

LETTING STUDENTS UNDERSTAND WHY STATISTICS IS WORTH STUDYING

Eric R. Sowe

The University of NSW, Australia

e.sowe@unsw.edu.au

Whether they are studying statistics as a disciplinary major or through service courses, students will be more motivated towards what they are learning, and will retain a richer recall of it, if they feel they are doing something worthwhile. I have previously argued that three elements in teaching are salient in giving a sense of worthwhileness: showing that statistics is interesting, useful, and substantial. The first two of these elements are already well discussed. But letting statistics be seen as a substantial discipline, in the sense of being resilient to challenging questioning prompted by students' own curiosity, has not been previously addressed in the statistics education literature. Here I show the kinds of challenging questions which serve this goal. The answers given need not be overly-detailed: what matters is that they satisfy students' curiosity. In this way they strengthen the students' sense that statistics is worth the effort of study.

“A powerful idea communicates some of its strength to him who challenges it.”

Marcel Proust (1943, p.191)

INTRODUCTION

Whether they are studying statistics as a disciplinary major or through one or more service courses, students will be more motivated towards what they are learning, and will retain through time a richer recall of it, if they feel they are doing something worthwhile. What gives students the sense that studying statistics is worthwhile? In Sowe (1995), I have argued that three elements in teaching are salient in developing this sense of worthwhileness: showing that statistics is interesting, useful, and substantial.

Ways of actively demonstrating that statistics is interesting and useful are already well discussed in the statistics education literature.

Approaches to teaching that bring out the intrinsic interest of statistics are explicit in the portrayal by Bradstreet (1996) and Sowe (2005) of the power and rewards of statistical reasoning, and in the insights of Snee (1993) and Sowe (2001) on the appeal of experiential learning and intellectual excitement in statistics classes. Such approaches are also implicit in the essays collected in Peck *et al.* (2005), that showcase the fertility of statistical methods.

Approaches for convincing students of the practical utility of statistics are explicit in the recommendations of Anderson and Loynes (1987) and Smith (1998) on student projects and assignments; of Peck, Haugh and Goodman (1998) on statistical case studies; and of Hahn and Schmee (1987) on bringing practitioners into the classroom. Such approaches are also implicit in the papers by Gani (1983) and Wallman (1993), that highlight the importance of statistical literacy.

Many of these approaches are as feasible to implement in a service course in statistics as they are during a disciplinary major.

However, letting statistics be seen from the outset also as a discipline of substance – that is, one that is in no way shallow, unsystematic or fragmentary – is a goal that has not been previously addressed in the extensive statistics education literature. [In this context it is interesting to note that Moore (1992) interprets statistics as a discipline of substance in another sense, namely that it is a discipline in its own right – a ‘respectable’ discipline, as he puts it – and is not a subfield of mathematics.] It is something which students (especially those in service courses) – if they think about it at all – mostly take on trust, but that need not be so.

In this paper, I look at how the substantial nature of statistics can most tellingly be made apparent. My experience in teaching statistics and econometrics over more than thirty-five years suggests that it is preferable not to be too didactic about this but, rather, to progressively provide the kind of information that will let students arrive at that understanding themselves. Such

students are, moreover, well prepared for discovering the longer-term intellectual rewards of statistical research.

I shall take the term ‘statistics service course’ to mean a course designed for students who are pursuing a statistics-using discipline, but who do not intend to become statisticians. Not all statistics service courses are at an introductory level, but they are often stand-alone, that is, they are not part of a progressive sequence in statistics.

STATISTICS AS A SUBSTANTIAL DISCIPLINE

Especially within the limited purview of a service course, students may gain the impression that statistics, as a discipline, is a bundle of esoteric analyses, based on complex concepts, for solving abstruse problems.

Even if they are shown that such problems are of practical importance and that the solutions have intrinsically interesting features, students may nevertheless miss the realisation that all the statistical techniques they study are based on clearly-defined and logically defensible assumptions, are systematically built up, are coherently inter-related, and are versatile in application.

In short, they miss the realisation that statistics is a substantial discipline.

What is the consequence of missing this realisation? It may impede students’ appreciation of the outstanding worth of statistical techniques, and consequently stifle their motivation towards anything other than superficial study. By contrast, a student who grasps that statistics is a discipline of substance will have a strengthened conviction that it is professionally empowering and, thus, worth the effort of careful study.

How, then, might the substantial nature of statistics be made apparent?

It is the thesis of this paper that an effective way for students to discover that statistics is a substantial discipline is by observing its resilience to challenging questioning.

Since thoughtful people enjoy delving into matters that pique their curiosity, the challenging questions can ideally be expected to come from students themselves. If the teacher shows commitment to his/her educational role and has developed a rapport that encourages a free flow of questions from the class, then challenging questions are sure to appear, scattered among more mundane inquiries.

To every challenging question from a student (whether or not it falls within the scope of the course syllabus) the teacher has a cardinal responsibility to provide an answer which fully satisfies the student’s curiosity. Such an answer need not necessarily be highly detailed. Unfortunately, students are not always academically, socially or culturally at ease in posing challenging questions. How to proceed in such circumstances is considered later in this paper.

WHAT ARE ‘CHALLENGING QUESTIONS’ IN THIS CONTEXT?

Challenging questions are questions that go to the essence of statistics as a substantial discipline. They are questions that seek solutions to enigmas of statistical theory, or greater insight on the foundations, structure, limitations or interconnections of theoretical techniques, or fuller guidance on the valid and appropriate application of those techniques in practice.

In both introductory and higher-level statistics courses, challenging questions arise as a student

- (i) queries logical deficiencies or fuzziness in classroom or textbook exposition, *or*
- (ii) probes as-yet-unexplained matters that occur to and intrigue him/her.

In higher level courses, challenging questions are typically also prompted when a student feels

- (iii) not enough has been said about the subtleties or limitations of a particular technique, or about what alternative techniques may be available for the same objective, *or*
- (iv) inadequate justification has been given for the choice of statistical analyses in some applied statistical study.

QUESTIONS THAT MIGHT ARISE IN AN INTRODUCTORY STATISTICS COURSE

In this and the following section I offer a selection of such challenging questions, gathered from my classes at all levels. The questions are broadly in students’ wording. They

relate to topics in 'mainstream' statistics education. Specialised fields that use statistics, such as biometrics, econometrics and psychometrics, can elicit challenging questions of their own.

- (a) "The textbook says ' $P(A) = 0$ means the event A is impossible.' When a fair coin is tossed we say ' $P(\text{coin shows head}) = P(\text{coin shows tail}) = \frac{1}{2}$,' and that surely implies ' $P(\text{coin lands on edge}) = 0$,' but that's not impossible in a muddy field, is it?"
- (b) "In a histogram, why must the frequency in each class interval in principle be in proportion to the area of the corresponding rectangle, rather than just to the height of the rectangle?"
- (c) In the context of the binomial distribution "what happens if the successive trials are not independent? What happens if there are more than two possible outcomes at each trial?"
- (d) "How can human heights and weights be said to be normally distributed when the range of the normal is infinite?"
- (e) "What is 'randomness'? How do you know when a sample is random?"
- (f) "What would statistics be like if there wasn't the Central Limit Theorem?"
- (g) "For valid inferences about a diverse population, why isn't it more important to have a representative sample than a random sample?"
- (h) "I've been told that it's always valuable to know the standard error of a point estimate, but in newspaper reports of opinion polls I've never seen a standard error mentioned. Why not?"
- (i) "Since the optimal properties of a point estimator – unbiasedness, consistency and efficiency – are all defined in terms of the sampling distribution of the estimator, what comfort are they to me when in practice I have just one sample?"
- (j) "If I can be only 95% confident that the population mean lies within the range of values specified by a confidence interval, then in the case of my particular sample it might not. So why am I better off with a confidence interval than with just a point estimate?"
- (k) "What do you do if you want to use a t-test but you think the data are far from normally distributed?"
- (l) "If you're testing a mean and, it's practically more important to avoid a type II error than a type I error, how can you be sure your risk of a type II error is actually smaller than your risk of a type I error given that you don't know the actual value of the population mean?"
- (m) "In least squares regression with Y as the regressand, why are residuals measured parallel to the Y axis, and not in some other direction?"
- (n) "Why do we use least 'squares' in regression estimation, rather than some other power of the residuals?"

Each of these questions shows an inquisitive mind at work. A few of the questions were asked spontaneously as new material was introduced, but the majority arose after students had had time to reflect on what they were learning.

Questions (a), (d) and (i) show how easily, and even unsuspectingly, a thoughtful student can pierce the polished veneer of an oversimplified textbook exposition. Questions (b), (m) and (n) underline the importance, in teaching, of justifying what may, to some students, seem quite arbitrary conventions. Questions (c) and (k) attest to the fact that interested students may not be content to stop short at syllabus boundaries. Questions (e) and (f) are typical of deep-thinkers who crave a 'big picture' understanding. Questions (g), (j) and (l), by contrast, display perceptive thinking about various 'close focus' issues. Question (h) highlights the reality that statistical methods are often misused by non-statisticians, and not always innocently. Many beginning students find this a fascinating theme, yet it is rarely given serious attention in first courses.

Though the inquiring students may not initially realise it, several of these questions (for example, (a), (e) and (i)) are quite profound. Answering such questions at introductory level without getting into a spiral of technicalities is by no means easy, and certainly requires some forethought from the teacher.

Collectively, these questions identify students who are seeking broader, deeper and/or richer insights into statistics than the course syllabus may be designed to provide. By responding

appropriately, the teacher can maximise the opportunities these questions offer to strengthen students' sense that statistics as a discipline is worth the effort of study.

QUESTIONS THAT MIGHT ARISE IN A HIGHER-LEVEL UNDERGRADUATE COURSE

- (o) "If you get a set of data that have some errors of measurement, but you don't know exactly what sorts of errors they are, is there any point in doing statistical analyses with those data? I can't see anything about this in the textbook."
- (p) "I've noticed that the best estimator of the population mean is the sample mean; the best estimator of the population median is the sample median; and the best estimator of the population variance is the sample variance. Is that a pattern I can rely on for finding best estimators?"
- (q) "The textbook says that, for a given confidence level, the shorter a confidence interval the better. I can see that an equal-tail confidence interval for a normal population's mean is always the shortest possible interval because the normal is symmetric. But the textbook also gives an equal-tail confidence interval for that population's variance, based on the chi-square distribution – which is nonsymmetric. How can that always produce the shortest possible interval?"
- (r) "If a biased estimator may be more efficient than an unbiased estimator, how do I know which one to prefer, since in practice I can't tell how big the gain in efficiency is relative to the amount of bias in the biased estimator?"
- (s) "In maximum likelihood estimation, couldn't it happen that the likelihood function just keeps increasing and never reaches a maximum?"
- (t) "How do I decide what level of significance to use in any statistical test? And if I test several parameters within the same estimated model, should I use the same level of significance for each test – even if the null hypotheses are not all of the same form?"
- (u) "I've heard that, if I evaluate the same statistic in two different samples, then even a small difference between the two values can always be made statistically significant if only I use big enough samples. Is that really true?"
- (v) "If the goodness of fit in regression, measured by R^2 , never decreases when I add another regressor, it seems that I could make a really well-fitting model even if I added in totally irrelevant variables. Isn't there something wrong here?"
- (w) "How do you realistically interpret a regression-based 95% prediction interval for the regressand when you're working with time series data? What worries me is how you can think of having repeated sample realisations over the regression estimation time period. The real world has only one set of data values over any time period."
- (x) "If I do repeated regressions with the same set of data, looking for a 'best' model, then I am data mining and that is bad. But suppose I hit on the 'best' model first up, then I have found it 'respectably.' But it's the same model either way, isn't it?"
- (y) "I've learned that 'correlation does not imply causation,' but now I'm reading a paper about 'Granger causality' that claims to detect causation in a regression relation by a correlational procedure. Isn't there a contradiction here?"
- (z) "If computer output for an estimated simple regression model shows that the (upper-tail) p-values for the intercept and slope are 0.004 and 0.002, respectively, doesn't that mean the slope parameter is twice as likely as the intercept parameter to be greater than zero?"

As might be expected of higher-level students, more questions here than in the previous section touch on the philosophical foundations of statistics. This can be complex territory. In responding to questions (r), (t), (w), (x) and (z), in particular, the teacher would do well to prepare carefully.

Of the remaining questions, (q), (s) and (u) suggest the questioners are probing below the surface of textbook explanations. Questions (v) and (y) aim to resolve paradoxes. Question (p) seeks the truth on an intuitively-plausible principle that a student has stumbled upon. And question (o) reveals a student concerned about a possible gap between theory and practice.

As in the previous section, each of these questions represents in some way a student's challenge to taught ideas. If these questions receive satisfying answers, the questioners will not

fail to observe the resilience of the discipline to their challenges. In this way, memorable impressions may be generated of the substance and power of statistics.

ELICITING CHALLENGING QUESTIONS

Asking questions in class is sometimes seen as an indicator of students' involvement with what they are learning. Yet, it is a deceptive indicator, for there are several reasons why some students – even students who feel involved – rarely or never ask questions.

Firstly, there can be a reluctance to look foolish in case one's question seems trivial or irrelevant to others. Secondly, a shy student may wait to see whether his/her unvoiced query will be resolved as the teacher proceeds. Thirdly, a student who is 'lost' in the flow of ideas may be unsure about exactly what question he/she needs to ask and lets the moment pass. And fourthly, cultural conditioning may restrain some questioners, lest their questions be taken as implying criticism of the teacher for being obscure.

All these constraints impact on a service course, as they do on any other course. But there is a further constraint peculiar to a certain kind of service course. If the course has developed a reputation as one that may be 'passed' largely on the strength of rote learning, because the assessments hardly probe deep understanding, then students may simply not bother to ask questions.

If a reluctance to ask questions of any kind becomes entrenched, it will certainly be difficult, even for a teacher committed to enhancing student learning, to elicit the kinds of challenging questions set out in the previous two sections.

What can be done?

The teacher should work from the outset to welcome both spontaneous and reflective student questioning. In large classes a standing invitation can be extended to email questions after class, with answers posted for everyone to read on the course webpage. In smaller classes, questions posed in class should be responded to at once. In both settings, replies should be constructive, and considerate of the questioner: curtness or condescension self-evidently have no place in teaching that aims to show that the study of statistics is worthwhile. The teacher, too, can make it a lively feature of the classroom experience to pose questions to students. As in all other facets of teaching, the behaviour the teacher models becomes the currency of the class.

In a different category of obstacles to free-flowing questioning is the social conditioning within some cultures that ordains that the teacher is an authority whose pronouncements are not to be questioned. This is, for students who are so influenced, a particularly powerful inhibition, and one which requires subtle strategies pursued over some length of time to counter. An overview of such strategies can be found in Sinclair and Wilson (1999, pp. 44-47).

If a rapport has developed that takes for granted a free flow of questions in class, then, as already mentioned, challenging questions are sure to appear. Not all students may have the perceptiveness or self-confidence to pose such questions, but those who do should be given positive reinforcement in the presence of their classmates. If discussion of a question would stray too far from the current topic, the teacher may hesitate to allow it to distract the class at that moment. The questioner could be answered fully out of class and the core of the matter revisited in class on the next suitable occasion.

Should a challenging question not be forthcoming from the class at a point where the teacher can see one that is instructive, he/she should not hesitate to step into the breach. Better the question be posed by the teacher than omitted altogether.

FURTHER BENEFITS OF ASKING CHALLENGING QUESTIONS

Challenging questions are always *searching* questions. Conceiving and formulating such questions represent first steps in independent thinking. Well-judged answers given to these questions, then, can do more than strengthen a student's sense that statistics is worth the effort of study. They can, as well, bolster the student's confidence that he/she is progressing well in 'learning how to learn independently' what is already known in statistics. It is then not such a big step for the student to begin asking *researching* questions, that is, to begin exploring what is not yet known in statistics.

ACKNOWLEDGMENTS

I am pleased to acknowledge the helpful comments of the anonymous referees.

REFERENCES

- Anderson C. W. and Loynes, R. M. (1987). *The Teaching of Practical Statistics*. New York: Wiley.
- Bradstreet, T. E. (1996). Teaching introductory statistics courses so that nonstatisticians experience statistical reasoning. *The American Statistician*, 50, 69-78.
- Gani, J. (1983). The relevance of statistical training. In D. R. Grey *et al.* (Ed.), *Proceedings of the First International Conference on Teaching Statistics*, (pp. 50-70). London: The Teaching Statistics Trust.
- Hahn, G. J. and Schmee, J. (1987). Practitioner and academician co-teaching: An idea to consider. *1987 Proceedings of the Statistical Education Section*, (pp. 199-203). Alexandria, VA: American Statistical Association.
- Moore, D. S. (1992). Teaching statistics as a respectable subject. In F. Gordon and S. Gordon (Eds.), *Statistics for the Twenty-first Century*, (pp. 14-25). MAA Notes #26. Washington DC: Mathematical Association of America.
- Peck, R. *et al.* (Eds.) (2005). *Statistics: A Guide to the Unknown* (4th edition). Belmont: Duxbury.
- Peck, R., Haugh, L. D., and Goodman, A. (Eds.) (1998). *Statistical Case Studies: A Collaboration Between Academe and Industry*. Philadelphia: Society for Industrial and Applied Mathematics.
- Proust, M. (1943). "Madame Swann at Home" volume 3, part 1 in C.K. Scott Moncrieff (transl.) *Remembrance of Things Past*. London: Chatto and Windus. (French original published 1918).
- Sinclair, A. and Wilson V.B . (1999). *The Culture-Inclusive Classroom*. Melbourne: Melbourne Business School, University of Melbourne.
- Smith, G. (1998). Learning statistics by doing statistics. *Journal of Statistics Education*, 6. <http://www.amstat.org/publications/jse/v6n3/smith.html>.
- Snee, R. D. (1993). What's missing in statistical education? *The American Statistician*, 47, 149-154.
- SoweY, E. R. (1995). Teaching statistics – Making it memorable. *Journal of Statistics Education*, 3. <http://www.amstat.org/publications/jse/v3n2/soweY.html>.
- SoweY, E. R. (2001). Striking demonstrations in teaching statistics. *Journal of Statistics Education*, 9. <http://www.amstat.org/publications/jse/v9n1/soweY.html>.
- SoweY, E. R. (2005). From a logical point of view: An illuminating perspective in teaching statistical inference. *International Journal of Mathematical Education in Science and Technology*, 36, 801-811.
- Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association*, 88, 1-8.