NUMERACY AND STATISTICAL REASONING ON ENTERING UNIVERSITY

Therese Wilson and Helen MacGillivray
Queensland University of Technology, Australia
tm.wilson@qut.edu.au

The relationship between mathematics and statistical reasoning frequently receives comment (Vere-Jones 1995, Moore 1997); however most of the research into the area tends to focus on mathematics anxiety. Gnaldi (2003) showed that in a statistics course for psychologists, the statistical understanding of students at the end of the course depended on students' basic numeracy, rather than the number or level of previous mathematics courses the student had undertaken. As part of a study into the development of statistical thinking at the interface between secondary and tertiary education, students enrolled in an introductory data analysis subject were assessed regarding their statistical reasoning, basic numeracy skills, mathematics background and attitudes towards statistics. This work reports on some key relationships between these factors and in particular the importance of numeracy to statistical reasoning.

INTRODUCTION

Each year, hundreds of students enrol in introductory statistics courses at Australian universities: some by choice, many as a compulsory component of a chosen program. Over the past fifteen years, the study of statistics has received increasing emphasis at the school level, with Chance and Data now forming a component of all Australian primary and junior secondary curricula. Yet it is not only the formal Chance and Data knowledge with which a student embarks on tertiary study. A number of components such as basic numeracy, mathematical inclinations, attitudes towards statistics and self-efficacy (perception of own ability) may also impact on a student's ability to develop statistical thinking.

Current pedagogical thinking acknowledges that students construct their own knowledge by combining current experiences with existing understanding. Hence, to better understand the development of students at the interface of secondary and tertiary education, we need first to determine and understand their relevant background factors, including attitudes and preparation, which may need to be acknowledged or adjusted before they can successfully progress in their statistical understanding.

METHODS

Students involved in this study were enrolled at the Queensland University of Technology in the subject "Statistical Data Analysis 1" during the first semester of 2004 or 2005. The study involved 389 students in 2004 (briefly reported on in Wilson and MacGillivray, 2005) and 301 in 2005 who completed a number of instruments during the first week of semester. Between 65 and 70% of enrolled students completed each of the instruments. Approximately 57% of these students were enrolled in an Applied Science degree (most in Life Sciences), 15% in a Mathematics degree and 9% in an Education degree. The remainder were in a range of courses having some connection with science. Background information was supplied by the students regarding previous mathematics subjects studied (at high-school or elsewhere), the results obtained in these and also their tertiary entrance score.

A multiple-choice and short answer Statistical Reasoning Questionnaire (SRQ), consisting of 20 items in 2004 and adjusted slightly to 21 items in 2005, was designed for the study. This questionnaire was informed by the Statistical Reasoning Assessment (SRA) (Garfield 2003) and the work by Jane Watson in Australian schools (see, for example, Watson and Callingham, 2003). Questions were developed and selected to reflect the statistical thinking considered appropriate at the Australian school/tertiary interface and necessary for development in a wide range of disciplines. Specifically, questions addressed the five aspects of reasoning about data summarised by Garfield (2003): representations of data, statistical measures, uncertainty, samples and association. A few questions on particular aspects of the high-school curriculum were included on the questionnaire but not in the analyses reported here, which include only the 15 non-curriculum specific questions common to the two years. Students' numeracy was assessed by a questionnaire consisting of 21 multiple-choice items, designed around those aspects of numeracy commonly assumed in an introductory data analysis subject, such as: use and understanding of fractions, decimals and percentages; substitution of numerical values to evaluate simple expressions; and rearranging simple equations and inequalities. So as to enable very simple questions to be asked, students were asked not to use calculators. As the items on this questionnaire assess skills at or below a standard relevant to approximately 14 to 15 year olds in Queensland, and the stated assumed knowledge for the data analysis subject is an algebra and function-based core senior secondary mathematics course, the results of incoming students on this questionnaire are of interest in themselves. In 2004, students' attitudes towards statistics and their self-efficacy in the area of mathematics and statistics were measured via a Likert scale. Items from the survey were grouped to measure attitudes in relation to: affect (feelings engendered by statistics); value in the subject; motivation to study statistics; links perceived between mathematics and statistics; use in society; difficulty of the subject; self-efficacy. Due to time constraints and the need to maximise student cooperation, only the self-efficacy component of the attitudes survey was repeated in 2005, as other components had contributed negligibly to the information obtained in the previous year. Details and full analysis of the questionnaires and data are reported in Wilson and MacGillivray (2006a, 2006b).

RESULTS

The statistical reasoning questionnaire (SRQ) highlighted both strengths and weaknesses at this school/tertiary interface, including some student misconceptions of relevance to the development of school curriculum. One example of student misconceptions can be found in two questions related to reasoning about statistical measures, with responses demonstrating confusion regarding the role of the mode in samples from quantitative data. Over the two years, 9% of students believed that the sample mode was an appropriate estimator of the "true value" and 19% believed that the mean was also the most common value. This reflects the problems caused by the persistent grouping in junior school syllabi, despite advice, of mode with mean and median as "measures of centre."

In the area of reasoning about uncertainty, 45% of students demonstrated an outcome-approach to probability (Konold 1989) by claiming that a weather prediction giving a 70% chance of rain was very good if it rained on 85%-100% of days for which such a prediction was made. This result is comparable to that (46%) reported by Konold (1995).

In the area of reasoning about samples, students were asked to decide which, if any, of three samples demonstrated greater than expected variation from a theoretical set of proportions. One sample was significantly different to the theoretical model, while the other two were well within acceptable variation. Forty-four percent of students correctly identified the single unusual sample; 37% felt that none of the three samples threw doubt on the theoretical proportions; 12% chose the extreme sample together with others; while the remainder chose samples other than the truly anomalous sample. This suggests that most incoming students have an understanding of sample variation, but that the extent of acceptable variation is difficult for them to assess.

When asked to choose between a two and three dimensional pie chart for inclusion in a report, 30% of students preferred the three-dimensional version, though only 8% gave a "prettier is better" explanation for their choice. Other students who preferred the three dimensional version indicated that their choice was based on the size or their perception of clarity of the graph, possibly reflecting the popularity of the three-dimensional pie chart in public documents.

The distribution of results and ordering of questions from that found least to most difficult in the basic numeracy questionnaire were so similar between 2004 and 2005 that the two cohorts were almost indistinguishable. Details and full analysis of the numeracy questionnaire are reported in Wilson and MacGillivray (2006a). The analysis includes subjecting the results of the basic numeracy questionnaire to a Rasch analysis to assess the value of the questionnaire and to better understand the levels of numeracy assessed by the items. The Rasch analysis provides evidence of a high level of construct validity and acceptable levels of reliability. It is also seen that questions can be regarded as reflecting five levels of numeracy which could be used to break the questionnaire into two components: introductory numeracy (levels one to three) and intermediate numeracy (levels four to five). The component into which a question falls is determined by its level of complexity, with questions in the intermediate component generally being distinguished by consisting of multiple steps and/or requiring a synthesis of ideas. For example, the questions involving percentages are represented in level one (the lowest level) through to level four with sample problems shown in Table 1.

Table 1: A range of questions involving percentages

Level	Abbreviated question	Success rate
1	5% of 200 students receive a grade lower than D. How many is this?	96%
2	Express 1/6 as a percentage	78%
3	How many students in a class of 200 receive an A grade if 5% receive a grade lower than D and 85% receive a grade of B to D?	69%
4	20% of 80 males and 30% of 120 females is what percentage of the mixed group?	52%

The analysis included the use of general linear models to determine which variables influenced scores on the statistical reasoning questionnaire. A considerable number and range of variables are available for inclusion in the analysis: numeracy total; tertiary entrance score; years since school; self-efficacy; senior mathematics result; gender; completed higher level mathematics or otherwise; repeating or non-repeating student; first semester student or otherwise; mathematics student or otherwise; year (2004 or 2005). Other variables that might have been considered in such a study tended to be difficult or unreliable to measure, as well as increasing the complexity of correlated predictors. From these variables, all possible main effects and 2-way interactions were included in the initial model and backwards elimination was used until all remaining variables were significant at the 5% level. Wording of the background information surveys in 2004 led to some non-reporting of lower tertiary entrance scores in 2004. This wording and the associated selective non-reporting were avoided in 2005, but the inclusion and interpretation of tertiary entrance scores in any analysis in 2004 or when the two years are compared or combined requires considerable caution and care.

For the two years combined, and without tertiary entrance score, the best model for scores on the SRQ involves as significant predictors only the numeracy total ($p<0.001$), self-efficacy ($p=0.011$) and gender ($p=0.039$), and has an adjusted R^2 of 20%. When the two components of numeracy suggested by the Rasch analysis are separated, both introductory numeracy ($p<0.001$) and intermediate numeracy ($p=0.003$) are significant predictors of statistical reasoning. However, many of the variables which are not significant in explaining SRQ scores are significant predictors of the numeracy total, with the best model involving senior maths result ($p<0.001$), mathematics student or otherwise ($p<0.001$), gender ($p=0.002$), completed higher level mathematics ($p=0.003$), self-efficacy ($p=0.003$) and year ($p=0.035$), with an adjusted R^2 of 28%.

Including tertiary entrance score in the analysis, whether for the two years combined or separate, does not diminish the importance of numeracy total as a predictor for the statistical reasoning. In the case of considering the two years combined, the model becomes more complex with interactions between numeracy and other variables becoming significant, and the interpretation is made very difficult by the unrepresentative nature of the subset of the 2004 cohort who did report their tertiary entrance score. When the years are considered separately with tertiary entrance scores included, the separate models are less complex and more interpretable within the circumstances of each year's reporting, but in every case, numeracy is highly significant ($p<0.001$) no matter what other variables are included.

IMPLICATIONS FOR SCHOOL AND TERTIARY CURRICULA

The numeracy questionnaire is of a level most appropriate for the end of junior high school, but inclusive of essential mathematical foundations. Its ongoing use assists students in identifying any weaknesses in their mathematical background that may impede (or indeed may have already impeded) their quantitative confidence. The majority of the students in the study completed a senior algebra and function based mathematics course at high school, and many of the students demonstrated a high level of competence with the basics of the numeracy survey. However the results of the numeracy survey indicate the importance in school curricula of consolidating and reinforcing skills and understanding in handling percentages, fractions, ratio, inequalities, order of operations, simple algebraic thinking, and, in particular, situations requiring more than one step or combining more than one skill. Students who do not complete the above senior course do not tend to receive such consolidation, and non-algebraic, non-calculus senior high school courses need to address these issues.

At the introductory tertiary level, the surveys and results demonstrate the importance of detailed knowledge of the diversity of student backgrounds entering introductory statistics courses in order to be accurately informed of the prior numeracy skill level of the students. Such detailed knowledge is necessary to develop strategies for dealing with diversity in incoming numeracy, while maintaining course standards and the interest of more capable students. The Statistical Reasoning Questionnaire is designed to evaluate students' statistical thinking at the end of schooling which, over the past 10-15 years, has included significantly increased exposure to chance and data. Results of the questionnaire help to demonstrate that mere exposure to data and some statistical terminology does not necessarily develop statistical thinking.

The Statistical Reasoning Questionnaire requires very few calculations and none that would be classified as requiring intermediate level numeracy. SRQ scores, however, were highly dependent on both the introductory and intermediate components of numeracy even after other background variables had been allowed for. The ability to think and reason statistically is clearly related to a student's numerical ability. While we do not claim that improving numeracy will necessarily improve statistical understanding, nor that an introductory statistics course should over-emphasise numeracy, attempts to teach statistical thinking without acknowledging the links with basic mathematical skills are likely to flounder at best, and at worst impede development of statistical thinking.

REFERENCES

Garfield, J. (2003). Assessing statistical reasoning. *Statistics Education Research Journal*, 2, 22-38.

Gnaldi, M. (2003). *Students' Numeracy and their Achievement of Learning Outcomes in a Statistics Course for Psychologists*. Unpublished M.Sc Thesis, University of Glasgow, Faculty of Statistics.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, 6, 59-98.

Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education,* 3.

Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, 65, 123-137.

Vere-Jones, D. (1995). The coming of age of statistical education. *International Statistical Review*, 63, 3-23.

Watson, J. and Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, 2, 3-46.

Wilson, T. and MacGillivray, H. L. (2005). Numeracy counts in the statistical reasoning equation. Presentation at *55th Session of the International Statistical Institute*, Sydney Australia.

Wilson, T. and MacGillivray, H. L. (2006, in preparation). Counting on the basics: Mathematical skills amongst tertiary entrants.

Wilson, T. and MacGillivray, H. L. (2006, in preparation). Statistical reasoning and confidence on entering tertiary study.