A critical goal of introductory science courses is to develop students’ understanding of key concepts and principles. Typically, introductory biology courses include two major parts—lecture and laboratory. In lab, students have opportunities to engage firsthand in scientific inquiry. However, students are exposed to the bulk of the subject matter in the lecture setting. Problems with the lecture format are well known. Instructors tend to present large amounts of information, and students tend to transcribe, but not process, what the teachers tell them. In large classes, students are inhibited from asking questions, and it is difficult for the instructor to be attuned to what students actually learn. Even when lectures are well organized, students tend not to be engaged with the subject matter in ways that lead to understanding (McKeachie et al., 1986; Cooper & Robinson, 2000; Gardner, 1994; Vernier & Dickson, 1967).

An increasingly popular way to deal with these kinds of lecture problems is to use “active learning” strategies to get students more involved in the class (MacGregor, Cooper, Smith & Robinson, 2000). Students report they like the class experience better than straight lecture, and that active learning exercises help them maintain their attention better throughout the class period (Bligh, 2000).

Active learning is a step in the right direction, but it does not guarantee that students will understand the subject matter. Students need to be more than active; they need to engage the subject matter in ways that lead to a deeper understanding of it. Research indicates that understanding is more likely to develop when students engage in activities such as analysis, evaluation, interpretation, prediction, and explanation (Bransford, Brown & Cocking, 1999; Coleman, 1998; Coleman, Rivkin & Brown, 1997).

In biology, we want students to do the same kind of thinking that scientists do when they try to explain biological phenomena, such as interpreting data, making predictions, and explaining phenomena. These are “sense-making” activities through which students can develop deeper understanding of the subject matter (Perkins, 1998; Chi, deLeeuw, Chiu & LaVancher, 1994). We have designed problem-solving modules for introductory biology lectures in which students engage in this type of thinking. In a typical module, students are presented with a set of data about a biological phenomenon. They develop an explanation to account for the data, and then get feedback from the instructor about their explanations.

In this study we examine the effects of a problem-solving module on student understanding of evolution and phylogenetic trees. In large lecture sections, students were presented with data about several animals and then worked in small groups to produce a phylogenetic tree to explain the data. The instructor collected and visually projected several of the models and analyzed them in front of the class. Instructor feedback focused on helping students develop their understanding of the concepts and revise any misconceptions of the material. This sequence was repeated two more times during the class period as students were presented with additional data.

Assessment revealed that these in-class modules resulted in significant improvement in student understanding.

**Methods**

**Course Background**

The module was tested during the fall 2003 semester in two different introductory biology courses, one for non-majors (BIO 103) and one for majors (BIO 105). Both courses are designed for freshmen, have no pre-requisites, and course sections range in size from 40-150 students. These are traditional entry-level biology courses with sections on Ecology, Cells, Genetics, and finally Evolution. We tested the effectiveness of the module by assessing 577 students in six different lecture sections taught by five different instructors. Several of the lectures were videotaped and monitored by external reviewers.

**Classroom Procedure**

Students in the Test Group were presented an in-class module with three iterations of problem solving in which they analyzed a set of data to develop a phylogenetic tree for...
a group of animals. The instructor first gave a brief lecture on phylogenetic trees and the history of evolutionary theory. In this lecture the instructor drew two trees on the board and showed a diagram of a third tree from the textbook. Next students were shown images of seven different animals (bear, sea otter, seal, porpoise, whale, hippopotamus, and penguin), along with information about each animal’s diet and habitat. Based upon these observations, groups of two to three students drew phylogenetic trees. Students were asked the question, Who is the whale's closest land relative?, using the phylogenic tree they generated. The instructor collected the diagrams and projected five to ten of these in class using a Samsung SVP6000 Document Camera. The instructor analyzed the trees and provided specific feedback to the class, pointing out key concepts and responding to misconceptions reflected in the diagrams. Next, the instructor gave a short lecture on anatomical evidence for evolution (e.g., analogous, homologous, and vestigial structures). The students were then shown the skeletons of the seven animals, focusing on the hind limbs and pelvis. The students revised their trees, and the instructor collected and projected five to ten of these, again noting any misconceptions to the class. The students were then given molecular data for the seven animals, and asked to make a final revision of their trees. The instructor then collected and projected the trees. The DNA evidence suggests that the whale’s closest land relative is the hippopotamus. Finally the instructor showed the students a recently-discovered fossil that is the whale/hippo link and guided the students through a discussion of convergent evolution with whales, sea otters, and penguins—each representing modern relatives of ancestral land animals that independently migrated to the sea.

In sections of the courses serving as Control Groups, students were taught using exactly the same PowerPoint slides and shown the same data. However, the Control students were not given the opportunity to analyze the data themselves. Instead, the instructor drew the correct answer on the board and explained the reasoning used to generate the answer.

Assessment of Student Understanding

Students reveal understanding to the extent that they can use knowledge appropriately to interpret, analyze, and evaluate new information and solve new problems. This view of understanding is consistent with contemporary research on human understanding (Perkins, 1998). We developed three instruments to evaluate student understanding.

A formative assessment tool was developed to measure students’ in-class performance on the module. Students’ diagrams of phylogenetic trees were scored by two of the authors (Cooper and Hanmer) using a rubric based on ancestry (e.g., no living species being ancestors to other living species), grouping (e.g., use of a logical scheme for classification), and accuracy (e.g., appropriate use of the chosen scheme).

Two summative assessment tools were developed to measure students’ ability to transfer their knowledge of evolution to new problems (see Appendix). A short answer question presented data on the diets and anatomies of five animals. Students were asked to draw a phylogenetic tree consistent with the data. In addition, multiple-choice questions were used to evaluate students’ understanding about what they are drawing when they make a tree. Five multiple-choice questions tested the students’ ability to group animals by physical attributes. Four multiple-choice questions tested student understanding of the underlying relationship of ancestry and evolution used to generate their trees. We developed two versions of each test, so that the students were not given exactly the same questions on the pre- and post-tests. The test was given as a pre-test before the unit on evolution had begun, and on the final exam, approximately two-three weeks after the module had been used in class.

Results

Formative Assessment

In the Test classrooms, groups of students were asked to draw phylogenetic trees in class, the drawings were collected, and scored using a standardized rubric that focused on Grouping, Ancestry, and Accuracy. In the first iteration, students tended to group the animals based upon either diet or habitat. Several common misconceptions appeared, such as living species serving as ancestors to other living species (Figures 1A and 2A). The trees were more accurate in the second iteration based on skeletal data, however, some misconceptions persisted (Figures 1B and 2B). By the third iteration based on molecular data, most trees were accurate, e.g., hippos were identified as the closest living
land relative to whales and porpoises (Figures 1C and 2C). Quantitative evaluation of the diagrams showed the same trend. Group scores showed a statistically significant improvement in all three criteria as determined by student t-test. The greatest gain was observed in Ancestry, in which 51% of the groups made at least one error in assigning modern species as ancestors to other modern species in the first iteration and only 7% made this error in the last iteration, for a total gain of 44%. Significant gains of 22% and 29% were also observed in Grouping and Accuracy, respectively (Figure 3).

**Summative Assessment**

The gains between individual student’s pre- and post-test short answer scores were analyzed by repeated measures of analysis of variance. Three predictor variables were taken into consideration (Question Topic: Grouping or Ancestry, Course: BIO 103 or BIO 105, and Treatment: Test or Control). The analysis revealed a significant three-way interaction effect on test score gain by all three predictor variables ($P = 0.017$). Students in Test lectures showed significantly greater gains than did students in Control lectures in the short answer question in all groups with the exception of the BIO 103 Ancestry questions (Figure 4).

The multiple choice scores were pooled by question topic (Grouping or Ancestry) and tested for difference in proportions. The gain in scores from Test lectures were statistically higher than Control lectures for BIO 105 Grouping and Ancestry questions, and for BIO 103 Ancestry questions. No significant gain was observed between Test and Control lectures for the BIO 103 Grouping questions (Figure 5).

**Discussion**

**Significance of Results**

Students in the Test lectures made significant progress during the lecture period in their ability to develop a model to explain data about the relatedness of animals. We were surprised that initially 51% of the students made errors in assigning ancestry to different species, even though they had been shown three diagrams of phylogenetic trees in a traditional lecture format just before they drew their own trees. However, after the students drew three trees and received formative feedback, only 7% of the groups made this error in their phylogenetic trees. It appears that students perform better after doing something three times than after just hearing it three times.

In analyzing the summative results it is important to keep in mind that students in both Test and Control lectures received exactly the same lecture material and raw data. The Test Group analyzed the data itself and drew its own diagrams; the Control Group was shown the correct answer by its instructor. Both groups showed significant gains in summative short answer scores. This suggests that the module works fairly well in a straight lecture format. However, when students in the Test group were given the opportunity to analyze the data themselves, make their own models, and receive feedback on their work in class, we saw a significant improvement in three of the four predictors being tested. The improvement was approximately 10%, which is equivalent to an entire grade on an exam. The summative assessment was two to three weeks after the lecture, suggesting at least some retention had occurred.

One could argue that practicing drawing diagrams in class made it easier for students to draw diagrams on an exam. However, gains were also seen in three of the four pre-
dictors being tested in the summative multiple-choice questions, some of which tested students’ understanding of the theory behind evolution. We found that the students could use existing logic skills to group animals, resulting in higher than expected pre-test scores on the Grouping questions, which could explain the absence of a difference between Test and Control in the BIO 103 Grouping questions. The largest gains and largest differences between Test and Control lectures were observed in the Ancestry questions. The Ancestry questions were more conceptual in nature, and required students not only to group similar animals, but to know how these groupings are explained by evolutionary theory. The significant increase in gain on these conceptual questions by students in the Test group is the strongest evidence that receiving formative feedback leads to a deeper conceptual understanding of evolution.

**Pedagogical Advantages of Using the Problem-Solving Module**

The problem-solving module benefits teachers in several ways. First and foremost, the module makes students’ understanding of the concepts open to observation and analysis in class. In traditional lectures, the instructor can only guess as to how students understand the subject matter. In contrast, when students work a problem-solving module they externalize their understanding of the concepts. Consequently, instructors can identify difficulties and misconceptions. Instructors can then respond—on the spot—to these problems, rather than discover them later on examinations. Moreover, instructors can give specific and pointed feedback to a large group of students during the learning process. In most large classes, if students receive feedback it is usually after they have taken a test and it is too late to facilitate their learning. In addition, by identifying areas that students either already know or can master easily, the instructor can change the content or format of the course to focus on areas of difficulty while spending less time on topics the students master easily.

The modules also benefit students in several ways. Working together in small groups facilitates student interactions and encourages the formation of friendships that make class more fun. As students work together on problems there is considerable discussion of concepts and peer instruction going on. These activities may spill over into a willingness to share questions on other topics. Finally, students seem to enjoy the change of pace and chance to see each others work afforded by the modules.

The main drawback of using such modules in large lectures is that it takes additional time in class for students to solve the assigned problem, and for the instructor to collect, project, and respond to their work. We found that the instructors took approximately 50 minutes to deliver the module, while those in the Control Group presented the material in 30 minutes. Because of this, we only use modules to introduce and strengthen what we consider to be crucial concepts that students need to master to understand a topic.

By integrating lecture with the problem-solving modules, students are exposed to the traditional content, while spending more time in class on the most difficult concepts. Both the majors and non-majors Test Groups did not perform below the Control Groups on any of the assessment tests, indicating that the modules did not interfere with their mastery of the other traditional content in the course.

In addition to the specific effects of the module on learning about evolution, we believe that the students also learn that scientific models must be modified as new data become available. In future work, we hope to show how the use of multiple problem-solving modules in lectures throughout a semester contributes to students’ scientific thinking.

In conclusion, there are several advantages to incorporating problem-solving assignments with formative assessment and feedback into large biology lectures. Students are more engaged in applying concepts from lecture to solving real problems. Students also receive critical feedback during...
the learning process, which supports the development of their understanding. An iterative process of student problem solving followed by formative feedback leads to a deeper conceptual understanding of difficult biological theories.

Acknowledgments

We would like to thank the following biology instructors at the University of Wisconsin-La Crosse for their patience and cooperation in using their classrooms as a laboratory: Anne Galbraith, Tim Gerber, Mark Sandheinrich, and Dan Sutherland. We would also like to thank the UWL Statistical Consulting Center run by David M. Reineke and his two student assistants, Susan Pederson and Craig Buettner, for their hard work analyzing our data.

References


ideally, we will place an ad here so as to try to avoid breaking up the appendix box.
Short Answer Question
Given the data in the table below, draw a diagram of an evolutionary tree indicating how the five species are related.

<table>
<thead>
<tr>
<th>Species</th>
<th>Canine Teeth</th>
<th>Preferred Food</th>
<th>Forelimb</th>
<th>Nails/foot</th>
<th>Skin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bat</td>
<td>Yes</td>
<td>Insect</td>
<td>Wing</td>
<td>5</td>
<td>Fur</td>
</tr>
<tr>
<td>Mouse</td>
<td>No</td>
<td>Plant</td>
<td>Leg</td>
<td>4</td>
<td>Fur</td>
</tr>
<tr>
<td>Robin</td>
<td>-</td>
<td>Insect</td>
<td>Wing</td>
<td>4</td>
<td>Feather</td>
</tr>
<tr>
<td>Rat</td>
<td>No</td>
<td>Plant</td>
<td>Leg</td>
<td>4</td>
<td>Fur</td>
</tr>
<tr>
<td>Shrew</td>
<td>Yes</td>
<td>Insect</td>
<td>Leg</td>
<td>5</td>
<td>Fur</td>
</tr>
</tbody>
</table>

1. The closest relative to the bat would be:
   a. Mice
   b. Rats
   c. Shrews
   d. Robins

2. The closest relative to the rat would be:
   a. Mice
   b. Bats
   c. Shrews
   d. Humans
   e. Robins

3. You would expect to observe the most DNA sequence similarity between a mouse and a:
   a. Bat
   b. Rat
   c. Shrew
   d. Human
   e. Robins

4. If a beaver has no canines and 4 toenails per foot, it would be most closely related to:
   a. Bat
   b. Rat
   c. Shrew
   d. Human
   e. Robin

5. Which types of data will produce the most accurate information about how species are related?
   a. Skeletons and DNA sequences
   b. Skeletons and feeding habits
   c. DNA sequences and feeding habits
   d. Habitat and feeding habits
   e. Habitat and DNA sequences

6. Shrews and bats have fur while robins have feathers. This suggests that:
   a. Only robins can fly.
   b. Shrews and bats have a more recent common ancestor.
   c. Robins and bats have a more recent common ancestor.
   d. Robins descended from bats.
   e. Bats descended from robins.

7. Humans are more closely related to shrews than mice. This means that:
   a. Humans descended directly from shrews.
   b. Humans and shrews have legs.
   c. Shrews and mice don’t have canines.
   d. Humans and shrews have a more recent common ancestor than humans and mice.
   e. Humans and mice have a more recent common ancestor than humans and shrews.

8. A branch point in an evolutionary tree represents:
   a. A living species that gave rise to a new species.
   b. An extinct common ancestor to species found on the branches.
   c. An extinct ancestor to just one of the species found on the branches.
   d. A specific mating between two different species.
   e. A time when natural selection did not occur.

9. The lines in an evolutionary tree generally represent:
   a. A single breeding between two members of the same species
   b. A single breeding between two members of different species
   c. Many generations leading to the formation of a new species
   d. One species changing into another species in a single generation
   e. A convenient way to indicate which species have similar traits, but are not necessarily related