IS A PASS GOOD ENOUGH IN TERTIARY STATISTICS?

ANNE WILLIAMS
Queensland University of Technology

Research in mathematics education and cognition suggests that knowledge, its acquisition, and its transfer are complex processes. In this study, end of semester totals are recorded for a large group of students in a traditional introductory subject in statistics at the tertiary level. A group of eight students are selected for observation while solving four statistical problems during the semester. This paper illustrates the capacity of students to pass an exam, yet highlights their limited understanding of statistical concepts, their narrow mechanical focus, their poor study habits and the lack of statistical transfer in atypical problems. Alternatives to the traditional approach are suggested.

The last decade has seen increased interest in statistics education. In particular, three international conferences on the teaching of statistics (see Davidson & Swift, 1987; Grey, Holmes, Barnett & Constable, 1983; Vere-Jones, 1991) have addressed some major issues involved in the teaching of statistics, namely: why should statistics be taught, what should be taught, to whom, by whom, and (most important of all) how.

With respect to introductory tertiary courses in statistics, there is some concern (see Grimmer, 1990; Moore, 1988; Morley, 1992; Taffe, 1991) that such courses often leave students dissatisfied, and fail to encourage enrolment in further statistics subjects. Despite this concern, a glance through students' records at any university would reveal that all but a small percentage of the huge numbers of students studying introductory subjects in statistics pass the subject. This raises questions about the usefulness of the knowledge acquired, and the type of statistical knowledge that is assessed.

Research into the nature of knowledge and the process of its acquisition emphasises the complexities involved. There is some agreement in the literature (e.g., Anderson, 1982; Long, 1989) that mathematical knowledge consists basically of two components: knowledge of content (facts, concepts, skills, procedures, strategies) and knowledge about how learning takes place (specific purpose procedures and higher order thinking strategies such as planning, analysing, monitoring, evaluating, modifying and predicting). Long (1989) suggests that for successful problem solving, both components should be used. To be used meaningfully, knowledge must be organised internally as patterns which are continually updated and modified. These patterns form the foundations for understanding. Research by Lawson (1984) and Perkins and Salomon (1987) imply that it is the knowledge of these higher order skills that allows students to solve novel problems. Thus it may be that current teaching practices in statistics do not encourage the second type of knowledge.

This paper focuses on one aspect of a larger study that explored the relationship between attitude and performance in a compulsory introductory subject in statistics. The purpose of the paper was to: (i) assess student understanding of statistical concepts in both typical text book as well as novel problems; and (ii) discuss the implications arising from the results.

METHOD

At the beginning of the first lecture of a compulsory inferential statistics subject, a large group of students, most of whom were enrolled in accounting, management or accounting/law degrees, filled in a questionnaire. This was designed by Wise (1985) to assess quantitatively the attitudes students hold towards statistics. A total mark ranging from 29 to 145 was tallied on this Likert scale of 29 items. Scores from 40 to 88 were classified as low attitudes: scores from 105 to 145 were classified as high attitudes. Eight students were selected for individual observation while solving statistical problems, one student for each possible combination of attitude (low, high), age (less than 25 years, at least 25 years) and gender.

After the final exam, the end of the semester results were recorded for 456 students, those who had not withdrawn from the subject. End of semester results were an aggregate of exercises which were worked and assessed in various ways throughout the semester. Most were the typical text book exercise, in which the question was clearly defined, an hypothesis test was required and numerical information was supplied.

Students had minimal statistical experience prior to this subject. It was understood that they had some knowledge of basic concepts such as mean, standard deviation and sampling techniques, but these were revised in the first week. Teaching and learning were "traditional", in that lectures were teacher-centred and exercise-driven. Topics in the course ranged from data collection, estimation and hypothesis testing, to multiple regression and nonparametric tests - a heavy workload. To assist students in doing hypothesis tests, a six-step solution had been emphasised throughout the semester. These steps involved: stating the hypothesis, significance level, test statistic (formula) and decision rule, performing the calculations, and reaching a conclusion.

The eight students were observed and audio-taped individually as they worked through four statistical problems (see Appendix 1), using a calculator when necessary. They were asked to talk aloud while they worked. The term "problems" is used here in the broad sense, to refer to typical statistical exercises as well as novel ones. The first two problems were worked during week 9 of a 14 week semester. This week was selected because it was just after the midsemester test. Had students prepared well for their exam, their statistical knowledge would then be at a premium. The midsemester test was multiple choice, and consisted mainly of exercises similar to those in the text. The remaining two problems were worked in week 14, just two weeks prior to the final exam.

The attributes of the four problems (see Appendix 1) are summarised in Table 1 below.

Table 1 - Summary of the Problems

Attributes	Problems 1	2	3	4
Similar to test book exercises	Yes	No	Yes	No
Open-ended	No	Yes	Yes	Yes
Hypothesis test expected	1-sample (mean)	2-sample (means)	corr/ regr	corr/ regr

Problem 1 was a typical text book exercise. Problem 3 was also typical, except for the fact that what was required was not clearly defined. Interpretation was left to the student. As such, it was "open-ended". Problems 2 and 4 were similarly open-ended, but they lacked numerical information.

RESULTS

In the study group, approximately eighty-six percent gained a pass or better. The average was 62, just below a credit award. In the sample of eight students, six passed (four passes, two distinctions) and two failed. The problem solving attempts by the latter group are summarised below for each problem. "P" refers to a student who passed the subject, "F" refers to a student who failed.

Problem 1 was recognised by all eight students as being similar to problems that had been worked before. They immediately summarised the given information with its statistical notation. All students recognised and stated that this question involved hypothesis testing, and set out to perform the steps highlighted for such problems in lectures and tutorials. Null and alternate hypotheses were given correctly in five cases (3P, 2F). Only three students (3P) used the correct test statistic (formula). Hereafter the responses differed. At one extreme, one student (1P) became confused, stating "I'm thinking of five different ways" and continued no further.

At the other extreme, another (1P) proceeded thoughtfully through the six-step procedure, explaining his choices, retracing his steps when his expected answer was not satisfied, then following through again to a final conclusion. Between these extremes, all other students (4P, 2F) were able to work through the exercise and make a conclusion of rejection of the null hypothesis. Four of these (3P, 1F) went beyond this rejection and actually answered the question. Only one (1P) looked at the numerical information for guidance in interpretation.

A number of different approaches were suggested to solve Problem 2. Even the student (1P) who could not complete the previous problem was fluent in explaining her strategy, but she stated that "This [problem] is great with no numbers...You can use judgement. That may be more beneficial than averages and stats." One notable finding in this problem was that only four students (3P, 1F) suggested the need for a statistical test, two of whom nominated a t-test, the appropriate test. By contrast, the other four students based their answers on qualitative judgement.

No student could offer a complete solution for Problem 3. One student (1P) could volunteer no information whatsoever. Another (1P) could only suggest (incorrect) hypotheses and a test statistic. Another (1P) tried to apply the chi-square test, the one topic he had revised, for 10 minutes, without success. The remaining three pass students (3P) wrote hypotheses related to correlation or regression analysis (demonstrating some knowledge of the problem type) and explained the direction of the relationship. All concluded that they had not revised this section yet, and could go no further.

Only one student (1P) attempted to state hypotheses for Problem 4, but like his attempt at Problem 3, his use of the chi-square test was incorrect, and his efforts ended in confusion. The student (1P) who could provide no information in the previous (quantitative) problem stated that "probably it's something to do with regression and correlation. It's the only chapter I've not looked at", but she could go no further. Another student (1P) could not even start, and suggested trying this exercise after the exam. The other three pass students (3P) explained the expected relationship between fuel consumption and weight, and stated immediately that this problem was concerned with regression analysis and correlation.

DISCUSSION AND CONCLUSION

While the examination results were pleasing, the individual performances emphasised problems in the understanding of statistical concepts.

Firstly, with a typical text book example, student attention was focused on the mechanics of the solution, and not on the analysis of the real situation. For example, in Problem 1, most students could use the six-step method to obtain a conclusion, yet only half could interpret this result, and only three could use the correct statistical technique. Hence, the emphasis on hand calculations and "getting the mathematics done" appeared to limit understanding, and lessen the ability to evaluate. The statistical knowledge acquired was likely to be content knowledge. The second type of knowledge had not been acquired.

Secondly, there were indications of poor study habits. For example, the text book question attempted late in the semester (Problem 3) could not be completed by anyone in the sample. Only half could suggest the topic to which this question related. Given the comparatively better performance in the earlier text book example (worked just after the midsemester test) one can only conclude that students postponed study of this subject until just before assessment. Such a strategy is likely to encourage rote learning and cramming as coping strategies, with little time for the development of deep understanding. As Keefe (1988) notes, learning this way requires constant practice and special cues for recall: the distinguishing features of the information are easily lost. Hence, the long-term benefits derived from studying this way would be minimal.

Thirdly, there was limited evidence of the transfer of statistical knowledge to novel problems. For example, only half the students suggested using statistical techniques to solve Problems 2 and 4. For the remaining students, there was no knowledge transfer from the current statistical subject they were studying, nor from the topics they had studied for the midsemester exam. Given their difficulties in selecting statistical techniques in Problems 1 and 3, this result is not surprising. A sound statistical knowledge foundation did not exist in the first place. Hence the higher order skills which Perkins and Salomon (1987) implied would facilitate transfer would be underdeveloped. Factors such as over-emphasis on mechanical calculations, poor study habits and rote learning (mentioned previously) would be intertwined in explaining this underdevelopment.

Now to the question: Is a pass good enough in statistics?

This study highlighted the rudimentary understanding students have of statistics, regardless of whether or not they pass; it illustrated the sketchy knowledge of key concepts just prior to the exam, and the limited skills in knowledge transfer.

Clearly it is time to consider the goals of introductory statistics courses. If the aims are to acquire the mechanical skills of statistical analysis, then this study demonstrates that students can achieve this through last minute cramming. However, research literature suggests the benefits may not be long-lasting. If the goals are in fact broader, and involve understanding, acquisition of higher order thinking skills, appreciation of the relevance and utility of statistics, then the traditional method of presentation and emphasis is not the way to achieve these.

In an era of mass tertiary education, there is a need to change the teaching and learning emphasis in statistics. New strategies are needed. For example, greater use of the computer for all statistical problems, including exam problems, would de-emphasise hand calculations, freeing time for reflection and experimentation (see Lock, 1987). This would increase understanding, and perhaps even lessen the dissatisfaction with introductory statistics subjects. The statistical data presented in the various media could be analysed and interpreted to ensure contextual relevance, thus facilitating transfer (see Kapadia & Andersson, 1987). Even if this approach were no more effective in terms of pass rates, the opportunity to see statistics in context would be advantageous. Students' perceptions about statistics may be enhanced and that may be critical to their decisions about its future usage.

REFERENCES

Anderson, J.R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.

Davidson, R. & Swift, J. (Eds.). (1987). Proceedings of the second international conference on teaching statistics. Victoria, BC: University of Victoria.

Grey, D.R., Holmes, P., Barnett, V. & Constable, G.M. (Eds.). (1983). Proceedings of the first international conference on teaching statistics, (vols. 1-2). Sheffield: University of Sheffield.

Grimmer, M. (1990). A learner-centred approach to teaching statistics. Auslink, 13-14.

Kapadia, R. & Andersson, G. (1987). Statistics explained: Basic concepts and methods. Chichester, West Sussex: Ellis Horwood Ltd.

Keefe, J.W. (1988). Profiling & utilizing learning style. Reston, VA: National Association of Secondary School Principals.

Lawson, M.J. (1984). Being executive about metacognition. In J.R. Kirby (Ed.). Cognitive strategies and educational performance (pp. 89-109). Newcastle, NSW: Dept of Education.

Lock, R. (1987). The role of computers in statistics instruction. Proceedings of the second conference on the teaching of statistics, (pp.87-95). Oneonta, New York: State University of New York.

Long, E. (1989, November). The role of domain-specific knowledge in classroom problem-solving. Paper presented at the AARE conference, Adelaide.

Moore, D.S. (1988). Should mathematicians teach statistics? The College Mathematics Journal, 19(1), 3-7.

Morley, C.L. (1992, July). Rediscovering the management statistics curriculum. Paper presented at the ANZAM conference, Sydney.

Perkins, D.N. & Salomon, G. (1987). Transfer and teaching thinking. In D.N. Perkins, J. Lochhead & J. Bishop (Eds.). Thinking: The second international conference. Hillsdale, NJ: Lawrence Brothers Assoc.

Taffe, J. (1991). Panel Session: Teaching statistics - mathematical or practical model. In R. Davidson & J. Swift (Eds.). Proceedings of the second international conference on teaching statistics (pp. 332-336). Victoria, B.C., Canada: University of Victoria.

Vere-Jones, D. (1991). Proceedings of the third international conference on teaching statistics (vol. 2). Voorburg, Holland: International Statistical Institute.

Wise, S.L. (1985). The development and validation of a scale measuring attitudes towards statistics. Educational and Psychological Measurement, 45, 401-405.

APPENDIX 1

Problem 1: A particular brand of wine contains on average 750 mls per bottle. To check the quality control, a sample of 36 bottles was randomly chosen from a consignment of wine, and the contents of each bottle were measured. Each bottle was found to contain an average of 746 mls with a standard deviation of 2 mls. At the 5% level of significance, does this sample indicate that the wine level is now different from 750 mls?

Problem 2: The manager of a firm is recruiting students from your University, who, like yourself, have some background in statistics and have completed this subject. He wishes to find out whether students enrolled in an Accounting degree are as good at statistics as those enrolled in a Management degree. How would you suggest he investigate this?

Problem 3: Peter Gordon is frustrated because he believes that his low student evaluations are due to his large classes (e.g., he can't learn the students' names, and there is less discussion). He samples 10 classes in the School of Business in order to test his hypothesis.

Class Size X	Class Rating of Instructor Y
10	4
20	4
20	5
30	4
50	2
30	3
40	4
40	3
40	2
50	3

a) Write down the hypothesis warranted by the data.
b) Investigate the problem and draw the conclusion warranted.

Problem 4: The advertising agency for a car manufacturer is interested in the answer to the question: Does the weight of a car relate to its fuel consumption? If you were in the employ of the agency, and, given that you have now completed this statistics subject, how would you investigate this question?