# Retention of knowledge between statistics courses: results of a pilot study

Alice Richardson

University of Canberra, Canberra, ACT, 2601, AUSTRALIA
*Alice.Richardson@canberra.edu.au*

**Abstract**

This paper reports on a pilot study of knowledge retention from the first to subsequent statistics courses. The first aim of the study was to assess the ability of a Statistics Concept Inventory (SCI) to measure the statistical knowledge of students after their first Statistics course. The second aim was to use the Inventory to assess students' retention of material from an introductory Statistics course into a subsequent statistics course. The usefulness of the SCI was partially supported, as many questions were better answered by the more able students. Retention was found to vary across the topics in the SCI.

*Keywords:* Statistics Concept Inventory, Statistics education

## 1. Introduction

What happens to students' statistical skills between their first course and subsequent ones? The aim of this paper is to report on work in progress at the University of Canberra to study knowledge retention.

This first statistics course has already been the subject of a large body of research in the last few decades. One of the most recent outcomes has been the GAISE report (Aliaga et al. 2005) setting out recommendations for the teaching of introductory statistics.

The first course in other scientific disciplines has also received attention. Zhang and Lidbury (2006) identified difficulties with language as contributing significantly to problems students experience in studying Genetics at the University of Canberra (UC). They identified ten strategies from the teaching of foreign languages, and implemented a selection of them in a Genetics classroom with much success. Zhang, Lidbury, Bridgeman, Yates, Rodger, & Schulte (2008) applied the same principles to first year Biology, Chemistry and Physics at a number of Australian universities. A selection of language teaching strategies was also implemented in Introduction to Statistics at the University of Canberra (UC) (Richardson and Zhang 2008a, b). The stategies are given in Table 1. This paper reports on initial attempts to assess retention of knowledge from that revamped introductory statistics course to subsequent courses.

Concept inventories to measure student learning in a first course exist in a number of disciplines too, including Physics (Yeo 2001), Biology (Khodor, Halme, & Walker, 2004; Michael, 2007; Smith, Wood, & Knight, 2008) and Chemistry (Bowen & Bunce, 1997; Krause, J. Birk, R. Bauer, B. Jenkins, & Pavelich, 2004; Michael, 2007). The Statistics Concept Inventory (Stone et al. 2003) consists of 25 questions, concentrating on descriptive and inferential statistics with a few

graphical and probability questions (Table 2). The topics appear in a haphazard order in the SCI.

However, research into the second course in statistics varies. Many textbooks assume that the second course will concentrate on regression and cover that topic in a fairly traditional manner e.g. Mendenhall and Sincich (1996). Others, such as Love (1998), employ a much broader project-based course. This project aims to forge a link between the two bodies of research, discovering the nature of retention between the first and second courses, and the success of current tools at measuring this retention.

Table 1: Language teaching strategies implemented in Introduction to Statistics

| |
|---|
| 1. Providing a list of terms and placing these terms in relation, in groups. |
| 2. Using warm up activities such as matching terms to definitions for revision purposes. |
| 3. Using online language exercises such as crosswords, gap-fill (Cloze) exercises. |
| 4. Using of flashcards for vocabulary revision. |
| 5. Providing stimulus questions for lecture and tutorial materials online, thus encouraging students to prepare before the lecture. |

Table 2: Distribution of topics in the Statistics Concept Inventory

| Topic Area | Number of questions |
|---|---|
| Inference | 7 |
| Probability | 3 (plus 2 omitted) |
| Descriptive | 8 (plus 1 omitted) |
| Graphical | 4 |
| Total | 22 (plus 3 omitted) |

## 2. Experiment

The aims of this pilot study are as follows. We want to know if the SCI has captured student learning, such as is measured by final grades, well. We are also interested in whether SCI questions discriminate well between students overall, and whether the different categories behave differently in terms of their discriminatory ability. Finally, we are interested in how well material is retained from first semester to subsequent semesters

In semester 1, 2009, the SCI was administered to a group of 27 Introduction to Statistics students at UC in the last two weeks of semester. This unit is an introductory statistics unit with 4 hours per week contact for 13 weeks. The questions were re-ordered to follow the order of topics in the course, and split into two sets of 11. This is because the SCI was being used primarily as a revision tool in the context of weekly online quizzes consisting of about 10 multiple choice questions. Furthermore, students were offered two chances to complete the test, the same conditions as their other weekly quizzes. The results from the first attempt are used here. There is a chance that some students would have answered without thinking, merely to access the feedback for each question.

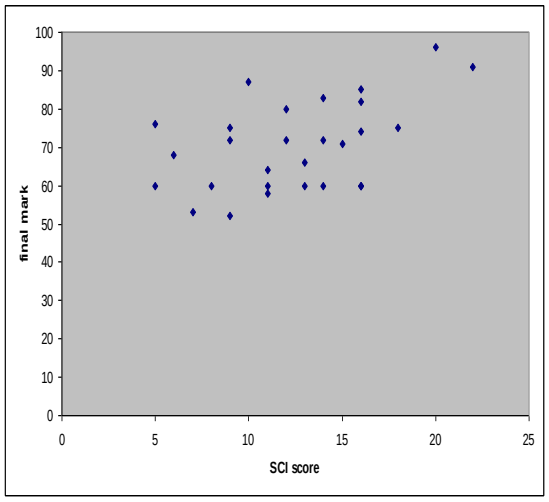
In semester 2, 2009, the 22 questions were again administered, in their original order and in one paper-based test, to 4 Nonparametric Statistics students. This course represents the only later-year course in this semester. Students participated on a voluntary basis. Two of them were graduate Statistics students, and had already studied three or more other stats units. One student came directly from Introduction to Statistics in semester 1, and one was a graduate Nutrition student who had studied introductory statistics at another university.

## 3. Results

First we want to know whether the SCI has captured student learning well. The correlation between SCI score and final mark of the 27 Introduction to Statistics students was 0.49 ($p = 0.005$), suggesting that there is a moderate but

significant association between increased SCI score and increased final mark. See Figure 1.

Figure 1: Scatter plot of SCI score and final mark




Second, we are interested in whether questions discriminate well between students overall. In particular, is there a significant difference in mean final mark between students who get a particular question right and students who get a particular question wrong? There were only 8 significant results (with $p < 0.05$) out of 22. Even without a correction for multiple tests, it does not therefore appear that these questions discriminated well between students.

Third, do the different categories behave differently in terms of their discriminatory ability? Discriminatory ability is defined as a significant difference in mean exam mark for those who got a question right compared to those who got a question wrong. Figure 2 shows the t test statistics of the 22 SCI questions by category. An analysis of variance produces $F = 3.01$ ($p = 0.058$). This suggests that with the quantity of data we have, it is not possible to tell for sure whether the different categories behave differently in terms of their discriminatory ability.

Fourth, we are interested in how well material is retained from first semester to subsequent semesters. There was one student from Introduction to Statistics in semester 1 who studied Nonparametric Statistics in semester 2. She scored 22/22 in semester 1, and 21/22 in semester 2. The incorrect one was an Inference question. Table 3 shows the results by category for the other 3 Nonparametric Statistics students. It appears that descriptive concepts are well retained by the students, along with inference. Graphical concepts and probability are not as well retained. It should be noted that these graduate students are not necessarily in their second statistics course at this point, but may be in their sixth or seventh.

Figure 2: Discriminatory ability of SCI questions by category

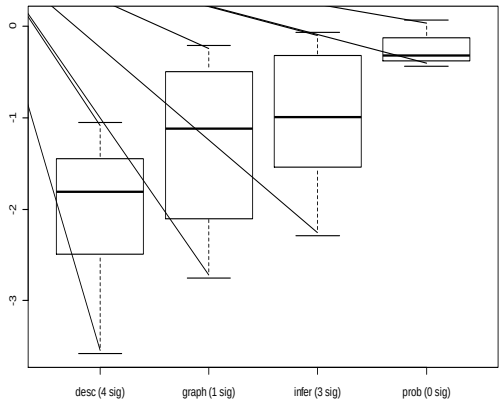



Table 3: Number correct by Nonparametric Graduate students in different categories

| Student | 1 | 2 | 3 |
|---|---|---|---|
| Inference (7 qns) | 3 | 4 | 4 |
| Probability (3 qns) | 0 | 1 | 1 |
| Descriptive (4 qns) | 1 | 2 | 2 |
| Graphical (8 qns) | 5 | 6 | 8 |
| Total (22 qns) | 9 | 13 | 15 |

## 4. Discussion

A distinction should be made between retention of knowledge from the first course to subsequent ones, and capstone courses. A capstone Statistics course, using a book such as Spurrier (2000) or Chatfield (1996), aims to integrate knowledge from many previous courses, not just one. Typically such a course is case study-based, involving a mixture of basic techniques such as exploratory data analysis with advanced techniques such as generalised linear and multilevel models.

Future work will naturally concentrate on collecting more data, separating the graduate and undergraduate data, and standardising the administration of the SCI. More data should permit adjustment of results for factors such as time between the first course and the course in which the SCI is administered, or the number of courses between the first course and the one in which the SCI is administered.

## References


Aliaga, M., Cobb, G., Cuff, C., Garfield, J. (Chair), Gould, R., Lock, R., Moore, T., Rossman, I., Stephenson, B., Utts, J., Velleman, P., and Witmer, J. (2005). *GAISE College Report.* [online] http://www.amstat.org/education/gaise/GAISECollege.pdf.

Bowen, C. W., & Bunce, D. M. (1997). Testing for Conceptual Understanding in General Chemistry. *The Chemical Educator* 2. New York: Springer-Verlag.

Chatfield, C. (1995). *Problem Solving: A Statistician's Guide*. London: Chapman and Hall.

Khodor, J., Halme, D. G., and Walker, G. C. (2004). A hierarchical Biology concept framework: a tool for course design. *Cell Biology Education* **3**, 111-121.

Krause, S., J. Birk, R. Bauer, B. Jenkins, & Pavelich, M. J. (2004). Development, Testing, and Application of a Chemistry Concept Inventory. Paper presented at the 34th ASEE/IEEE Frontiers in Education Conference, Savannah, GA, October 20 – 23.

Love, T.E. (1998). A project-driven second course. *Journal of Statistics Education* [online] **6(1)**. http://www.amstat.org/publications/jse/v6n1/love.html.

Mendenhall, W. and Sincich, T. (1996). *A Second Course in Statistics: Regression Analysis*. Upper Saddle River, NJ: Prentice Hall.

Michael, J. (2007). Conceptual assessment in the biological sciences: a National Science Foundation-sponsored workshop. *Advances in Physiology Education* **31**, 389-391.

Richardson, A.M., Lidbury, B.A. and Zhang, F. (2007). One potato, two potato, three potato, four: the use of *Hot Potatoes* software in science language comprehension. *Uniserve Science Symposium Proceedings, Science Teaching and Learning Research including Threshold Concepts*, Sydney 27 – 28 September, 2007 p83 - 88.

Richardson, A.M. and Zhang, F. (2008a). Language support for Statistics learners. Paper presented at OZCOTS 08, Melbourne July 3 - 4, 2008.

Richardson, A.M. and Zhang, F. (2008b). Activating multiple senses in learning Statistics. *Uniserve Science Symposium Proceedings, Visualisation and Concept Development*, Sydney 3 – 4 October, 2008 p98 - 103.

Smith, M. K., Wood, B. W., and Knight, J. K. (2008). The Genetics Concept Assessment: a new concept inventory for gauging student understanding of genetics. *CBE-Life Sciences Education* ***7***, 422-430.

Spurrier, J.D. (2000). *The Practice of Statistics: Putting it all Together*. Pacific Grove, CA: Duxbury.

Stone, A., Allen, K., Rhoads, T.R, Murphy, T.J., Shehab, R.L. and Saha, C. (2003). The Statistics Concept Inventory: pilot study. 3rd ASEE/IEEE Frontiers in Education Conference, November 5 – 8, Boulder CO.

Zhang, F., and Lidbury, B. (2006). It's all foreign to me: learning through the language of genetics and molecular biology. *Uniserve Science Symposium Proceedings, Assessment in Science Teaching and Learning.* Sydney. p153-159.

Zhang, F., Lidbury, B., Bridgeman, A., Yates, B., Rodger, J., and Schulte, J. (2008). Language difficulties in first year Science - an interim report. *Uniserve Science Symposium Proceedings, Visualisation and Concept Development*, Sydney 3 – 4 October, p159-164.