Concept Mapping and Learning in Tertiary Statistics Classes

Lyn Roberts & Beverley Moriarty
University of Ballarat & Central Queensland University

Presented at joint ERA & AARE Conference:
Educational Research: Building New Partnerships,
Singapore, 25-29 November, 1996

Concept maps have been used extensively in science education both to promote and to measure meaningful learning. This study examines their use in measuring tertiary science students' understanding of fundamental concepts in statistical inference. A class of nineteen students enrolled in a second unit of statistics were asked to draw concept maps from lists of terms associated with the definition of a statistical problem or investigation, and with techniques of statistical inference. Each of the two maps was drawn early in the semester and then again at the end of semester. The maps were assessed using a scoring scheme which was adapted from a number of other schemes presented in the literature.

Pre and post scores on the two statistical concept maps were examined to determine changes over the semester. They were also correlated with scores on the assignment, which was the formal assessment for that unit. While there was no significant improvement in map scores over time, significant correlations were found between scores on the initial problem definition map and the assignment, and also between the assignment and the second statistical inference map.

The classroom exercises involving concept mapping encouraged a different approach to cognitive processes. They are a worthwhile contribution to teaching and learning in tertiary statistics courses.

Introduction

This paper describes an experiment on the use of concept maps in a tertiary statistics class. Students drew maps to illustrate their understanding of the relationships between terms used in two areas of statistics, namely defining a statistical problem and hypothesis testing. Quantitative and qualitative evaluations of the maps were used to measure the extent of the students' understanding of fundamental concepts, and to assess changes in this understanding over the semester.

Many authors (e.g. Boland, 1995; Garfield & Ahlgren, 1988; Singer & Willett, 1990) have described the difficulties associated with teaching service courses in statistics. Watts (1991, p.290) writes that the problem which inhibits the learning of statistics as opposed to other disciplines such as physics or biology is that "the important fundamental concepts of statistics are quintessentially abstract". Garfield (1995a, p.31) says that "teachers should not underestimate the difficulty students have in understanding the basic concepts of probability and statistics". Concept mapping may assist students to develop this conceptual understanding, as well as providing teachers with another way to assess it.

The use of concept maps in science education developed in the 1980s, and was given wide dissemination through the book "Learning How to Learn" (Novak & Gowin, 1984). A concept map is a diagram intended to illustrate the understanding of the relationships between concepts involved with a particular area of study. A list of words describing important aspects of a topic is assembled. The words are arranged so that similar terms are near each other, and may be sorted into a hierarchy from most general to most specific. Links are then drawn between the concept words, and statements written to describe or explain the links. The map is intended to demonstrate the level of deep understanding of the map maker through the illustration of connections between concepts.

The technique of concept mapping has been used extensively in high school science classes (Jegede, Alaiyemola, & Okebukola, 1990; Lehman, Carter, & Butler, 1985; Roth & Roychoudhury,. 1992). Recently some statisticians have attempted to incorporate the techniques into their tertiary statistics classes in an attempt to assess students' conceptual knowledge (Lipson, cited in Garfield, 1995b; Williams, 1995). Williams has given detailed descriptions and analysis of maps of concepts of hypothesis testing drawn by six students. Quantitative evaluation of these maps was limited to a counting of the number of statements given by each participant. This paper presents a more detailed scheme for scoring such maps.

If maps are to be used for quantitative evaluation, for example for the purposes of assessment, there needs to be a valid and reliable system of giving a score to a map. Many different scoring schemes are to be found in the literature (e.g. Harnisch, 1994; Malone & Dekkers, 1984; Mason, 1992; Novak & Gowin, 1984; Stensvold & Wilson, 1990; Stuart, 1985; Wallace & Mintzes, 1990). The existing published schemes were not easy to use in the context of maps of statistical concepts. An alternative scoring method is proposed. This new scheme was used to score student maps drawn early in the semester, then again after several weeks of instruction. Map scores were also correlated with scores on the formal assessment for the unit.

Methods

The group of 19 students selected for this study were enrolled in a second unit in statistics which was specifically designed for biology students. These students had all passed a general introductory unit in statistical inference for science students which involved four hours per week of class time for one semester in the first year of their university course. All participants were told at the outset of the purpose of the research, and if they agreed to participate were asked to sign a consent form. While they were aware that participation in the concept mapping aspect of the classroom process was not compulsory, they saw it as a useful adjunct to their learning and, at least in the early stages of the project, all participated willingly.

In the first week of the course students were given an introductory lecture on concept mapping. As part of the work to be handed in for the first week, they were asked to draw a concept map using a list of terms taken from the first chapters of a statistics text. The terms, which were seen to be fundamental to the definition of a statistical analysis, were:

Categorical, Continuous, Data, Discrete, Experiments, Interval, Measurement, Nominal, Numerical, Ordinal, Population, Qualitative, Quantitative, Ratio, Sample, Sampling unit, Statistics, Survey, Variable.

The maps were returned and discussed in class. Some misconceptions had been displayed, and a group discussion attempted to clarify them. This was a valuable classroom exercise. The teacher / researcher was able to identify and attempt to clarify some points about fundamental statistical concepts upon which some students were obviously confused. This aspect of the study is discussed in another paper (Roberts, 1996).

Two weeks later, there was an appropriate point in the lecture sequence for a second concept map to be set. The students had revised a range of statistical tests, so a list of terms connected with aspects of statistical inference was given:

Accept, Alternative hypothesis, Anova, Assumption, Categorical, Chi-square, Correlation, Difference, Hypothesis test, Independent, Mann-Whitney, Mean, Normal, Null hypothesis, Numerical, One tail, Paired, P-value, Regression, Reject, Relationship, Samples, Significance, T-test, Two tail, Variance, Wilcoxon.

The assessment for this part of the unit involved a statistical analysis of data collected on a biology camp. Students had to decide on the biological phenomenon to be investigated, to plan the sampling and data collection techniques, and then analyse the data. The concepts set for both the problem definition and the inference technique maps were directly relevant to the assignment.

For the concept maps to be a valid indication of understanding of concepts necessary for the assignment task, there should be a strong correlation between assignment scores and map scores. Students were asked to redraw both maps at the end of the thirteen week semester, in order that changes in understanding could be measured.

Scoring Concept Maps

The initial problem definition maps were used to investigate several of the scoring schemes described in the literature, and to develop a scheme which is suitable for this subject. Scoring schemes which have occurred in the literature can be divided into two basic types. The first counts the occurrence of features such as propositional links, branching points or levels of hierarchy, allocates a point score to each, and adds them up. The second rates the maps against specific criteria. Novak and Gowin (1984) proposed counting the number of valid propositions, the number of valid levels of hierarchy, the number of valid crosslinks, and the number of examples given. Hierarchy levels and crosslinks scored five points each, propositions and examples one point. This method caused some difficulties with the statistics maps. While only valid propositions are counted, there is no way of distinguishing between students who make many incorrect statements such as "Discrete is qualitative", and those who have not put words to the links at all, but have demonstrated some understanding by the way the map is structured. Mason (1992) rates maps on each of six criteria: number of concepts, focal concepts, validity of linkages, number of linkages, horizontal versus vertical flow, and semantic categories of links. Harnisch (1994) also rates maps against a set of criteria. Yet another scoring scheme advocated by Stuart (1985) combines aspects of counting features and rating against criteria.

The most common problem with the scoring of statistical concept maps was in dealing with incorrect propositional links or links with no words. A lesser problem was how to rate misconceptions in hierarchy levels. The scoring scheme proposed for statistical concept maps gives a score for each of five features, based on specified criteria. It places considerable weight on the propositional statements, but also allows some credit for links which seem logical although unlabelled. As these maps have been drawn from a list of terms provided, there is no advantage in counting the number of terms or propositions on the map. However, as students can include extra terms or omit some they do not understand, there is a component of the score for terms used. Specific criteria are given for each score for all of the components. Examples were seen to be less important than correct propositions and hierarchy levels, since very few students included them, but they can earn a maximum of two points. Since there are two ways of scoring the links, namely lines and statements, the maximum score for completely correct linkages is 8. Because there is no uniquely correct hierarchy for these concepts, hierarchical levels were seen as less important than the understanding of the meaning of the terms as displayed through the links. Hierarchies then are scored out of 6. This gives a total score out of 20. Detailed criteria for scoring are given in Table 1 below.

Table 1: Proposed scoring scheme for Statistics concept maps

Feature	Detailed specifications	Point score
Terms	Given list plus extras	4
used	All on list	3
	Left out 1 or 2	2
	Left out 3 to 5	1

	Left out >5 terms	0
Linking lines	All logical connections	4
	One or 2 invalid or missing connections	3
	3 to 5 invalid or missing connections	2
	Half of connections invalid or missing	1
	Majority of connections incorrect or missing	0
Propositions on links	All present and correct	4
	Most statements correct	3
	Most statements present, several wrong	2
	More than half links have missing or inct statements	1
	No statements given	0
Hierarchy	Logical flow from general to specific	6
	One or 2 minor misconceptions in the hierarchy	5
	One major misclassification in hierarchy	4
	One major and 1 or 2 minor misclassifications	3
	Several misconceptions in the hierarchy	2
	About half of levels incorrect	1
	Majority of incorrect levels	0
Examples	Examps given for terms on lowest hierar level	2
	Some examples given	1
	No examples given	0

Results

The first concept map on problem definition was done by 18 of the 19 students in the class. These maps were independently scored by the researcher and one other experienced teacher of statistics using the proposed scheme. A correlation of 0.81 was obtained between the two sets of scores. This is comparable with the results of Mason (1992) who reported an interscorer reliability of 0.8 for her scoring scheme.

Scores were obtained for two maps, problem definition and statistical inference, each done both early in the semester and at the end of the semester. Unfortunately not all students did all maps. There was a distinct decrease in the response to the maps set at the end of semester. Only nine students completed both problem definition maps, and nine completed both statistical inference maps. (They were not the same students completing all four maps.) The students realised that the second maps did not contribute towards their assessment, and so neglected them in favour of more pressing assessable tasks. The students who completed the second maps had a wide range of scores on their first maps, so it does not appear that they self-selected on the basis of ability.

The scores on the maps constructed at the end of semester were lower than those done at the beginning. Paired t-tests showed the differences to be not statistically significant ($p=0.43$ for problem definition map; $p=0.056$ for inference map). One student commented on the second problem definition map that it was more difficult to do when

the topics had not been recently discussed in class. While teachers would hope that they have engendered deep and lasting learning, that does not seem to have occurred in this case.

Between the first and second attempts, students worked on a practical assignment which involved aspects of problem definition, the need to identify the correct statistical procedures to use, the ability to carry out and interpret the results of those procedures, and the ability to present a coherent report. The concept maps were thus relevant to two of the four components of the assignment. Students were encouraged to obtain feedback from the lecturer during the planning stages of the assignment work. While several of the students had very little idea initially on how to define the sampling units and variables for their project, most were able to clarify their thoughts as they worked through the project. Most students checked in advance that they had chosen the correct statistical method to use, and so there was very little discrimination in the marks for that section. The range of marks for the assignment was mainly related to differences in skills of interpretation and presentation.

There was a moderate correlation between scores on the first problem definition map and on the assignment ($r=0.42$, $p=0.042$, $n=18$) and also a high correlation between the assignment scores and the second statistical inference map ($r=0.77$, $p=0.008$, $n=9$). This perhaps indicates that those with a clearer initial idea on how to define a statistical problem could cope better with the assignment. Those who performed better on the assignment then had a better understanding of the relationship between the concepts associated with different aspects of statistical inference.

Conclusion

It had been hoped that the maps drawn at the end of semester would demonstrate the extent to which student understanding had increased, and hence reassure the teacher that the objectives of the unit had been achieved. Unfortunately this was not the case. The students did not perceive the final maps to be relevant to their assessment in the unit. Many students either did not complete the final two maps, or did not seem to put in as much effort as they had with the earlier maps. In future studies some extra incentive will need to be provided to encourage full participation. Perhaps having a second map drawn some time after the initial one, but not so close to the end of semester, may give a more accurate assessment of the extent of any improvement.

Some aspects of the concept map scores showed significant positive correlation with the formal assessment for the unit. The topics for the concept maps were directly related to two of the objectives of the assignment, namely the ability to formulate the problem for a statistical investigation, and to decide which statistical test to use. Because many of the students sought feedback on these aspects during the preparation of their assignments, the discrimination in marks came more from other aspects of the assignment than from the subjects of the concept maps. The significant positive correlation between scores on the first problem definition map and the assignment was based on the largest sample size of all used in the study. This first map was the one taken most seriously by the majority of the class. That those who scored better on this map which displayed their understanding of how to define a statistical problem tended to score higher on the assignment, reinforces the argument that concept maps can be used to assess understanding.

Concept maps have been shown to be a useful adjunct to other classroom methods in the teaching of statistics. Misconceptions can be identified and corrected, and students are encouraged to thing about the fundamental concepts of the subject. The scoring method proposed for these maps can also contribute to formal assessment in a way which complements traditional assessment methods. Concept mapping is a technique which should continue to be used in university statistics courses.

References

Boland, P.J. (1995). Discussion on the symposium on the teaching of statistics in higher education in Ireland. The Statistician, 44(1), 55-58

Garfield, J. (1995a). How students learn statistics. International Statistical Review, 63,25-34.

Garfield, J. (1995b). Research report: Research papers at ICOTS4. Teaching Statistics, 17(1), 29-31.

Garfield, J. and Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: implications for research. J.Res. Math. Educ., 19(1), 44-63

Harnisch, D.L. (1994). Concept mapping approach and its applications in instruction and assessment. presented to Annual Meeting of the American Education Research Association (AERA) and National Council on Measurement in Education(NCME), New Orleans.

Jegede, O.J., Alaiyemola, F. F., & Okebukola, P. A. (1990). The effect of concept mapping on students' anxiety and achievement in biology. Journal of Research in Science Teaching, 27(10), 951-960.

Lehman, J. D., Carter, C. & Butler, J. K. (1985). Concept mapping and achievement: results of a field study with black high school students. Journal of Research in Science Teaching, 22(7), 663-673.

Malone, J. & Dekkers, J. (1984). The concept map as an aid to instruction in science and mathematics. School Science and Mathematics, 84(3), 220-231.

Mason, C. L. (1992). Concept mapping: A tool to develop reflective science instruction. Science Education, 76(1), 51-63.

Novak, J. D., & Gowin D. B. (1984). Learning How to Learn.

Cambridge: Cambridge University Press.
Roberts, L. (1996). Using concept maps to measure statistical understanding. presented to Sydney International Statistics Conference, Sydney, Australia.
Roth, W-M. & Roychoudhury, A. (1992). The social construction of scientific concepts or The concept map as conscription device and tool for social thinking in high school science. Science Education, 76(5), 531-557.
Singer, J. D., and Willett, J. B., (1990). Improving the teaching of applied statistics: putting the data back into data analysis. The American Statistician, 44, 223-230.
Stensvold, M. S., & Wilson, J. T. (1990) The interaction of verbal ability with concept mapping in learning from a chemistry laboratory activity. Science Education, 74(4), 473-480.
Stuart, H.A. (1985). Should concept maps be scored numerically? European Journal of Science Education, 7, 3-8.
Wallace, J. D., & Mintzes, J. J. (1990). The concept map as a research tool: Exploring conceptual change in biology. Journal of Research in Science Teaching, 27(10), 1033-1052.
Watts, D. G., (1991). Why is introductory statistics difficult to learn? and what can we do to make it easier? American Statistician, 45(4), 290-291.
Williams, A. (1995). The concept map in statistics: assessing conceptual knowledge. presented to Australian Association for Research in Education Conference: Directions: Yesterday, Today, Tomorrow. Hobart, Tasmania.