Conceptual Challenges Facing A-Level Statistics Students: Teacher and Examiner Perspectives

James Nicholson & Gerry Mulhern
Queen's University Belfast, School of Psychology
University Road
Belfast BT7 1NN, Northern Ireland
j.r.nicholson@qub.ac.uk g.mulhern@qub.ac.uk

1. Summary

This study sought to investigate perceptions of students' conceptual challenges among A-Level statistics teachers and examiners. The nature and extent of participants' insights were assessed using a questionnaire administered in either written form or via a semi-structured interview. The questionnaire comprised two sections: (i) free-response questions in which participants were asked to list the three most significant conceptual challenges faced by students; and (ii) an attitude scale designed to assess agreement with specific statements regarding possible conceptual challenges. Each section addressed five topic areas: regression and correlation, estimation, sampling methods, distribution modelling, and general statistical thinking. Forty-nine participants completed the questionnaire, though not all teachers were familiar with all of the topic areas. Results revealed interesting patterns of agreement and disagreement among participants with regard to students' conceptual difficulties and concomitant factors.

2. Introduction

> *What many teachers are only beginning to fully appreciate, but many children do understand - even if they don't articulate it - is that, in future, it won't be how much you know that's important, but your ability to analyse vast amounts of data, and decide what to do with it. The winners will be those who can make that judgement.*
> (Taylor, 2000; p. 38)

Stochastic reasoning is of considerable relevance to contemporary society. There is much more 'raw data' available than at any time in history and unprecedented ease of access to powerful analytical tools. The increase in technological capability suggests that these trends will continue and indeed accelerate. Moreover, many professional people make critical decisions, sometimes involving life or death, in situations where at least part of the information is probabilistic.

In the UK secondary schools, Advanced-Level statistics is usually studied as an optional part of the mathematics curriculum for academically competent 16 - 19 year olds, although the same topics also form part of most introductory statistics courses in tertiary colleges. Recently, new courses and assessments in A-level statistics have been published and are due to be introduced in September 2000. These place more emphasis on the applied nature of statistics and, consequently, a correspondingly greater emphasis on interpretation is likely. Moreover, since assessment requirements appear to exert a powerful influence on what happens in the classroom, such a change in emphasis has considerable implications for the future.

It is also increasingly important that reliable methods of overcoming cognitive obstacles in achieving a real understanding of stochastic principles are developed and disseminated. Shaughnessy (1992) has commented on the similarity between the factors hindering effective teaching of stochastics, and the impediments to the teaching and assessing of mathematical problem solving. He argued that it is not surprising that there should be these similarities, since the cognitive processes are alike, involving the construction of models of physical phenomena, the development and implementation of strategies, and the comparison and evaluation of different possible approaches.

Hawkins (1996) observed that the statistics teacher must teach the student *when and how* to use various techniques, not just the computational skills required to solve a problem. Hawkins (1996) cited Bradsheet's (1996) recommendation that statistical reasoning should be taught before statistical computation, although she acknowledged that this may not be easy for teachers to do, especially if they themselves lack confidence in statistical reasoning: 'Statistics is largely the exercise of common sense. Many teachers still need to be convinced of this, however, and until they are [convinced] their teaching is unlikely to be really appropriate or effective.' (p. 65). Gal (1996) identified two distinct aspects of handling data - *generative* skills, where students act upon data (i.e. *doing* statistics), and *interpretative* skills, where students form opinions about the meaning of the data. Gal (1996) argued that a question posed needs to be direct, in the sense that it poses a specific question. Using an example of a dataset concerning boys' and girls' views about toys, he found that participants produced more meaningful responses to the question 'In your opinion, do boys and girls have more or less the same kinds of preferences, or do they have different preferences?' than to the more generic question 'What in your opinion can the toy company conclude from the results in this table?', which led to much less informative responses.
Gal (1998) further argued that, in most cases, the reasonableness of students' opinions about data cannot be judged without having access to both the opinion and the arguments upon which it is based. He argued that teachers need to establish a 'culture of explaining' in the classroom. In our view, this places a onus on examiners to provide a framework of meaningful assessment within which students can grow and develop skills of statistical interpretation and judgement in realistic contexts. Recent practice has

not been entirely encouraging. While examination boards typically have published solutions and markschemes, in cases where candidates were expected to offer a short paragraph of analysis or interpretation, boards generally have offered a few key words or phrases, giving little guidance as to what a 'good response' should look like. Conclusions about hypothesis tests have also often been given in an abbreviated form, without any reference to the role of evidence in reaching the conclusion.

In our view statistical judgement is inherently difficult to assess, because the reasoning process may be complex. We would also argue that such assessments must be informed by educators' insights into the conceptual challenges faced by students in making statistical judgements. Gal (1996, 1998) has suggested that the context for a judgement needs to be explicitly defined, and that students should be encouraged to make explicit their reasoning processes in arriving at a judgement. We agree with Gal's view, but would caution against too literal an application of these principles. An interesting illustration of the problems that can arise was found in an examination paper issued by the Northern Ireland Council for Curriculum, Examinations and Assessment (CCEA, 1998), who asked candidates to calculate a correlation coefficient between two variables (C and R), and then to *'state, with a reason, whether you think there is a significant correlation between C and R'.* The markscheme indicated that the two responses *'Yes because [the] value is quite large'* and *'No because [the] value is quite small'* should both be awarded full marks. While it is important that examiners should be open to candidates making different interpretations or judgements, where these are supported by coherent and cogent reasoning, the quality of assessment items should be such as to avoid the impression that statistical judgement is merely a matter of personal opinion.

The present study sought to explore the insights of teachers and examiners into the conceptual challenges faced by A-level students learning statistics. The study is part of a larger project which aims to develop diagnostic and support materials for A-level statistics teachers and students.

3. Methodology

Participants: 40 teachers and 9 examiners were invited to take part in the study. Participants were drawn from a cross-section of examination boards in Northern Ireland and England based on existing contacts.
Materials: The nature and extent of participants' insights were assessed using a questionnaire administered, where possible, via a semi-structured interview. In order to extend the geographical range, some questionnaires were distributed by post and completed in writing.

The questionnaire comprised two sections: (i) free-response questions in which participants were asked to list the three most significant conceptual challenges faced by students; and (ii) an attitude scale designed to assess agreement with specific statements regarding possible conceptual challenges. Each section addressed five topic areas: regression and correlation; estimation; sampling methods; distribution modelling; and general statistical thinking. All teachers were currently teaching A-level statistics, although, due to the varying structure of different syllabuses, some teachers were not familiar with all topic areas.

Procedure: Questionnaires took approximately 30 minutes to complete. For the open-ended questions, notes were made on the tapes of the interviews, and subsequently analysed along with the written responses from the postal questionnaires.

4. Results and Discussion

Due to the exploratory nature of the study and, hence, the relatively small numbers of examiners, data for examiners and teachers are presented together. Where obvious differences arose, these have been highlighted. Overall, the biggest single issue which arose in the survey was the difficulty that even able students have with statistical interpretation. Examiners saw interpretation and inference as a bigger problem even than teachers did. Perhaps it has not been seen as quite such an issue by teachers because typically an examination paper has identified not more than 15% of the marks for interpretation, and it has been common for it to be less than 10%.

4.1 Importance and difficulty

Teachers and examiners were asked to list the 3 most important concepts to be grasped by A-level statistics students, based on their personal experience. Table 1 presents the five most frequent responses produced by participants.

Table 1: Most important concepts: frequency of response.

Interpretation - relating to the real world	18
Relationship between population and samples	16
Hypothesis testing	15
Making judgements	14
Probability	13

Participants were then asked to list the 3 concepts that A-level statistics students find most difficult to grasp, again based on their personal experience. Table 2 presents the four most frequent responses. Comparison of Tables 1 and 2 reveals strong correspondence between the patterns of ‘importance’ and ‘difficulty’ responses, with the exception of *making judgements* which occurred only 8 times in the case of the latter.

Table 2: Most difficult concepts: frequency of response

Interpretation - relating to the real world	19
Probability	18
Hypothesis testing	16
Relationship between population and samples	14

In the case of *probability,* several participants commented that the superficial treatment of the probability laws earlier in the curriculum, whereby only independent and mutually exclusive events are encountered and hence where students learn the mantra *'and' means multiply; 'or' means add* caused substantial difficulty at a later stage when the full range of contexts were encountered.

Other concepts identified by participants more frequently as difficult than as important were *distribution of the sample mean* (important = 1; difficult = 9) and *dealing with continuous and discrete random variables* (important = 4; difficult = 9). Five participants commented that a *different type of thinking* was required for statistics than for mathematics, and this idea seemed to underpin comments from a much larger number of teachers in relation to difficulties associated with decision-making and interpretation in context. In the semi-structured interviews, the general idea of randomness was identified as a problem for students, since the mathematics taught to date has been largely deterministic and statistics largely descriptive.

Again, in relation to difficulty, several other less frequent responses clustered around a single theme. 'Permutations and combinations' (7), 'expectation' (5) and 'calculation of standard deviation' (2) all relate to mathematical techniques underpinning statistics, while responses such as 'different type of thinking', 'it is not an exact science' and 'different from previous mathematics' (10 in total) address the contrast between formal mathematical and stochastic thinking. The following comments taken from a structured interview typify this latter theme:

> *It is a method of thinking - almost a philosophical understanding of what the whole subject is about, how the real world interacts with mathematics - wider than just interpretation. Even if things are running well, there will be variation and it is the business of statistics to account for and model the variability.*
>
> *The logic behind hypothesis testing offers an insight into the difficulties students face moving from the world of 'pure mathematics' to the world of inferential statistics. They seem to get a definite answer - a test has a 'correct outcome' for a specified level of significance, but that answer is not definite in certain broader senses - it is shrouded in doubt, in that it may or may not be correct. This requires a firm grasp of some subtle ideas - that a decision may be 'correct' in that you would always make*

that decision in the same circumstances, with the same knowledge, but yet be 'wrong' in that, as the future unfolds, another choice would have been more advantageous.
While pupils may understand in an abstract sense that extrapolation beyond the data may be problematic, they find it difficult to apply this abstract notion to concrete scenarios, for example, to identify in a particular practical context how an existing linear relationship might break down as a variable range is extended.

Two participants who teach groups with differing backgrounds offered some interesting perspectives. One teacher, working in a further education college with students who are generally relatively weak at mathematics, found that difficulties with computational aspects of the course hindered progress in conceptual development. Another, working with able students who had taken GCSE Mathematics a year early, followed by GCSE Statistics, indicated that they struggled to apply the mathematical techniques, learned a year previously, when required in a statistical context. It is interesting to note that two such diverse groups of students were seen to struggle with procedural aspects of statistics.

4.2. Relating importance and difficulty

A measure of commonality between views of importance and difficulty was constructed as the number of common items in the list divided by the total number of items in the two lists. This gave a 'score' lying between zero and one, for example, if one list had 4 concepts and the other had 3 with only 1 concept appearing in each list (i.e. 2 occurrences), then the commonality was calculated as 2/7. This rather crude measure of association between importance and difficulty proved remarkably symmetrical in its distribution over the possible range with approximately 20% of participants scoring 0 (no commonality) and 25% scoring 1 (identical lists).

It is possible that these extremes may be indicative of different pedagogic approaches. Specifically, at least some of those scoring close to 1 may be thought of as outcome driven, in that they placed importance on a topic because they consider it to be difficult, with obvious implications for student attainment. Obviously, others scoring close to 1 may coincidentally consider a topic to be both important and difficult, but not necessarily be outcome driven. Conversely, by definition, those scoring close to 0 perceive a clear distinction between conceptual importance and difficulty. Inevitably, such a suggestion is tenuous on the basis of these data, particularly in the absence of any external measurement confirming the teachers' pedagogic strategies, but we suggest that this might be an area of interest for future research.

4.3. Attitudes to students' understanding

A total of 28 statements were used to elicit participants' attitudes towards factors influencing students' understanding of stochastics. Participants were asked to indicate the extent of their agreement with each statement using a 5-point scale. Only six items were found to produce a strong consensus among participants, with 70% or more of the respondents in agreement. These were:

- 'students seem generally aware of the dangers of extrapolation too far beyond the range of the known data' - agree;
- 'in regression and correlation, new technology has reduced the number of students who make errors in generating regression lines or correlation coefficients' - agree;
- 'in point estimation, students confuse the concept of a population parameter (as a property of the underlying distribution) with that of the sample statistic (which is a function of the data)' - agree;
- 'students assume that large samples always give accurate results' - agree;
- 'students have a good understanding of what the level of uncertainty specified in a sampling method tells us' - disagree;
- 'becoming better at the procedural aspects helps because, as some of those processes become more automatic and less demanding of intellectual attention, it leaves more thinking capacity to address the big picture' - agree.

A lack of a consensus in the remaining 22 items raised more questions than it answered. A variety of factors, which are beyond the scope of the present study, may have given rise to differing views among participants. Among others, we would speculate about differences between the student groups with which various teachers are working (e.g., in terms of ability, or the amount of statistics taught); differences in teacher qualifications and experience; or the extent to which teachers may or may not be 'outcome driven'.

The latter suggestion is supported by data from the semi-structured interviews, in which responses to questions relating to conceptual challenges facing students were frequently couched in terms of particular examination questions which had caused problems for the students. This is consistent with the commonly held perception that assessment drives what goes on in the majority of classrooms.

5. Conclusion

The increasing dependence of contemporary society on processing information creates both opportunities and challenges to future generations. As Taylor (2000) points out, those who can handle and analyse large quantities of data will be at an advantage. The challenge facing statistics educators, both teachers and examiners, is to find effective ways of helping students develop these skills in dealing with stochastic contexts.

6. Acknowledgement

This work was supported by a grant from the Nuffield Foundation.

References

Bradsheet, T.E. (1996). Teaching introductory statistics courses so that nonstatisticians experience statistical reasoning. *The American Statistician, 50,* 69-78.

CCEA. (1998). *Examination paper for Module S2 in February 1998.* Belfast, UK: Northern Ireland Council for Curriculum, Examinations and Assessment.

Gal, I. (1996). Assessing students' interpretation of data. In B. Phillips (Ed.) *Papers on Statistical Education - ICME-8.* Hawthorn, Aus: Swinburne University of Technology.

Gal, I. (1998). Assessing Student Knowledge as it Relates to Students' Interpretation of Data.

I.S.P. Lajoie (Ed.) *Reflections on Statistics: Learning, Teaching, and Assessment in Grades K-12.* Mahwah, NJ: Lawrence Erlbaum Associates.

Hawkins, A. (1996). Teachers of statistics: Needs and impediments. In B. Phillips (Ed.) *Papers on*

Statistical Education - ICME-8. Hawthorn, Aus: Swinburne University of Technology.

Shaugnessy, J.M. (1992). Research in probability and statistics: Reflections and directions. In D.

Grouws (Ed.) *Handbook of Research on Mathematics Teaching and Learning.* New York: Macmillan.

Taylor, I.C. (2000). A tale of a PM and his mouse. *TES Online, Times Educational Supplement, 11 February,* 38.