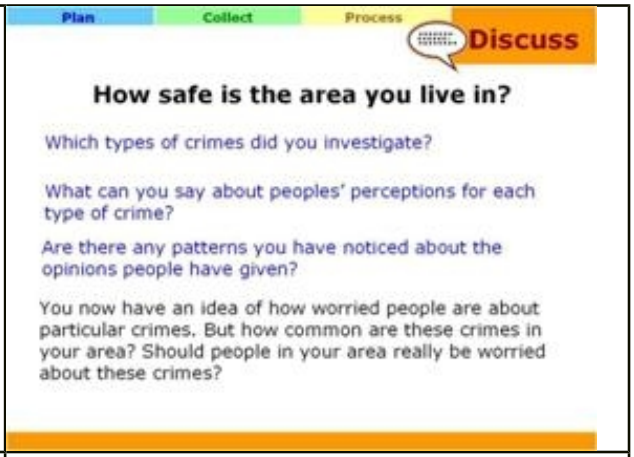

Figure 13 Discussion summary

3.2. Feedback from teachers

The resources developed for the RSSCSE/QCA project were trialed in 43 schools from a random sample selected by the QCA. The teachers using the resources were asked to provide detailed feedback using an online form in which they were asked to evaluate the effectiveness of the material from their own experience. The feedback form contained 26 statements and the teachers used a sliding scale (from 0 for disagree to 100 for agree) to express their views. The teachers who responded in this way expressed strong support for the use of both the PSA and materials of the type we provided. In addition over 75% of the teachers providing online feedback indicated that they would use the resources with older pupils. The teachers were also asked for open-text written feedback and we provide here a selection of the comments from teachers who trialed the resources.

> "I didn't use the pupil worksheet as I found it a bit too wordy for my students to cope with. Instead, we covered the main points on the worksheet orally in our discussions. The PowerPoint was a fantastic resource which I will definitely use again and I liked the way that lots of different techniques were employed in looking at one problem."
>
> "I have spent 3 sessions on the 'how safe is your area?' task with my year 8 group (mixed ability from level 3-7) and they are getting on brilliantly. It has worked really well as we are now in the last week of term and have had time to put their work into a display which hopefully lots of other students will see. Lots of useful discussions and some surprise that Derby seems to be worse than the national average on just about every crime!"
>
> "From a teachers point of view I think that this worked well, the first two lessons are tight on time if you want pupils to use what they have learnt for their own data. We did not do the back of the sheet as this was covered by what they did in the computer room. They enjoyed working in groups and discussing how the different steps should be carried out. By setting it within the school it made it more relevant to them and gave them a chance to have there say."
>
> "The students were engaged from the introduction. It led to a very purposeful whole class discussion about crime and safety. Having computers available was definitely more useful than handing out data sheets for our area as the students felt more in control. Excel for charts proved effective and working in small groups to produce a display was a good end to the project. The best bit from a teaching point of view was that I saw the students interpreting their results which they normally struggle with and this is definitely a key part of Statistics GCSE coursework. I spent 4 sessions (about 3 hours) which worked well."

The resources were amended on the basis of the detailed comments provided by classroom teachers and referees.

4. Developing a New Assessment Regime

A regime change in teaching and learning leads to the need for a regime change in the corresponding assessment. For example, one that grades students' knowledge and skills in problem solving could not just rely on asking, for example, questions about the mechanics of calculating summary statistics, or the ability to draw bar charts and/or histograms. The regime needs to focus on the thinking behind each stage of the process, the links between the stages and an ability to entertain the fact that there may be more than one solution to a statistical problem. In this section we describe the way in which we developed an assessment regime taking into

account these key points and the thinking needed to learn through statistical problem solving.

After considering various options for the PSA we decided to develop an *online* assessment regime which involves students setting themselves up as advisors and critiquing the work of others. Having recognized that high level skills are demanded in applying the PSA (see Table 1) it was felt that assessment needed to occur in a supported environment, particularly when considering that the initial design of the learning materials was for use with 12 year old students. However, we now believe that older students, for example at UK A level study (learners up to the age of about 18) and at first level at university, will benefit from the approach we have adopted, both for learning from being taught through the PSA and being assessed on the approach.

We were also confronted with the need to create a time-constrained assessment and were concerned that it should be accessible to students with a wide range of ability. By necessity, S&HD has a high level of literary demand and by using an online tool features such as audio commentary and zooming can be used to make the assessment as accessible as possible. Similarly, the choice of a familiar context for the problem at the core of the assessment puts students at ease and helps to generate genuine interest in the problem they are trying to solve. A computer-based tool has the added advantage that large sets of data can be used where appropriate, and are easily stored on a computer database. We believe that the assessment environment not only draws out students' understanding in these areas but also presents S&HD in a manner similar to the way it is used in practice by professional statisticians. Finally there is a gradual move towards the use of online assessment from within the QCA, which was also one of the recommendations by Smith (2004).

We chose as a context for the assessment a scenario associated with the purchase of a new mobile phone: *Getting the best deal*. We felt that it was important that the familiar context should be easily accessible and should also have the potential to be supported with audio files, allowing the reading aloud of text within the assessment.

The assessment, which comprises three sections, A, B and C, was designed to take one hour to complete. The questions used are described in detail in Table 2 and we now summarise the main results of the online assessment.

Section A is designed to test the candidates' knowledge of the PSA and presents them with a modified version of Figure 1 (see Figure 14). Candidates are presented with four descriptions of the activities undertaken during the different stages of the cycle and are asked to drag and drop the descriptions into the correct locations. The next question presents them with the cycle and asks them to explain in their own words why it is described as a cycle and why this description is important.

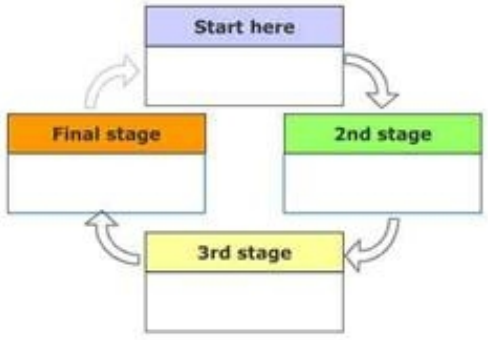

Figure 14 Template for the first task in the assessment

Section B starts by describing a case in which a student, 'Ayesha', has used the problem solving approach to decide which mobile phone would be the best value for money. Ten statements summarising the process are presented and the candidates are asked to read them and to drag and drop each into its correct 'box' of the cycle diagram (Figure 14). This is followed by eight further questions that carefully probe the candidates' understanding of what Ayesha did at each of the different stages, and the extent to which she was correct in her work.

Section C describes another child, 'Andy', who also wants to buy a mobile phone but who has not completed the problem solving cycle. Instead the candidates are led through eight slides in which available information is presented and the candidates are invited to

help make some decisions. After some of these decisions the candidates are told what was *actually* done but they cannot return to previous pages to change their responses. The next stage presents the collected data and encourages students to interact with them. The candidates are asked to undertake calculations, make comments on the data and to consider suitable graphical presentations and summary measures. They must then type in responses as part of a discussion of what has been found and finally make a decision as to which phone should be purchased. The last three questions explore the fact that the data used by Andy was not all that was available and leads them to comment on how the problem solving cycle could be revisited if different questions are posed.

Table 2: Questions in the on line assessment 'Getting the Best Deal'

Question	**Task**
A. Overview of the problem solving approach	
A0	The descriptions of the stages of the cycle must be dragged and dropped into the correct position.
A1	Explain why the process is called a cycle and why that description is important.
B. Ayesha's investigation Ayesha has been told she now has to pay for her own mobile phone. She wants to use what she has learned about the problem solving approach from her math teacher to achieve this. She has written up her investigation in rough using ten statements.	
B0	The ten statements, describing Ayesha's investigation, must be dragged and dropped into their correct positions in the cycle.
With reference to the **'plan'** statements in Ayesha's investigation	
B1	Look at the four statements and then indicate whether someone else could read this plan and then carry it out for themselves (Yes, No or Don't know)†.
B2	If they answer 'no' to B1, describe what other information they would need (by typing into a text box)‡.
B3	Indicate which of the statements states the problem Ayesha is trying to solve ('Don't know' is one possible answer).
B4	Suggest another method Ayesha could use to choose a day at random (text box).
Question	**Task**
With reference to the **'collect'** statements in Ayesha's investigation	
B5	What is the mean length of Ayesha's phone calls (text box)?
B6	Comment on her choice of shop (text box).
B7	Would collecting more information on different deals help? Explain your answer (text box).
With reference to the **'process'** statements in Ayesha's investigation	
B8	The process statements used by Ayesha are presented showing her calculations. Why did Ayesha use the number 30 in the calculations (text box)?
B9	Do you think it is reasonable to expect she'll use her phone the same amount each month (Yes, No or Don't know)? Explain your answer (text box).
B10	Are there other things you think she should do with the data (text box)?
With reference to the **'discuss'** statements in Ayesha's investigation	
B11	Would you advise Ayesha to change to the deal given that the calculations are correct (Yes, No or Don't know)? Why (text box)?
B12	Ayesha thinks that £56 is a lot of money and 840 minutes is a long time. It is based on data from one day. What might have happened on this day for it to be higher than normal (text box)?
B13	What could Ayesha do to check if her use on this day is similar to other days or not (text box)?
C. Andy's investigation Andy considers advertisements for two phones that only differ in the amount of time and number of texts that are included in the package. The two adverts are displayed.	

With reference to the **'plan'** stage of Andy's investigation	
C1	Andy's friends say that the deals are the same. Do you agree (Yes, No or Don't know)?
C2	Can you give an example of when one deal might be better than the other (text box)?
Andy decides to investigate how much he uses his phone to see which is the better deal. The candidates' task is to help him with his investigation. (After the candidates have committed to an answer they cannot return to it later in the light of what Andy actually decided. They are told that their responses will be judged on the basis of whether their reasons and explanations are correct, not on whether they choose the same course of action as Andy.)	
C3	What data do you think he should collect (text box)?
C4	Over what number of days, weeks or months should he collect data? Explain your choice (text box).
C5	How could he collect the data (text box)?
A handwritten statement of Andy's problem and plan are then presented for consideration.	
C6	Why does he decide to choose a bill at random (text box)?
C7	Why are his mobile phone bills a good way to collect the data (text box)?
With reference to the **'collect'** stage of Andy's investigation	
A table of the data that Andy decided to collect (number of calls, total call duration and number of texts for each of 31 days).	
C8	The call durations are in seconds and look messy. Suggest how Andy could make this better (text box).
With reference to the **'process'** stage of Andy's investigation	
He decides to change the seconds to minutes and round them to the nearest half minute	
C9	Andy has done the first few, can you do the rest below? (Four numerical data boxes are there for the answers to be entered and an on screen calculator, together with instructions for its use, is also available.)

Question	**Task**
C10	One of the rows in the table doesn't look right to Andy (the row is presented). Can you spot why (text box)?
C11	Can you explain why it's happened (text box)?
C12	Is it a problem (text box)?
C13	Complete the frequency table for the number of calls made each day for the month. (Numerical data boxes labeled 0, 1, …, 12 are there for the answers to be entered.)
C14i	Look at the chart below (line graph of daily call duration). Mark whether you think it is helpful or unhelpful (or not sure) in deciding which package to use.
C14ii	Look at the chart below (bar chart for number of calls per day). Mark whether you think it is helpful or unhelpful (or not sure) in deciding which package to use.
C14iii	Look at the chart below (pie chart for number of calls per day). Mark whether you think it is helpful or unhelpful (or not sure) in deciding which package to use.
C14iv	Look at the chart below (scatter plot of number of calls vs number of texts). Mark whether you think it is helpful or unhelpful (or not sure) in deciding which package to use.
C15	Candidates are asked to undertake calculations for the number of calls made daily. (Numerical data boxes labeled mean, median, mode and range are there for the answers to be entered.) On screen calculator is provided.
With reference to the **'discuss'** stage of Andy's investigation	

C16	What does the mode tell you about how Andy uses his phone (text box)?
C17	How long does he spend on the phone in the month in all (text box)? Candidates can use the data and charts on previous screens.
C18	Which package do you think he should choose and why (text box)? Superimposed line graphs of number of texts and number of calls for a different month are presented. An anomaly is highlighted.
C19	Can you give a possible explanation for what happened here (text box)?
Referring back to the **'plan'** stage of Andy's investigation	
Andy had noticed that lots of phone deals seemed to suggest that the more texts you send the fewer calls you make. Two sample deals are presented.	
C20	Write a plan for an investigation to look at whether it is true that **the more texts a person sends, the fewer phone calls they make.**

†Where 'Yes, No or Don't know' is indicated there are three check boxes.

‡Where 'text box' is indicated the candidates are invited to "Type your answers here".

Our approach to designing a grading scheme and then allocating marks for the assessment involved adapting the delightfully simple grading scheme for UK A level coursework in statistics provided by the MEI. This allocates the assessment questions to domains for grading and uses a very simple mark allocation scheme. In our regime we use five domains: the first to allow for the holistic view of the problem solving approach and the remaining four to correspond to the four different stages of the problem solving cycle in Figure 1. Following suggestions by Garfield (1994, example 2) we allocated marks that correspond to the responses candidates make to each question being incorrect (0 marks), partially correct (1 mark) or correct (2 marks). Each of the five domains could also be given different weightings if an examiner so wishes. The resulting mark allocation scheme is shown in Table 3, where we have indicated a drag and drop question by 'D&D'.

Table 3: Assessment sheet for *Getting the Best Deal*

Domain	Mark	Question	Description	Comment	Mark
Holistic view	0, 1, 2	A D&D	Places descriptions in correct locations.		
	0, 1, 2	A1	Clear statement of what cycle means. and why it is important.		
	0, 1, 2	B D&D	Places statements into correct stages. of the cycle.		
Ayesha					
Plan	0, 1, 2	B1 & B2	Gives a clear justification for choice of response to B1.		
	0, 1, 2	B3	Correctly identifies the statement that is the problem to be solved.		
	0, 1, 2	B4	Gives a clear description of an alternative method that could be used to choose a day at random.		
Collect	0, 1, 2	B5	Correctly calculates the mean.		
	0, 1, 2	B6	Clearly comments on the choice made for the shop.		
	0, 1, 2	B7	Presents clear discussion of whether collecting more information would help.		

Process	0, 1, 2	B8	Clearly states why the number 30 was used.		
	0, 1, 2	B9	Gives a clear justification for choice of response to whether phone use could be expected to be the same in other months.		
	0, 1, 2	B10	Clear statement of other uses that should be made of the data.		
Discuss	0, 1, 2	B11	Gives a clear justification for choice of advice on whether deal should be accepted.		
	0, 1, 2	B12	Gives a clear statement of possible reasons for data to be an outlier.		
	0, 1, 2	B13	Makes a clear statement of how data should be checked.		
Andy					
Plan	0, 1, 2	C1	Selects the correct answer.		
	0, 1, 2	C2	Clearly states an example of how one of the deals could be better.		
	0, 1, 2	C3	Clearly states which variables should be measured.		
	0, 1, 2	C4	Gives clear explanation of the choice of time period over which data should be collected.		
	0, 1, 2	C5	Clearly describes how the data could be collected.		
	0, 1, 2	C6	Gives a clear explanation of why a bill is chosen at random.		
	0, 1, 2	C7	Gives a clear explanation of why phone bills are a good data source.		

Domain	**Mark**	**Question**	**Description**	**Comment**	**Mark**
Collect	0, 1, 2	C8	Makes a clear suggestion of how data tabulation can be made less messy.		
Process	0, 1, 2	C9	Correctly calculates call durations.		
	0, 1, 2	C10	Correctly identifies mistake in data.		
	0, 1, 2	C11	Gives clear explanation of why mistake is there.		
	0, 1, 2	C12	Gives clear statement of whether this poses a problem.		
	0, 1, 2	C13	Correctly completes the frequency table.		
	0, 1, 2	C14i	Makes a sensible response to question about time series plot.		
	0, 1, 2	C14ii	Makes a sensible response to question about bar chart.		
	0, 1, 2	C14iii	Makes a sensible response to question about pie chart.		
	0, 1, 2	C14iv	Makes a sensible response to question about scatter plot.		
	0, 1, 2	C15	Correctly calculates summary statistics.		

Discuss	0, 1, 2	C16	Gives clear interpretation of the mode.		
	0, 1, 2	C17	Gives correct interpretation of results from frequency table.		
	0, 1, 2	C18	Clearly states preferred choice of package and clearly explains reasons.		
	0, 1, 2	C19	Gives a clear interpretation of the new graphical information.		
Plan	0, 1, 2	C20	Constructs a clear plan for a further investigation.		
A mark of 0, 1 or 2 is allocated for each question asked of the candidates. These correspond to the responses being incorrect (0), partially correct (1) or correct (2).					

The assessment was given to a range of students who had experienced the problem solving teaching/learning resources developed at the material development stage of the RSSCSE/QCA review. The assessment materials are intended to test the students' ability to approach statistics in a holistic manner.

The assessment was well received by the classes who trialed it and teachers could choose to use correct answers in their teaching after the assessment had taken place. In addition to the factors that led us to choose an online approach to the assessment, we feel that this form of online assessment has several educational advantages over paper-based equivalents. We also believe that the assessment regime developed for this project has much educational potential for students of *all* ages and we are currently developing it for first-level non-specialist university students.

Garfield (1994) stresses the need for assessments that measure the understanding of a PSA that can also be viewed as an integral part of the teaching and learning process. The assessment we have produced is a prototype of a tool that could be used to judge the effectiveness of teaching materials at different levels. The structure and approach we have adopted is appropriate for a variety of ages and ability of students. The template could be adapted for a wide range of individual needs. It is our belief that there is potential for further developing the assessment and providing feedback in the following ways.

1. The assessment in its current form allows teachers to give formative feedback by providing responses for the whole group in a spreadsheet. The responses could be combined to produce a completed exam paper for each student; this could also be accessed on screen. Similarly, the teacher could select a particular response to a question and examine the spread of responses allowing students to become accustomed to the idea that there is often no single correct response to some questions in statistics.
2. The potential exists for feedback in the form of a checklist of skills with a guide as to how students have performed against each skill. This could work in partnership with a portfolio of assignments to demonstrate each skill.

In order to understand how this proposed approach to assessment was received by the pupils we approached teachers who, of necessity, had trialed the materials with their classes. Three of these teachers agreed to use the assessment with their pupils after delivering a class to review the PSA. A total of 58 pupils took the online assessment and submitted their responses. The pupils tested fell into two year groups, year groups 8 and 9, and they were all given the test during normal class time. Table 4 gives summary statistics for the scores achieved by all of the pupils within each of the domains.

Table 4: Test scores by domain

Domain	Minimum	Q1	Median	Q3	Maximum
Holistic	0.00	50.00	50.00	83.33	100.00
Plan	0.00	27.27	45.45	59.09	90.91
Collect	25.00	50.00	68.75	87.50	100.00
Process	3.85	11.54	19.23	46.15	88.46
Discuss	0.00	12.50	21.43	42.86	71.43

It is clear from the table of descriptive statistics above that a wide range of marks can be attained in each of the domains. Note in particular that there is relatively poor performance overall in the 'Process' and 'Discuss' group of questions. An examination of box

plots for each year group indicated that the observed poor performance could be due to the two year 8 groups performing less well than the year 9 group for these questions and also, although a lesser extent, in the 'Plan' questions. The performance at the 'Collect' stage is similar for all groups, regardless of age, and all three groups appear to have grasped the overall idea ('Holistic' domain) of the problem solving cycle to a similar extent.

In order to investigate this further we undertook an analysis of variance for the scores from all the questions (expressed as percentages) with domain and year group as main factors. Both of the main effects proved to be highly significant as was the interaction. The interaction plot is reproduced in Figure 15 and clearly shows the interaction to be between the class (D and P are the year 8 groups) and the 'Process' and 'Discuss' domains. These results are in line with what might be expected in that the more complicated tasks proved more difficult for the younger pupils. The time available for the test also appears to have been a factor in pupil performance as less than half of the pupils attempted questions C13 to C15 and questions C17 to C20. What is encouraging, however, is the apparent success in enabling pupils to understand the PSA and the relatively sound performance in both the 'Plan' and 'Collect' domains for all pupils.

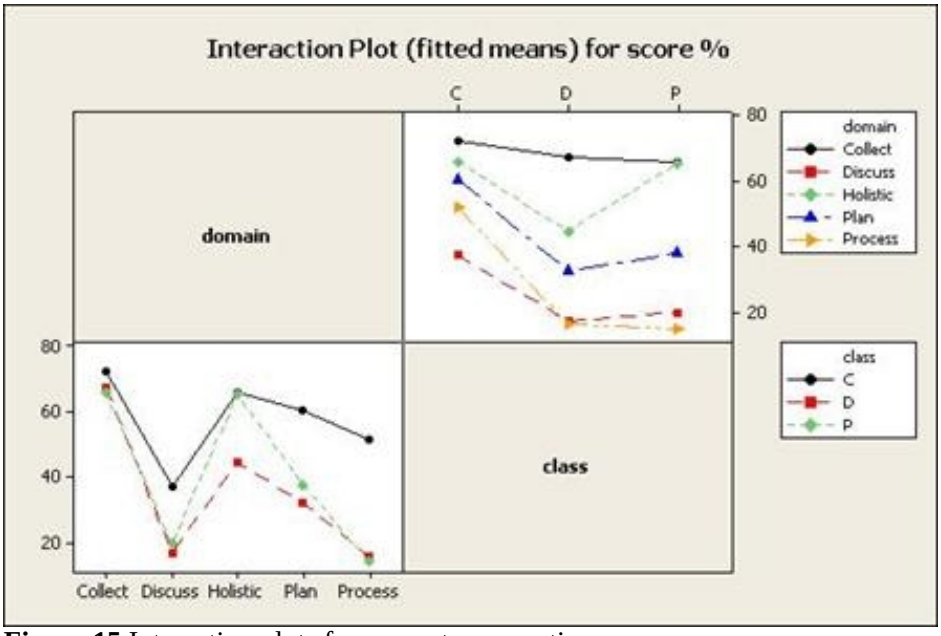

Figure 15 Interaction plots for percentage question scores

5. Conclusions

In September 2006 the UK Department for Education and Skills made a policy decision to abolish all coursework in the national mathematics examinations sat by 16-year olds in England. In their October 2006 report (MEI, 2006) the MEI observed that

> "Effective teachers use a variety of methods to encourage interest and understanding but the current high-stakes results culture in (British) education encourages less confident teachers to 'teach to the test'."

The abolition of coursework for handling data in particular means there is a danger that there are now key areas of the English mathematics curriculum that will not be assessed and, more seriously, therefore may not be taught. For this reason it is important to have replacements both for the material that *had* to be learned in order to complete the formerly compulsory handling data coursework, and for the assessment of that material.

In this paper we have presented materials that could be those replacements: a methodology for teaching and learning through a problem solving approach and a new assessment regime for grading learners after being taught statistics through that approach. Academic and professional statisticians are increasingly arguing for such an approach to be adopted in teaching at all levels: if this is done then the assessment methods used need to match the new way of teaching and learning. As problem solving involves a range of different levels of cognitive skills, the actual questions posed to students within the assessment need to be different and should take these skills into account.

Students of *all* ages benefit from skills acquired through problem solving. The template for teaching, learning and assessment that we have produced we shall be developing for use with older students, for example undergraduates who study introductory level courses at university. At this educational level, the problems used in teaching and learning are likely to be more complex and the questions used for assessment within each domain may depend upon more complicated scenarios. For example, the methods and techniques used in the 'Process' domain of the assessment may well refer to cutting-edge statistics, and the questions posed would need to reflect this. Similarly the 'Discuss' domain may contain, for example, questions that reflect decisions about the efficacy of large drug trials. The template we have produced will handle all these scenarios and we will be reporting their implementation elsewhere.

References

Anderson, L. W., Krathwohl, D. W. (2001), *A Taxonomy For Learning, Teaching, and Assessing: A Revision of Bloom's Educational Objectives*, New York Longman.

Bloom, B. S., Englehart, M., Furst, E., Hill, W. and Krathwohl, D R. (1956), *Taxonomy of Educational Objectives: The Classification of Educational Goals, by a committee of college and university examiners. Handbook I: Cognitive Domain.* New York, Longmans, Green.

Chatfield, C. (1995), *Problem Solving. A Statistician's Guide*. (2nd Edition). London: Chapman and Hall.

Chatfield, C. (2002), "Confessions of a pragmatic statistician," *The Statistician*, 51, 1 – 20.

Cobb, G. W. (1992), "Teaching Statistics," in: *Heeding the Call for Change*, ed. L. A. Steen, MAA Notes No. 22, Washington: Mathematical Association of American, 3-34.

Cobb, G. W., and Moore, D. S. (1997), "Mathematics, Statistics and Teaching," American Mathematical Monthly, 104, 801 – 823.

Franklin, C.A. and Mewborn, D.S. (2006), "The Statistical Education of Grades Pre-k-12 Teachers. A shared responsibility," *Thinking and Reasoning with Data and Chance*, 68th yearbook of the National Council of Teachers of Mathematics.

Garfield, J. (1994), "Beyond Testing and Grading: Using Assessment To Improve Student Learning," *Journal of Statistics Education [Online], 2(1),* *http://www.amstat.org/v2n1/garfield.html.*

Garfield, J. (1995), "How students learn statistics," *International Statistical Review*, 63, 25-34.

Garfield, J. and Ben-Zvi, D (2007), "How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics," *International Statistical Review*, 75, 372 – 396.

Groth, R. E. (2006), "Engaging Students in Authentic Data Analysis," In: *Thinking and Reasoning with Data and Chance* (Eds, G F Burrill and P C Elliott). National Council of Teachers of Mathematics Yearbook.

MEI (2006), "Coursework in Mathematics, A discussion paper," *Mathematics in Education and Industry October 2006*, www.mei.org.uk/files/pdf/CourseworkMEI.pdf.

Rossman, A., Medina, E., and Chance, B. (2006), "A Post-Calculus Introduction to Statistics for Future Secondary Teachers," *Proceedings of the 7th International Conference on Teaching Statistics (ICOTS7).* International Statistical Institute, Voorburg, The Netherlands.

Smith, A. (2004), *Making mathematics count: the report of professor Adrian Smith's inquiry into post-14 mathematics education*, Department for Education and Skills Publications (2004) (DfES Publication No MMC, DfES/0546/2004).

Stuart, M. (1995), "Changing the teaching of statistics," *The Statistician*, 44, 45-54.

Stuart, M. (2003), *An Introduction to Statistical Analysis for Business and Industry – a Problem Solving Approach*, London: Hodder Arnold.

Wild, C. and Pfannkuch, M. (1999), "Statistical thinking in empirical enquiry," *International Statistical Review*, 67, 223 – 265.

John Marriott
Royal Statistical Society Centre for Statistical Education
School of Science and Technology
Clifton Campus
Nottingham Trent University
Nottingham NG11 8NS, UK
John.Marriott@rsscse.org.uk

Neville Davies
Royal Statistical Society Centre for Statistical Education
School of Science and Technology
Clifton Campus
Nottingham Trent University
Nottingham NG11 8NS, UK
Neville.Davies@rsscse.org.uk

Liz Gibson
Royal Statistical Society Centre for Statistical Education
School of Science and Technology
Clifton Campus
Nottingham Trent University
Nottingham NG11 8NS, UK
Elizabeth.Gibson1@rsscse.org.uk
