Who wants to be… The Use of a Personal Response System in Statistics Teaching

Dr Ernst Wit
University of Glasgow

ernst@stats.gla.ac.uk

Service courses of statistics can be among the most recalcitrant. Undergraduate students do not always see immediately the relevance of the course to their own field so that interaction with them tends to be difficult. Add on top of that the large class size, and interactive teaching may seem impossible. The development of handsets as used in *Who wants to be a millionaire?* has proven to be a possible tool to enhance interaction and stimulate learning. In this article we describe this personal response system (PRS) and its implementation within a statistics service course to first year psychology students.

Keywords: personal response system, statistics teaching, handsets, service courses, multiple-choice questions, interactive teaching

In the popular TV programme *Who wants to be a millionaire?* the candidates can ask the audience to help them with difficult questions. The audience picks up its handsets and votes for several possible answers. The votes are anonymous and show up as a bar chart on the screen with percentages for the different answers. Essentially it is this technology and the implementation thereof in a statistics course taught to non-statistics undergraduate students that is central to this article.

It is not the first time that technological means have been proposed to aid education. However, review studies have shown that any evidence of improvement is tentative at best or plainly absent at the worst [2]. Especially when technology is introduced to replace existing non-technological activities, benefits are often not clear at all. Introducing technology in the classroom just for the sake of it does not necessarily help the learner and can be sheer folly. Although this may seem obvious, the current technology push, stimulated by industry and government alike, give computer-aided instruction a certain glamour that is hard to resist by educators.

When we decided to introduce the *personal response system* (PRS) into the classroom, it was therefore in response to a specific weakness in the then current delivery. Identifying problems and evaluating the system after the glitter of novelty has worn off is an essential aspect of the introduction. It is important to assess whether the handset system meets basic didactic standards before celebrating its inauguration into the statistics curriculum.

Draper et al [3] have described general educational principles of the PRS and how it might be able to fulfil successfully a niche requirement in certain types of classes. In this article we draw heavily on their theoretical foundations. Their practical and technical assistance, which they have communicated to us personally, has been invaluable in implementing the system in our curriculum. Whereas the system has been used in Psychology, Computing Science, Philosophy and Biology, this is the first time the system has been implemented in the mathematical sciences. Each discipline brings with it specific requirements and peculiarities, which can only be judged and resolved in practice. Evaluation of the PRS is in progress, and therefore results are provisional but encouraging.

2. Justification and Methods

2.1. A case study: "Statistics for Psychologists"

An integral part of the teaching by many statistics departments is the service teaching for other departments. The statistics department of the University of Glasgow has provided statistics instruction for students in the life sciences, chemistry, nursing and others. The largest of the service courses is the class taught to 200-250 first year undergraduate psychology students. Throughout one academic year the students receive four hours of statistics instruction per week together with a weekly two-hour computer lab in which the lecture material is put in a practical context. The computer labs serve to bridge a gap between the theoretical lectures and ultimate use the students make of statistics within their psychology studies. It focuses on translating scientific questions into statistical hypotheses, exploring data subjectively via tables and plots and interpretation of computer output when performing a formal

analysis. Techniques discussed in the course are wide-ranging and include, among others, design of experiments, one-sample and two-sample inference, ANOVA, simple and multiple regression, simple probability models, categorical data analysis and factor analysis. A team of four staff members teaches the course.

Whereas this performative approach to statistics is appreciated by students and the host department alike, the statistics department feels that it is also essential for students to acquire some theoretical comprehension of the central concepts of the course. For this reason, in addition to the lectures and the computer labs, it devised a one-hour mass tutorial every two weeks. This tutorial was also meant to give students a chance to express what they had not understood very well in the past two weeks and to ask additional questions, for instance with respect to issues of formal assessment.

Whereas the course, both in terms of formal lectures and computer practicals, has been a sustained success over almost ten years by all available standards - such as student evaluations at the end of the course as well as the perceived impact of the course after several years by these same students - the tutorial system has had a difficult and unimpressive record. After a few weeks into the year, attendance rates would go down to a mere 10-15% and during the tutorials students seemed too shy to ask questions or respond to questions that lecturers posed them.

Both the lack of attendance and interaction required a rethinking of the structure of the tutorials. There was the option of dropping them altogether. However, the fact that our current delivery showed weaknesses did not necessarily mean that the tutorials did not have any educational value. It was at this stage that we considered implementing a radically different approach. In the University of Glasgow and the University of Strathclyde in Glasgow a personal response system has been used in several types of classes at different levels. The equipment and the know-how for using it were locally available, which simplified the implementation. Moreover, the system is not complicated, robust in practical sessions and requires little extra work on behalf of the teaching staff beyond setting up the receiver and handing out the handsets to the students.

2.2. Technical details of the personal response system

Several handset systems have been developed and the technology is advancing rapidly. At the moment, an infrared system with TV remote control-like transmitters and a little box-like receiver is by far the cheapest. One receiver is sufficient for some 50 handsets. Several receivers can be linked serially to deal with larger classes. The equipment for 150 students is a laptop, three receivers and 150 handsets and can be carried by two people, assuming that a data-projector is already present in the room. With a bit of practice it takes some 5-10 minutes to set up the system and hand out all the handsets. One classroom in the University of Glasgow has the system permanently installed and several courses have given students their own private handset. This would be clearly preferable and make routine implementation of the PRS within normal lectures quite trivial. Unfortunately, this room was not available this year at the scheduled tutorial times.

Figure 1 shows a schematic of the Personal Response System. The receivers are connected through the serial port to a laptop computer with the appropriate software. The handsets contain buttons corresponding to the numbers from 0 through 9 and each handset is identified by a unique ID number. When a button is pressed on the handset, the signal is registered by the receiver, the software records the corresponding number and displays the handset ID number to indicate to the users that their vote has been registered.

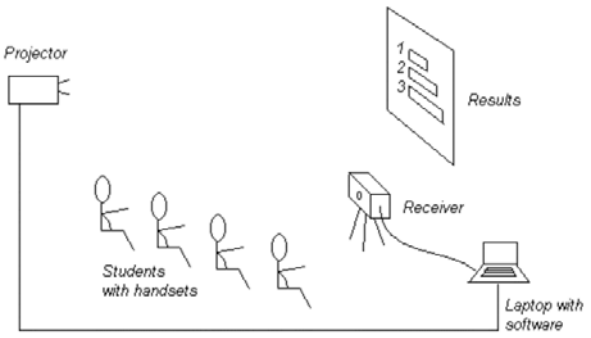


Fig 1 Schematic of the infrared Personal Response System

The multiple-choice questions are displayed on a separate screen by means of a traditional overhead projector. The students are asked to indicate their choice of answer by pressing the corresponding number on their handsets. The registration of each vote with the system we use takes a fraction of a second. During this period, the system is unable to register other votes. Consequently, it takes some time for all the votes to register. It is our experience that after a little practice 150 students are able to vote within 2 minutes.

2.3. Educational principles

What is the rationale for using the Personal Response System to replace the traditional answer-and-question sessions? There are several justifications possible. Some are specific to the course in which we applied the

system, whereas others are more general. The course in which the Personal Response System was implemented was a large class given in a large lecture room. This set-up imposes a necessary passivity on the audience, in which it never feels necessitated to re-express what it has learned. Even if it is asked a question, it can always hide behind the anonymity of the large class size. Whereas it should be acknowledged that learning could take place in many ways, re-describing the acquired information in one's own terms is an activity that powerfully promotes learning, as constructivist theories of learning have pointed out. First and foremost then this system encourages all those that attend the class to express their knowledge beliefs without any of the traditional inhibitors such as the fear of ridicule by the lecturer or their peers. No matter how much a lecturer stresses that making mistakes is encouraged, public humiliation seems a deeply rooted anxiety.

Another important use of the handset system is to initiate discussion. When everyone has made a choice and voted for a particular answer, the psychological investment in that answer turns the student from a passive attendee into an active participant for whom the outcome has some emotional value. A strategy we often pursued for difficult questions in which the correct answer was not clearly selected was to show the students the distribution of given answers and asked them to discuss with their neighbours if they would like to reconsider their own answer.

Moreover, besides encouraging students to engage with the class material, it provides useful feedback for the lecturers on the reception of their lectures. Periodic use of the handsets system gives the lecturers the opportunity to adapt the lectures to the perceived difficulties of the class in front of them. This is a form of interaction that goes beyond mere question-and-answer sessions. The handsets system can be used as a form of formative assessment.

2.4. Multiple choice questions

Multiple-choice questions are sometimes seen as restrictive and assessing only a certain type of knowledge. This is common but unnecessarily bleak view on this type of questioning. McBeath [5] describes how multiple-choice questioning can be used to evaluate everything from plain definitions to intrinsic interpretation issues. All levels from Bloom's famous taxonomy of educational objectives [1] can be addressed via multiple-choice questions.

In particular, for almost every question it is possible to ask a dual question. For example, rather than asking what the definition of a particular concept is, it is sometimes more useful to ask which concept is defined by certain definition. Besides basic definitions, the course focuses particularly on exploratory graphical analysis and interpretation of formal analysis. Many of the questions asked are one of the following types:

1a) What is the definition of this concept?
 b) What concept is defined by this definition?
2a) Which of these statements are correct about this plot?
 b) Which of these plots can you associate with the following statement?
3a) Which of these analysis methods is useful for this data?
 b) Which of these datasets could be analysed by this method?
4a) Which of the following interpretations corresponds to this output?
 b) Which of the following outputs warrants the following interpretation?

Testing the different directions of the link between the general and the particular is not only important in itself because understanding requires understanding all of these, but it also keeps the type of mental demand on the students fresh, even if one is in fact sticking to the same topic.

For the same reason it is sometimes helpful to use trick questions or so-called "brain teasers". It is important to recharge students' attention halfway through the hour of the tutorial. One can use notoriously misunderstood issues and turn them into multiple-choice questions where common misunderstandings are among the options.

What might seem contradictory in designing good questions is the importance of having a certain fraction of students making mistakes. Although once and a while, it is comforting for both students and lecturer to come across a question in which almost everyone has the correct answer, the steepest learning curve can be expected from those students that understand the general gist of the subject, but still make occasional mistakes. To make the most of these mistakes, one should encourage discussion among students after the first voting round. This involves the better students into the learning process. Their explanations hopefully serve to enforce knowledge among those with some gaps in their understanding as well as reinforce their own.

Occasionally, in a second round of voting the answer distribution gets worse. Apparently, there was not enough critical mass in the audience to spread the correct answer around. Such a situation calls for intervention by the lecturer. Either the lecturer can

follow the route of a traditional explanation, perhaps followed by an additional round of voting. The flexible nature of the handset technology can also be exploited by engaging into a set of branching questions to explore the origin of the confusion in the students. It can be, after all, more a matter of a confusing question, rather than a conceptual misunderstanding.

An option is to set up the tutorial as a sequence of branching questions, where what is actually presented depends on how the audience respond to the previous questions. It is even possible to include a question in which students express what part of the lecture material they found most complicated and consequently focus on that. However, this should be used with care, as it is well known that learners have difficulty distinguishing between superficial details and essential difficulties. Making a flexible branching tutorial requires extra work, but increases interactivity.

Designing a question is one thing, coming up with several answers is another. From our practical experience we have deduced four requirements that make a difference. They are:

(i) comprehensibility of the language

Every subject matter comes with its own jargon and it is all too easy to use language in either the questions or answers that distract from the issue that is really tested. For example, if the question is why it is preferable to use the one-sample t-test rather than sign-test when the conditions are satisfied, then the correct answer can be stated in several ways: a) t-test has a larger power b) a t-test is less likely to accept the null-hypothesis when the alternative hypothesis is in fact true. To be able to choose (a) one should not only understand the conceptual link between the different tests and their associated errors, but also the technical vocabulary of "power". If they make a mistake, it is not clear if they didn't understand the conceptual issue or just didn't know the jargon. It is therefore better to separate jargon questions from conceptual questions.

(ii) likelihood of the options

The common use of multiple-choice questions makes most students strategic respondents. Often, they purge the least likely choices by a process of elimination and then evaluate the remaining options on the basis of form alone. Whereas this ability serves them well during examinations, in this formative assessment it is a skill that actually prevents them from engaging with the question. It is therefore essential that the lecturer spends a certain amount of effort on devising answers that are logical in form and structure as well as plausible from the point of view of a strategic respondent.

(iii) preventing guessing by allowing the expression of ignorance

The absence of obvious penalties makes guessing the answer always an option. However, guessing diminishes the educational value of this formative exercise and should be discouraged. The only way to allow students not to guess is by including an option in which they admit their ignorance. We suggest that this "I-don't-know" choice is included by default. However, we found that we had to point out to students repeatedly that such an option actually exists and should be taken seriously. The other side of the coin, obviously, is that expressions of ignorance should be dealt with seriously too and not merely discarded.

(iv) limiting the number of options

Some thought has to be given to the number of choices that are provided. Our experience tells us not to exceed five answers, including the "I-don't-know" option. It is not easy to take in a lot of information from a screen. Whether it is plainly the font size, the clogging of the screen or some other, deeper psychological reasons, when confronted with a lot of choices, it takes a long time to sort through them before being able to answer. Perhaps if students had the answers in front of them on paper and were able to make marks with a pen, this process would improve. However, such an approach would conflict with the more interactive, open system that we propose. Therefore, limiting the number of options is the most advantageous strategy.

3. Implementation and Results

We have used the handset system during the 2002/3 autumn and spring terms in the Statistics for Psychologists course roughly once every two weeks. Lectures of the course took place from 1 to 2 pm, Mondays through Thursdays throughout the term. The handset tutorials replaced the Thursday lecture of every second week. The tutorials have been monitored by several observers (Steve Draper and Margaret Brown, Psychology, University of Glasgow, Gregor Kennedy, Education, University of Melbourne) whose comments have been included throughout this section.

We describe the implementation by the author of the tutorial on Thursday 5 December 2002. The tutorial started five minutes late at 1:10pm, because the class that occupied the room before us spilled over into our hour. Approximately 70 out of 210 students attended. This was slightly less than the attendance rates of tutorials before and after this one. When asked why only one third of the students showed up, several students suggested that there was a deadline tomorrow for another course. As Table 1 shows, these numbers

are, nevertheless, a clear improvement over previous attendance rates for the tutorials.

Table 1 Approx attendance rates for tutorial in S1C

Tutorial	1st–2nd	3rd–4th	5th–8th	9th–10th
Previous years	60%	25%	10%	10%
2002 – 2003	80%	50%	40%	n/a

This tutorial covered the three topics that had been discussed in the previous three weeks: hypothesis testing, correlation and simple linear regression. I suspected that not all topics were considered equally difficult, so I decided to let them vote with the handsets for their preferred topic of discussion. This meant that I had to prepare many spare questions, but it increased flexibility and interaction during the tutorial.

A majority wanted to discuss regression, whereas a smaller group wanted to discuss hypothesis testing. During the remainder of the hour four regression questions and three hypothesis-testing questions were examined. In the first round of voting almost all the questions split the audience, which was a sign that the questions addressed common difficulties with the material. Then the students were given the opportunity to discuss the questions with their neighbours and vote again. Afterwards, the correct answer with explanation was provided. An observer commented that if the students give many incorrect answers, then the ideal explanation includes a lot more than what the right answer is, and why it is the right answer; it should ideally say something about why the wrong answer is attractive but still wrong.

In this tutorial I introduced a 50:50 technique, in which, in several instances, I eliminated all but 2 possible answers, while explaining why the others were wrong. Normally this gave me the chance to address two separate issues in two separate instances: one after the first round of voting, the other after the second. An example is question 1.

Q1 The null-hypothesis for a Wilcoxon test could be

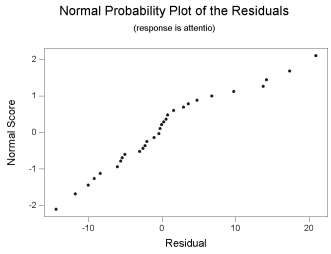
1. The population mean is 35
2. The sample mean is 35
3. The sample median is 35
4. The population median is 35
5. I don't know

The correct answer (4) was the most preferred answer followed by (3), but none of the answers received more than 35% of the vote. After the first voting round I eliminated answers 1 and 2 and explained that the Wilcoxon test makes inference about the median, rather than the mean. Discussion among themselves swung the vote distinctively (75%) in favour of the correct answer, but leaving a quarter still in the dark about the difference for inference between a sample and a population. The second explanation focused therefore exclusively on why hypotheses are framed in terms of population parameters rather than sampling summaries.

To keep the attention fresh and appeal to different problem solving skills, the type of questioning varied constantly. Questions 2 and 3 deal both with assumption checks in regression problems, but test reverse type of skills.

Q2 If your normal probability plot looks like this, then what is the case?

1. The data are linear
2. The data are normal
3. The data are not very linear
4. The data are not very normal
5. I don't know

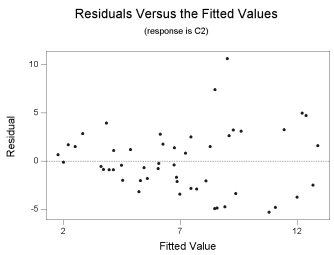


Question 2 simulates a situation that they can come across in practice when they are performing a statistical analysis. They have to interpret the figure in the light of the regression analysis they are performing, know that the probability plot tests for normality of the residuals, and interpret the non-linearity of the line as a violation of that assumption. Question 3 is an analytical question. Ideally the student analyses the question and deduce that an increasing variance with the response should correspond with increasing fluctuation of the residuals and that therefore the first figure is the correct one. So, whereas question 2 is an interpretation question, question 3 is a conceptual question aiming at the students' analytical skills.

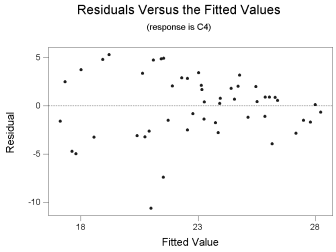
An important component in the design of every question is the existence of credible alternatives. Such brainteasers stimulate a careful consideration on behalf of the student. Although other, less propitious interpretations are possible, the wide range of answers in each first round of voting did suggest that this aspect of the question design had been successful. In each of the seven questions, the second round voting improved the percentage that got the correct answer. In one case, a third voting round was held. This round did not yield much discussion or any substantial changes in the voting pattern. At several occasions verbal responses

Q3 Which of the following plots shows that the variance is increasing with the response?

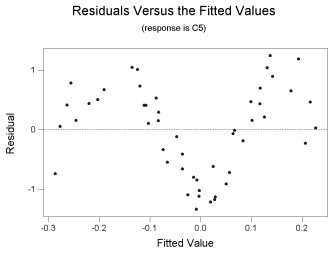
1.



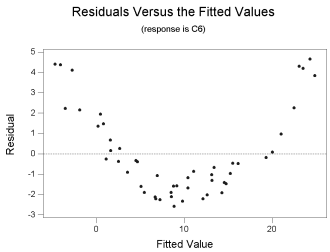
2.



3.



4.



5. I don't know

from the audience were solicited to explain their motivation for a particular answer, but very few people responded.

At 1:50 pm the class ended five minutes ahead of time as to allow sufficient time to stow the handset equipment before the start of the following hour. At the same time, one of the tutorial monitors interviewed three students. Despite the quiet nature of the class, all three were strongly enthusiastic about the handset tutorial. They all mentioned anonymity as an important reason why they like the personal response system. This is perhaps a partial explanation for the lack of verbal interaction during the tutorial. Other reasons for liking the tutorial were a) the way handsets allow them to produce their own answers rather than just listening to the one person who answers b) that they get on the spot answers c) that they get on the spot explanations of the right answer.

4. Discussion

There are clear benefits of the PRS. Using handsets is fun and breaks up the monotony of the lectures; it makes lectures more interactive and involves the whole class; students are able to contribute without fear of making a mistake; students have an idea how they compare to their peers; it gives lecturers a chance to see if students are doing as well as they think *before* the final exam. These are all valid reasons and explain why the author enjoyed using them in practice.

However, the current system certainly has limitations. The flexibility that this system offers pertains mainly to the lecturer. It does not give the students the opportunity to ask actively for clarification in the same anonymous way as they can answer questions. Although the technology for such a system is readily available, it is still not economically viable to implement. Also, whereas all formal examining in the course is done via open answers questions, the PRS relies on multiple-choice questions. Other issues that have sometimes been mentioned on the downside of the PRS are: setting-up the system takes time and there have been occasional problems with the projector; the "glamour" of the system can distract from the learning aspect or, worse, leads to asking questions just for the sake of it.

Our implementation of the PRS was exclusively introduced to deal with a weakness in the course programme. All the lecturers were trained in using the system and were assisted by one or two teaching assistants familiar with the system. The exclusive use of the handsets for the tutorials eliminated the danger of using the PRS just for the sake of it. In a study performed on March 6, 2003 among all the students attending a S1C handset tutorial, 74% thought that the use of handsets had been useful to very useful in aiding their understanding of statistics. Of the remaining students 22% did not have an opinion. At the same time, 87% saw more benefits than disadvantages in the use of handsets. Although it should be observed that the study was performed among those present, it is clear that students have a positive perception of this form of instruction. From the point of view of the lecturer it is clear that this system requires some extra work. Multiple-choice tutorial questions have to be prepared, which can be time-consuming. On the other hand, the handset system has given us useful feedback about the

performance of the students. Moreover, rather than frustrating hours in front of a fraction of the class, these handset tutorials have been rewarding if only by seeing the positive response from the students.

Although the data is in principle available, it is not yet easy to compare results across students for each question. Current development of user-friendly handset software will make this easier [6]. One of the unanswered questions is whether there are several types of students that are good at certain things and bad at others or whether everybody is equally likely to make a mistake. This type of analysis might be useful to assist the lecturer in designing the course.

References

[1] Bloom B S ed (1956) "Taxonomy of Educational Objectives," *Handbook 1: Cognitive Domain*. Longman, Green & Co, New York
[2] Draper S W (1998) "Niche-based success in CAL" *Computers and Education* vol 30, pp5-8
[3] Draper S W, Cargill J, & Cutts Q (2002) "Electronically enhanced classroom interaction," *Australian journal of educational technology*, vol 18 no.1, pp13-23
[4] Laurillard D (1993) *Rethinking university teaching: A framework for the effective use of educational technology* (Routledge, London) p103
[5] McBeath R J ed (1992) *Instructing and Evaluating Higher Education: A Guidebook for Planning Learning Outcomes*. ETP, New Jersey
[6] Mitchell C M (2003) PowerPoint add-in for Personal Response Systems, http://www.dcs.gla.ac.uk/~mitchell/PRS/Downloads.html

TechWatch Courseware News

More info from TechWatch online at http://ltsn.mathstore.gla.ac.uk/news/techwatch.asp

LiveMath Maker 3.5.4 in now available: Improvements and fixes include: Collect now really collects common factors, even the harder ones. Factor had an issue where it tried too hard on some complicated, multi-variable polynomials. Factor now tries the right amount,and works well with the improved Collect feature to crack even more polynomials. Calculate Limits have been improved.

NEW Mathematica Information Center and Graphics Gallary from Wolfram: The new Mathematica Information Center is a collection of nearly five thousand Mathematica programs and documents categorized for easy browsing and searching.

Operations Research v3 is out: v 3.0 of Operations Research, a Mathematica application package for constrained optimization, is available from Wolfram.

Maple T.A.v1.2 is now available from Adept Scientific: Maple T.A. provides schools with a complete infrastructure for creating, administering and grading free-form, online mathematics exams and assignments. Maple T.A. grades an assignments with the Maple engine, the most trusted software for exploring mathematics and developing mathematical applications. Randomization of questions in order to hinder cheating and encourage repetition of exercises. It also lets instructors include free-form mathematics questions on assignments, even when there are thousands of students in the class.

Word2TeX and TeX2Word products out from Chikrii Softlab Ltd: Word2TeX is a converter designed in order to use with Microsoft Word and enables Microsoft Word to save documents in LaTeX format. TeX2Word is a converter designed in order to use with Microsoft Word and enables Microsoft Word to open TeX/LaTeX documents.

TeX2XML project from Chikrii Softlab Ltd: aims to develop tools for converting From TeX/LaTeX to XML+MathML (including XHTML+MathML or any other XML-based markup language)

MathOCR project from Chikrii Softlab Ltd: Development work on optical recognition of mathematical expressions (both written by hand and printed) has been successful after two years.

Control System Professional Suite released by Wolfram: Designed for professionals, researchers, educators, and students using control theory and its applications.

MathML input improved in Perception: Perception V3.3.1 was recently released complete with a new wizard that both simplifies and improves the way Math ML can be added to questions.