Hypothesis Testing as a Core Component of a Virtual Experiment

Arianne Dantas  
Department of Physiology  
University of Melbourne, AUSTRALIA  
a.dantas@physiology.unimelb.edu.au

Robert Kemm  
Department of Physiology  
University of Melbourne, AUSTRALIA  
r.kemm@unimelb.edu.au

Debbi Weaver  
Higher Education Development Unit,  
Centre for Learning and Teaching Support  
Monash University, AUSTRALIA  
debbi.weaver@celts.monash.edu.au

Abstract  
We designed an interactive virtual experiment, requiring students to formulate hypotheses of expected outcomes of experiments, and then allowing them to test and review their hypotheses and receive immediate feedback. We report a trial of this approach, conducted as a paper-based exercise, to determine if the planned structure of the multimedia resource is an effective teaching tool, and provide some preliminary results from using the resultant multimedia module in a classroom. Our results are very promising as students appeared more engaged by the virtual experiments than previous teaching approaches. While students found forming relevant hypotheses to be a challenging task, there was a noticeable change in their attitude to learning and we hope that later analysis may show good learning outcomes from this encouragement of deeper learning practices in experimental investigations and consequently will encourage such enhancements of practical class teaching more generally.

Introduction  
Experimentally based scientific disciplines use practical classes for developing professional investigative skills, but it is not always possible to provide undergraduate students in large courses with experience in some areas because of cost and/or ethical considerations. Simulations are used widely as a replacement of such hands-on experience, but they may not always provide an adequate understanding of the real-world situations that arise in live experiments.

We have chosen to develop a virtual experiment to provide students with some of these experiences by the extensive use of video, provision of data files that can be played back and analysed as if recorded live with a data-recording system, and an embedded requirement for students to make hypotheses before undertaking various experimental investigations. This virtual experiment is designed to complement a multimedia module developed to simulate the control of a complex biological system (Weaver et al, 2000, 2002).

This case study is presented to inform others working in scientific disciplines by showing the process we used, from identifying a need, trialling the design in a paper-based study, to the development of a multimedia module and trialling its first use by students’ first use in classes. Although it is shown for physiology, the general principles of hypothesis testing and predictions, the use of video delivery and the ability of data play-back and analysis can all be widely used in biological, medical and physical sciences to engage students in authentic learning experiences.

Background  
Identifying a Need:  
To justify and obtain funding for this multimedia project it was essential to show a need for improved teaching that was not adequately fulfilled by existing approaches to learning. ‘Cardiovascular Responses’ is an integrated topic of Physiology drawing on background knowledge and understanding from several different body systems. It is an essential topic for all physiology and medical students to understand, but one that we consistently found difficult to
teach in an engaging and contextually-related manner. Traditionally, this topic was taught in a practical class, by surgical preparation of an anaesthetised rabbit, often with the surgery performed by students. A range of hormones and drugs were injected intravenously, and other physical interventions performed, and the animal's responses (usually heart rate and arterial blood pressure), were monitored.

At The University of Melbourne, we originally taught this class with 50-60 students working in small groups of 6 to 8, each with an experienced demonstrator. While the class was successful at teaching the experimental basis of cardiovascular control, not all students were fully engaged and the ability to integrate the findings with theory was very dependent on demonstrator skills and student involvement. In addition, it was expensive, labour-intensive, and difficult to staff effectively, given that some of the fine surgical skills are no longer required in Departments moving to more cellular physiology research. Also, the class was scheduled for 5 hours, and in recent years, the length of this class in a crowded curriculum became an issue. Over-riding other considerations was the ethical need to reduce the reliance on animal experimentation in our teaching.

We also tried providing a single anaesthetised animal as a demonstration and videoconferencing to individual workstations in a laboratory. Students were encouraged to come to the front of the class to view and discuss the experiment. This approach was largely unsuccessful, as few students interacted with the experiment, and most were at best, passive observers of the experiment from their workstation. The laboratory arrangement meant that the lecturer could make eye-contact with few students, and this teaching approach was soon abandoned.

We next tried a professionally made videotape of the entire experiment, with simultaneous recording of responses via PowerLab Chart™ software. Voiceovers were incorporated into the video to explain the procedures and observed results. This package was presented to students in a lecture theatre setting, with 40 to 200 students present. Academic staff would discuss each procedure, invite predictions from students of expected outcomes, play relevant sections of the video, and then discuss observed responses. This approach only engaged a small number of students present, and despite the best attempts by the lecturer, was generally a very passive experience for most students.

Concurrently with this trial, we developed an interactive CD-Rom tutorial (Weaver et al., 1999, 2000, 2002) to complement this practical class, by covering the key compensatory pathway involved in cardiovascular physiology, and extending this with a real-life example. The tutorial proved highly successful when run in a separate class, with students working in groups of 2-3 to discuss contextual and integrated physiological questions (Weaver et al., 1999, 2002). However it did not extend or apply the concepts into other real-life situations involving other cardiovascular interventions, and students did not consider challenges other than the example provided.

The success of this tutorial encouraged us develop an interactive program as a virtual experiment to cover the cardiovascular responses to the same pharmacological and physical interventions as in the original successful small-group practical class. This led us to a paper-based trial to first investigate the efficacy of such an approach to improve practical class teaching by facilitating active learning and encouraging constructivist learning. Such an approach has proved successful in practical class teaching at other institutions (Rodenbaugh et al., 1999). The computer is a key component of this approach when it is used as a cognitive tool (Jonassen & Reeves, 1996).

Aims and Objectives

Our long-term aim has been to develop an interactive computer program which will replicate a real experimental experience by allowing students to view high-quality video of all procedures, replay recorded cardiovascular responses in real time and directly measure important results from these screen tracings. Students are required to formulate and test their hypotheses of the cardiovascular responses to a variety of challenges or interventions. Formulation of hypotheses prior to observing results aimed to create a more engaging environment for the virtual experiment and prompt students to think at a deeper level about the physiological mechanisms occurring before and during the experimental procedures. A similar study has found that student learning is more effective, particularly in correcting any prior misconceptions, when students are encouraged to formulate predictions of experimental outcomes prior to conducting the experiment, rather than simply performing the experiment and recording outcomes (Modell et al., 2000).
Our first objective was to develop our ideas as a paper-based exercise, to determine the suitability of this hypothesis-testing approach as a learning exercise for our students, and to assess whether such an approach might achieve the desired learning outcomes. During this process, we used PowerPoint to storyboard the design as a basis of the future guidelines for the programmers to be contracted for the final multimedia module. Layout of screens for student hypotheses and interpretations was designed to be similar to what was to be incorporated in the multimedia module.

Our second objective was to use the results of the paper-based study to refine the PowerPoint storyboards and add screens suitable for showing Quicktime movies of all experimental procedures together with interactive Quicktime movies of the data recordings. The notes facility in PowerPoint was used to provide descriptions of all the behaviours required in the programming in each screen, as well as to indicate what feedback was to be provided to the students at each stage. We chose to deliver the program on an intranet with a Web-Browser using HTML coding, Flash and JavaScripting so that the authors could later revise some parts of the material with a Web authoring program without requiring expensive programming assistance.

Our third objective was to determine whether there were any differences in students’ attitudes to learning and levels of understanding when provided with this virtual experiment with its central theme of hypothesis testing.

Finally we wished to evaluate this approach to program development, using PowerPoint to provide detailed guidelines to programmers, as a prototype for use in three other funded projects. We believed this approach to be an effective way of adding extensive programming expertise to enhance storyboards produced by academics experienced with educational issues and multimedia development.

Methods

Our storyboarding of development required students to formulate their own predictions of experimental outcome, then view digital video of the procedures and results of the intervention, measure outcomes directly from embedded Chart™ recordings of responses and then revise their predictions accordingly. The movies were to be embedded in Web pages as Quicktime movies. The Quicktime movies of Chart™ recordings included a cursor on each trace that could provide the time and amplitude for any chosen point. The rest of the methods focuses on the initial paper-based exercise, since we consider this an important step in multimedia design before there is extensive commitment to expensive programming.

![Figure 1: Quicktime Video of Simultaneously Recorded Responses and Procedures](image-url)
The essential difference with students’ use of the multimedia module 12 months later was that they worked at their own pace, viewed the videos on their own computer, and received instant feedback on all their submitted predictions and interpretations, once their interpretation of each experimental procedure was submitted. Students played simultaneous videos of the data recording and the experimental procedure (Figure 1) and were able to analyse their own data by dragging the cursor to the appropriate section on the screen tracings to obtain numerical values. The students were also able to print out a record of their predictions and modifications after observing and measuring the results of the various experimental procedures and these were added to their written reports.

For the paper-based trial, two scheduled and assessed classes, each of 30 2nd year Science students, were conducted on consecutive days. At the beginning of each class, the academic in charge explained the process to the students, together with some introductory discussion about the surgical preparation involved.

The experiment was separated into sections, covering a discrete procedure or intervention. In each section, preliminary discussion explained the procedure involved, then students were asked to complete a prediction table or flow-chart in a workbook. These tables or flow-charts generally required students to arrange a selection of variables into the correct sequence of physiological events, and also predict whether that variable would be increased, decreased, or unchanged. An example of a prediction table is shown below as a screenshot from the multimedia module.

![Prediction Table and Flow Charts](image)

**Figure 2** Prediction Table and Flow Charts for Student Entry of Information Prior to Experiments

Students then viewed the videotape of the procedure and responses digitally projected for the whole class, and were also supplied with a graphical printout of relevant sections of the recorded responses (eg. the blood pressure and heart rate of the animal). Students were then given the opportunity to revise their predictions on a separate page in their workbook, and this was termed their final interpretation.

After the first procedure (prediction and final interpretation), students were taken step-wise through the logical processes in generating the prediction table or flow-chart by the academic staff for students who had failed to understand what was expected. At this stage, students were encouraged to ask questions if they still felt unsure about the process involved. After this first scenario, students were then asked to complete the remaining sections without any further feedback from staff about interpretations of results.

At the end of classes, student workbooks were collected and assessed by the authors. Photocopies were kept for evaluation, and the books marked and returned to students. All student decisions were recorded in a spreadsheet to allow for statistical analysis of the data.
The intention of this evaluation was to:
- allow us to identify possible student misconceptions and determine whether these misconceptions were corrected by the current approach.
- determine whether the tasks were at an appropriate intellectual level for our students

**Practical class report**
During the week following these sessions, students were required to prepare and submit a laboratory report, answering set discussion questions, which required explanations of some of the phenomena observed during this class. Students submitted this report before their prediction/interpretation workbook was returned to them.

**Results**

**Student responses to the Paper-based exercise**
At the start of the paper-based exercise, the students were completely bemused, and did not know where or how to begin. They were reluctant to predict outcomes of the procedures, and required a great deal of explanation and encouragement to think in terms of how the particular procedure or intervention would affect the heart rate and/or blood pressure, and how the cardiovascular system would compensate for this change.

Both groups displayed a great deal of difficulty understanding the terminology used in procedures where they were asked to decide whether particular drugs would cause more or less of a change in certain factors than those observed without the drugs present. The concept of observing a decreased increase in a variable was very difficult for the students to grasp. Accordingly, the terminology used in these questions has been changed to now use the words accentuated and diminished.

This was such a different approach to learning in a practical class, that there was a great deal of antagonism and complaints in the first class about the process involved and many students appeared indifferent. Overall students in both classes showed no more or less overall enthusiasm than any other practical class. Despite this, in the end of semester Quality of Teaching Surveys conducted by the Department they rated this class at 3.70 compared with the average score of all practical classes of 3.47 (5-point Likert scale) so that they must have found some value in the class.

**Analysis of student prediction and interpretation tables**
Student attitudes appear to be reflected in their results. While there was no significant difference in the performance of the 2 groups for the first half of the exercise, Group 1 students performed less well in the second half of the class. This result was in accordance with the observed attitudes of the students in Group 1, who did not seem to be applying as much effort to the tasks. Students in Group 1 were instructed not to worry too much about the prediction phase, but to try whatever they thought would be correct where as students in Group 2 were told to deliberate and attempt to complete the predictions to the best of their abilities.

Interestingly, at the end of each procedure, average student marks for the interpretation phase were regularly above 80% for each of the tasks set. For 5 out of the 7 procedures, student performance on the interpretation phase was significantly improved when compared with the prediction phase of the exercise. For the remaining 2 procedures, student performance for the prediction phase was much higher (compared with the other 5 procedures) and no statistical improvement was observed with the interpretation data. These 2 procedures were carried out immediately after the feedback provided by the academic following the first procedure, and were in a similar format to the first prediction flow-chart and covered similar physiological concepts. Subsequent procedures extended the theory covered, and incorporated different formats of predictions tables, and these changes may have created difficulty for some students.

**Practical class report**
The laboratory report submitted by each student during the week following this class was graded and the 2 student groups compared. Data was also gathered on these students’ performances in other practical classes for the remainder of the semester. For each practical session, students are required to submit a similar practical report, which are marked by their demonstrators and returned to the students.
Students from Group 1 performed statistically more poorly (p=0.03) with their written reports on this class that Group 2, but they also may have performed more poorly over all classes although the large difference could not be shown significant (p=0.07). Comparisons between each student’s performance in this class against their performance in other classes revealed no significant difference (p=0.30 and 0.71 for group 1 and 2 respectively). So although self-reported student satisfaction and our own observations indicated that students appeared to engage more with the subject material, their performance in post-prac questions did not reveal an improvement.

Preliminary Analysis of Student Use of the Multimedia Module

It is only possible to report some preliminary analysis of the outcomes of the use of the multimedia tutorial in classes. There is an extensive questionnaire to be analysed, observational notes to be collated and an analysis of an examination question testing their ability to predict an outcome of a novel experimental procedure based on the experiments undertaken in this module.

This article has focussed on the development of the module because of space limits. The detailed analysis of the evaluation and examination outcomes of the first use of the multimedia module will be discussed in a later article focussing on the learning of physiology.

Discussion

The Paper Based Trials

We were pleasantly surprised to note the high marks recorded for student interpretations (after viewing the experimental video), with most class averages being over 80%. Students were intellectually challenged during this class, but these results indicate that the tasks are achievable, albeit difficult for most of the class, and thus at an appropriate intellectual level for the target students. Students indicated they enjoyed this class, compared with other practical classes, as seen by the higher Quality of Teaching grade recorded for this session, and appeared to be engaged with the material during the session, even though they did express reluctance to think at a deeper level and expressed frustration at not understanding what was expected of them.

Students’ ability to formulate predictions improved for the first few procedures, but then fell back to lower levels for subsequent procedures. This coincided with a change in the format of the required response, but not in the formulation of hypotheses. We believe that the early improvement indicated students gaining familiarity with the layout and sequence of terms for the prediction flow-charts, rather than an improvement in their ability for formulate their predictions.

Observations and feedback from tutors indicated that students were unfamiliar with the subject material, and we believe a major difficulty in this area occurred in the timetabling of these sessions. While students had studied basic cardiovascular physiology in the previous semester, they had not yet attended the relevant lectures on control of blood pressure, which were scheduled for the week following this class. This may have accounted for some of the antagonism of the students, and for some of their reluctance to predict outcomes of the procedures. For the first use of the multimedia module, this class was scheduled later in semester, and so now appears after the relevant lectures, and importantly, after an interactive computer tutorial on the baroreceptor reflex (Weaver et al., 2000), which should alleviate the problem of lack of background knowledge and possibility improve student attitudes.

For most of the procedures studied, the interpretation mark was significantly higher than the corresponding mark for the prediction phase of the exercise. It is not possible from this study to determine whether this improvement derived from the process of forming a hypothesis and then testing this while watching the experimental video, or whether merely watching the video, without the prediction phase of the exercise, would have achieved the same results.

However, by comparing the two different groups of students, where one group was encouraged to take the prediction exercise seriously, but the earlier group was instructed to “just give it a try”, it can be suggested that if students do take this exercise seriously, their final interpretation marks are generally higher. This evidence is supported by the results from the submitted laboratory report, which asked students to further explain some of the observed phenomena. Students from the second group performed significantly better than the first group, indicating a greater understanding of the physiological processes involved.
However, at first sight, it might be tempting to conclude that taking this process seriously resulted in an improved learning outcome for students. Yet, the overall marks of the 2 groups of students for the other practical classes in their course suggested that Group 2 students may have exhibited a higher performance in all classes over the semester, although this was not statistically significant ($p = 0.07$), so no conclusions can be drawn about the different approach of the instructor, based on these practical marks. It is still possible that student understanding of the key concepts has improved, but this was not evaluated in this paper-based study.

**Preliminary Discussion of Student Use of the Multimedia Module**

The following is a brief description of the key features from observations and a brief perusal of the questionnaire data, prior to a more thorough statistical analysis that will be reported later in a more detailed article.

We found that use of Powerpoint for preparing flow-charts, screen designs and providing notes to guide the programmers was a powerful tool for the educational designers, now that we are out-sourcing our programming. We have found that Powerpoint is easily mastered by our academics and we are using this approach in our other multimedia developments. Our previous experience did allow us to present this information effectively for the programmers and they found also found this approach was very effective compared with the guidelines usually provided by their clients.

The majority of students found the module a very challenging and rewarding exercise and they took far longer to view the videos and consider their responses than in the paper-based trial and than in any form of teaching and learning in previous practical exercises on this topic. We are unable to say yet whether the students’ better grounding in the topic with prior use of the multimedia tutorial on cardiovascular control has improved their ability to make predictions and improve their interpretation of results compared with the previous paper-based exercise, but we hope detailed statistical analysis of their progressive submissions will reveal useful information. We thought that some students might be a bit put-off by the graphic nature of the surgical preparation videos, but most really appreciated how difficult the surgery was and that it gave them a much better appreciation of the nature of such experimental investigations.

An examination question was set in an open-book exam that was slightly more difficult than any of prediction tables and flow charts undertaken by students in this exercise. Preliminary indications are that many students obtained maximum marks, but more detailed analysis of the students’ answers (other than raw scores) will be necessary to separate their understanding of the relationships in flow charts and the predictions of directions of change. Some students also performed poorly and this perhaps indicates that this approach to learning and understanding does not suit everyone, although this is a difficult area of physiology and there are always students who show a very poor understanding of the topic.

**Conclusions**

We observed an increased engagement of students with the process of the virtual experiment when encouraged to earlier predict the outcome, compared with our observations of the passive experience, with little mental processing, in classes of similar students in previous years. Our data from the paper-based trial indicates that the process of prediction – experiment – interpretation may improve student consideration of the physiological mechanisms while conducting the experiment, and may assist in correcting student misconceptions. However, it appears important that students have an adequate grounding in the supporting theory prior to the practical classes to be able to make reasonable predictions in this complex cardiovascular control system. We are hopeful that detailed analysis of the student submissions in the trials with the multimedia module and examination performance will show some significant improvements in learning outcomes that reflect students’ greater engagement and discussion and possibly deeper approaches to learning. If so, our results would support the argument from Modell et al (2000), who found that encouraging students to formulate a hypothesis prior to conducting an experiment improved the students’ learning of the concepts involved.

**References**

CD-ROM

Journal Articles

Acknowledgements
This study was supported by an internal grant from the Teaching and Learning (Multimedia and Educational Technologies) Committee at The University of Melbourne, awarded to the authors.
The authors gratefully acknowledge the cooperation of ADInstruments (www.adinstruments.com) for assistance with converting Chart™ recordings from their data recording systems into appropriate formats for this program.