Abstract

A computer-simulated learning module has been developed for the teaching of experimental design, the sampling of populations of organisms, and data analysis and presentation, concepts that students from both science and non-science backgrounds often find difficult. Furthermore, traditional pedagogical methods for these concepts are often logistically difficult and therefore expensive, or restrictive, uninteresting and irrelevant to students. This interactive module incorporates text, video, audio, graphics and shockwave multimedia to simulate experimental design and sampling, and is coupled to Excel for data analysis and graphical presentation of derived parameters. The project was funded by a SIF grant, and has been successfully implemented in two First Year Biology units in the School of Biological Sciences, Monash University.

Keywords

Computer-simulated methods, Pedagogy, Statistical education

Introduction

First Year Biology is a major service course for second year biology subjects in the Faculties of Science and Medicine. A key challenge is to address the educational needs of a broad range of students from a diversity of language and education backgrounds. Part of the first year curriculum involves the teaching of basic experimental design and proficiency in the use of Excel for statistical analysis and graphical representation of summary data. Statistical literacy in biology is important for a number of reasons, including those related to the reading and evaluation of scientific papers and in the design and analysis of biological experiments (Fowler, Cohen and Jarvis 1998). There are numerous excellent texts that cover experimental design and data analysis (e.g. Sokal and Rolhf 1999; Quinn and Keough 2002), and others that deal with the general use of computing for the learning and teaching of statistics.

Whilst simulation software has been developed that assists students in learning the core concepts underlying statistical inference using pre-prepared data sets (e.g. delMas, Garfield & Chance 1999), methods of developing statistical literacy by means of simulated field trip scenarios which incorporate both experimental design and decision-making with regard to the timing and location of biological sampling, as well as analysis of the resultant data, appear to be largely unavailable in the teaching of undergraduate biology. This is despite the underlying acceptance of the value of computer simulation methods (CSM) in the teaching of statistical concepts (Mills 2002).
A computer-based, interactive module has been developed for the teaching of principles related to experimental design, sampling and quantitative methods in biology. The structure of the module is based upon an existing practical session that had been taught successfully for several years but which did not encompass issues regarding design or decision-making normally associated with field studies, and therefore offered little scope for flexible learning or integration of current technologies into an interactive environment.

The module operates at a number of pedagogical levels, the most important criteria of which are:

1. the stimulation of problem-based learning (PBL);
2. an introduction to the formulation and testing of scientific hypotheses;
3. the simulation of realistic field conditions and an understanding of the problems associated with biological sampling and data gathering in the field;
4. the development of skills in experimental design;
5. an understanding of the importance of measuring and interpreting variation in biological data;
6. instruction in data analysis and the appropriate presentation of quantitative biological information;
7. generation of suitable assessment material, including answers to questions, graphs, etc.

The high degree of interactivity of the module, which incorporates graphics, audio, video and text, enables students to be exposed to a more realistic situation of sampling a population of organisms. The module is also an effective tool for the teaching of statistical analysis of biological data, which have been traditionally derived from a dry laboratory practical where sampling has already been carried out and which are therefore unrealistic and lacking in relevance to actual environmental conditions.

**Design**

Module design commenced with formation of a project planning team comprising academic staff from the School of Biological Sciences, which compiled and submitted a proposal for a Monash Strategic Innovation Fund (SIF) grant. This was followed by an extensive literature search and discussions with other academic staff and multimedia professionals to establish the framework and locate suitable developmental tools with which to develop the module. A project outline was established and staff employed to carry out video production, graphic design and the writing of text. Integration of video, text and other materials was carried out by a multimedia professional using HTML and Shockwave software.

The development of this module conforms with the strategic directions of the Faculty and University as outlined in the Learning and teaching Operational Plan, the Science Ahead strategy for 1998-2002 and Leading the Way: the Monash Plan 1998-2002.

**Implementation**

The module is based upon investigations of the differential impacts of sewage on soil nutrient loads and the abundance of an invertebrate species at sites adjacent to an outfall pipe (Fig. 1). The structure of the module simulates that of field-based experiments and surveys, with an introduction, field and laboratory work and data analysis and presentation. Students are provided with class notes which detail pre-exercise preparation, key terms, the module URL, and instructions for loading the Analysis toolpack for Excel®. The notes also provide references for further reading and web links to statistics-related sites.

Students are guided through the module in a structured fashion, beginning with the rationale and aims of the survey, information about sewage generation and treatment and background detail regarding the species to be sampled during the exercise. Pages are numbered and students are able to progress through the module, reading text and accessing graphical, audio and video supplements to the text. Key words are hotlinked to a glossary providing further explanation and formulae where appropriate. A ‘Contents’ key tracks progression through the module, and a student is able use this facility to quickly revisit a page without having to scroll back manually through all previous ones.
Using postulated conditions related to the position of the outfall, movement of tides and prevailing flow of current, students randomly sample twenty cores within each of three locations (unaffected, mildly affected and strongly affected) on the shore (Fig. 2). Students are prompted to consider sample size, the location of each sampling area and whether to sample at high or low tide, which will affect the time available for sampling. These decision-making processes are based upon previously-obtained information about the ecology and distribution of the species in question. Sample data are generated by a random grid selector for each sampling site, within which are embedded values of worm abundance and percentage soil nitrate. The data are downloaded to Excel when sampling of all sites is complete.

**Figure 1.** Module template for the measurement and sampling practical.

**Figure 2.** Image of generated sample collections at each site at the simulated outfall. At each site, each white square represents a single sample of worm count and % nitrate.
Students then use Excel to carry out statistical analysis and graphical presentation of their generated data, instructions for which are detailed at each step in the procedure (Fig. 3).

Figure 3. Example of data layout and instructions for data analysis.

Assessment questions are embedded at specific points within the module, and students are prompted to consider each question in light of the relevant activity or following data analysis. Student answers to these questions, together with the sample data and summary figures generated from Excel are copied into the final page of the module (Fig. 4) and automatically returned via email for assessment and grading.

Figure 4. Template of the final page of the module, within which students copy raw data, site statistics, answers to questions and summative figures.
Results

Postgraduate students in the School of Biological Sciences carried out initial evaluation of the module, and provided feedback on its format, structure and learning outcomes. Based on their suggestions, minor modifications were made to some text and assessment components.

A trial of the module was conducted in 2001 using 420 1st year Biology students. In its first year of use, student learning was assessed via a 15 minute miniquiz (an invigilated, computer-mediated assessment task), using a combination of multiple choice and text-based questions and incorporating images where required. The pedagogical rationale for the interactive module is quite different to the ‘wet lab’ previously used for the teaching of the descriptive statistics, and as a consequence, each employed a different set of assessment questions. Thus, it was not possible to directly compare student outcomes. Nevertheless, it is highly likely that by completion of the module, students gained or increased their understanding of concepts related to experimental design and biological sampling and proficiency in the use of spreadsheet programs such as Excel® for data analysis and presentation. Further, student grades for computer-mediated assessment of the module were very strongly correlated with their final grade for the Unit ($r = 0.55, P < 0.001$). The overall mean grade for students who undertook the module (Table 1) was also within the range obtained for other assessment tasks in the Unit: weekly miniquizzes, $\bar{y} = 63.3 \pm 1.5$; practical reports, $\bar{y} = 64.1 \pm 2.4$; essay, $\bar{y} = 68.3 \pm 0.6$.

Students scores for questions related to descriptive statistics and variation were, in almost all cases greater than 50%, with a high degree of variation in the correlation (the ‘Discrimination’) between the question score and overall test score (Table 1). Negative discrimination values for some questions indicate a positive correlation between getting the question correct and getting a poor score for the test i.e. the bottom part of the class answered such questions better than the top part of the class.

<table>
<thead>
<tr>
<th>Question</th>
<th>% Correct (n=420)</th>
<th>Discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>The hypothesis formulated for the worm counts sampled in the Measurement and Sampling Prac was that the number of worms in samples from highly affected sites would be “greater than” those from unaffected sites.</td>
<td>56.0</td>
<td>0.97</td>
</tr>
<tr>
<td>The distribution in the figure to the right is “negatively skewed”.</td>
<td>50.1</td>
<td>0.44</td>
</tr>
<tr>
<td>If a distribution is positively skewed, the mean will be “greater” than the median.</td>
<td>50.4</td>
<td>0.12</td>
</tr>
<tr>
<td>The sample mean for a particular characteristic of a species is an estimate of the population “mean” for that characteristic.</td>
<td>57.2</td>
<td>-0.02</td>
</tr>
<tr>
<td>The expression: $\bar{y} = \frac{\sum X}{n}$ calculates the “mean” of a sample.</td>
<td>86.4</td>
<td>-0.16</td>
</tr>
<tr>
<td>The data on worm numbers collected in the Measurement and Sampling Practical is an example of “discrete”, categorical data.</td>
<td>47.7</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

Overall mean for computer-mediated assessment: $\bar{y} = 60.4 \pm 0.94$

Table 1. Student outcomes regarding questions related to sampling, variation and descriptive statistics. Text enclosed by inverted commas is the correct option or text to be typed. Discrimination is the statistical correlation between the question score and the overall test score.
Student responses to the module were in the main very positive (Table 2). Free-text feedback from students were varied and mostly related to the length of time required to complete the module (Table 2), and difficulties in accessing the videos. Some examples of positive comments included:

- “it was clear and concise, easy to understand”;
- “the instructions for constructing the bar graphs were excellent, as were the rest of the instructions”;
- “the practical was set out really well, it just took a long time to complete”;
- “this practical was very useful in improving computer skills and gave an insight into the possible practices of biology once I completed my course”;
- “I thought it was very good. It showed the real type of field work that is out there once we finish Uni”.

Some examples of less positive comments were:

- “this prac took far too long and assumed too much”;
- “the only negative thing about this prac is that it takes too long to do it and instructions should be given before we had to do the prac”;
- “the formulas given in the exercise should have been in the practical manual”;
- “the practical should be put into the context of what is discussed in recent lectures and practicals”.

<table>
<thead>
<tr>
<th>Question</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neither agree nor disagree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>This practical was easy to follow and well-structured</td>
<td>8</td>
<td>51</td>
<td>15</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>This practical helped me to develop knowledge of and skills in the use of spreadsheet applications such as Excel</td>
<td>9</td>
<td>48</td>
<td>19</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>This practical provided me with greater understanding of how biologists design and carry out scientific research</td>
<td>6</td>
<td>51</td>
<td>26</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>This practical provided me with a better understanding of how biologists analyse and present experimental data</td>
<td>6</td>
<td>60</td>
<td>24</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Duration</th>
<th>% Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1 hour</td>
<td>2</td>
</tr>
<tr>
<td>1-2 hours</td>
<td>12</td>
</tr>
<tr>
<td>2-3 hours</td>
<td>25</td>
</tr>
<tr>
<td>3-4 hours</td>
<td>27</td>
</tr>
<tr>
<td>&gt; 4 hours</td>
<td>34</td>
</tr>
</tbody>
</table>

Table 2. Student responses to specific questions regarding the structure and outcomes of the module.

Discussion

The mean grade for the computer-mediated assessment task, together with the very strong correlation between this and the overall Unit grade, suggests that it may be a more objective measure of student knowledge than the grade previously obtained for the wet lab, which was assessed by sessional staff using a standard marking schedule. Nevertheless, specific questions are required for testing student understanding of concepts related to experimental design and sampling, and their degree of competency in using spreadsheet software for data input and analysis and presentation of descriptive statistics. The low or slightly negative discrimination values for some questions suggest that they may require revision or rewriting.
The module currently takes the majority of students far too long, although this may also have resulted from the high variation of student expertise in using spreadsheet software such as Excel, as well by factors related to processor and modem speeds and Internet connectivity etc. Further editing of the module and its provision, together with associated software, on a CD-ROM should resolve some of these problems.

**Future Developments**

The immediate objective is development of a databased system of student delivery and tracking, incorporating quiz questions and results delivery. This will incorporate specific questions to determine student learning with regard to experimental design and sampling, and their competency in the use of spreadsheet software for data input and analysis, and graphical presentation of descriptive statistics. An intensive study will then be carried out to evaluate the effectiveness of the module in regard to these learning objectives.

Future plans include: (i) incorporation of a linked Contents table that will allow students to start or finish sessions at different stages, without having to recommence the module each time it is initiated; (ii) the development of modules that investigate other biological questions, including one that assesses the differential impacts of forest clearing on birds and small mammals and another that simulates the use of guinea pigs in laboratory experiments. The latter module will incorporate methods of analysis for data collected in the experiments and address ethical issues associated with the use of live animals for such purposes. In each case the sampling regimes and analysis will illustrate different problems and issues.

**Acknowledgements**

The module was originally conceptualised by Gordon Sanson and Gerry Quinn. The project was funded by a Monash University SIF grant. Project structure and timelines were devised by Gordon Sanson, Gerry Quinn and Martin Predavec. Bronwyn Burton carried out storyboard construction, including video production and text input. Programming by Clever Bits Pty Ltd with editing and revision by the authors.

**References**


Copyright © 2003 Gerry Rayner & Gordon Sanson

This article may be used by educational institutions and non-profit organisations for the purposes of research and teaching provided appropriate citation and reference are included. The article must not be modified. Any reproduction of this article must include this copyright statement. The copyright owners also permit Monash University to publish this article on the World Wide Web (prime sites and mirrors) and in CD-ROM format within the EET 2003 conference proceedings. Any other usage is prohibited without the express permission of the copyright owners.