

Cladogram Curriculum: A First Look Into A Teaching Prototype For The Fundamentals Of
Phylogenetics

Emily Schreiber

Vanderbilt University

Honors Research Thesis

Under the Supervision of Dr. Laura Novick

Vanderbilt University

April 7, 2010

Abstract

Past research has examined students' comprehension of and reasoning with evolutionary relationships depicted by cladograms (i.e., tree thinking). Cladograms serve as a fundamental resource when studying macroevolution and it is for this reason that students need to have a clear understanding of, and ability to reason with, these diagrams. Past research has found misconceptions and difficulties within the understanding of cladograms, particularly with students in upper level biology classes. This difficulty has left researchers calling for cladogram education to be introduced prior to upper level biology courses and perhaps as early as high school. This study created an early prototype of a curriculum in order to give students this basic cladogram education. The curriculum was effective in raising scores on previously tested skills for college students, with both stronger and weaker biology backgrounds. Although the curriculum was not effective for every tree-thinking skill tested, overall it led to higher performance for both groups.

Cladogram Curriculum: A First Look Into A Teaching Prototype For The Fundamentals Of Phylogenetics

Evolution is an important part of a biology education. For a biologist, evolution serves as the tool with which to organize all species. “Animals may be classified in some societies according to their usefulness or destructiveness to human endeavors. Others may group animals according to their roles in mythology. Biologists group animals according to their evolutionary relationships” (Hickman, Roberts, Keen, Larson & I’Anson, 2004). The concept of evolution is divided into two subcategories: microevolution and macroevolution. Microevolution looks at small changes to a specific species. Macroevolution looks at the entire history of life culminating in a single common ancestor for all species. Previous research has focused on microevolution (e.g., Anderson, Fisher, & Norman, 2002; Nehm & Schonfeld, 2008; Sandoval & Reiser, 2003). However, because of recent calls for teaching of macroevolution (Baum, 2005; Catley, 2006; Novick & Catley, 2007, 2009, 2010; O’Hara, 1997), my research focused on this important evolutionary biology concept.

Catley (2006) recently called for “a paradigm shift in evolution education so that both perspectives – micro and macro- are given equal weight, and perhaps even more importantly, that each is presented in such a way as to inform the other”. Catley made the case for this shift by explaining that a student must be able to understand macroevolution to grasp evolution as a whole. He presented the case that microevolution is taught to such a degree in schools that macroevolution is given little, if any, attention. Moreover, when instruction is given on these larger concepts, those giving the instruction often have a poor understanding of it themselves.

Basics of Cladistics

One of the main areas studied by evolutionary biologists is the relationship among species due to descent from common ancestors. The main goal of finding these patterns of common ancestry is to be able to reconstruct the Tree of Life. As the Hickman, Roberts, Keen, Larson, and Eisenhour (2007) biology textbook explains, there are two ways in which to classify species in order to create this tree. One way is using phenetic classification, which groups together species solely based on their overall similarity. However, using the phenetic classification, evolutionary history of the species is ignored. The second way, and the preferred method for evolutionary biologists, is to classify species using phylogenetic classification. Phylogenetics is the study of evolutionary relatedness and classifies species based on evolutionary principles. Phylogenies, or evolutionary trees, allow for inferences to be made about related species and also are the most common way to organize current knowledge, as the tree is slowly reconstructed. Figure 1 provides an example of a cladogram in the most commonly used type of the evolutionary tree: the tree format.

Phylogenies are read to depict the common ancestry of species, or taxa, from the bottom of the diagram to the top of the diagram. Thus, species depicted in a cladogram are not ancestors of one another, as one might believe by reading the diagram from left to right. Instead, species are descendents of common ancestors located at the branching points. In Figure 1, the ancestor of pig and camel is more recent than the ancestor shared by pig, camel, and human. One concept that is important in understanding how to interpret cladograms, or branching diagrams that depict evolutionary relatedness, is that the arrangement of the taxa on the ends of the branches is unimportant. The branches can rotate, similar to the branches on a child's mobile. Looking at the taxa on the far right of the cladogram in Figure 1, the branches can be rotated in many different ways and the structural integrity of the cladogram will remain constant. For example, the last

three taxa could be camel, pig, human; pig camel human; human, camel, pig; or human, pig, camel. There are different formats of cladograms, but the most commonly used format is the tree form, which is shown in Figure 1.

“Cladograms are a vitally important tool used by evolutionary biology because they represent and organize existing knowledge about species and higher-order taxa...they enable evidence-based inference and prediction and provide a conceptual framework for basic and applied biology” (Novick & Catley, 2010). As the popularity of this diagram in evolutionary biology has risen, it has become apparent that students have not been given adequate education, if given any at all, on these tools used so readily in the real world (Catley, 2006).

Necessity of Teaching Tree Thinking

As explained by Sandvik (2008), tree thinking is an ability that must be taught. To demonstrate this, Sandvik gave a questionnaire of four questions about evolutionary diagrams to university students in Norway who had previously taken courses in biology. Sandvik found that none of the students were able to correctly interpret the cladogram that was used in the questionnaire. These students had taken biology classes before, so this presumably would not have been the first cladogram they had seen, and yet, they were not able to answer the questions on this relatively simple cladogram. The conclusion drawn from this study was that the students had simply not been taught to understand phylogenetics. Sandvik hypothesized either that the students had learned with incorrect diagrams found in the textbooks or that they had learned with correct diagrams but had not gained a full understanding of phylogenetics. Whatever the reason behind it, Sandvik called for textbooks to catch up with current research and for students to be taught specifically how to read and interpret cladograms.

This notion that part of the problem lies in the textbooks students are using was reinforced by a textbook analysis done by Catley and Novick (2008). The researchers examined 31 textbooks, which contained 505 cladograms. Of those, 45% were tree diagrams and 55% were ladder diagrams. Ladder diagrams are another format of cladograms. Figure 2 depicts the ladder version of a cladogram. The textbooks also contained 192 evolutionary diagrams that are not cladograms, most of which, therefore, did not depict evolutionary relationships in a manner consistent with scientific evidence. The high school textbooks, which are used during what has been stressed as a vital early learning period, had the same number of evolutionary diagrams that were not cladograms as they had actual cladograms. In addition, these non-cladogram diagrams actually appeared more frequently than cladograms in the middle school texts. This research showed that textbooks are not teaching specifically with cladograms.

Moreover, research showed that when textbooks were teaching with cladograms they were the ladder format of cladograms that was proven proven to be not as easily understood by students. Previous research had shown that subjects had more difficulty understanding ladder cladograms than tree cladograms (Novick & Catley, 2010), and subjects performed worse on translation problems (translating from one diagram format to another) when a ladder was involved (Novick & Catley, 2007). If we expect students to be able to understand and reason with this important biology diagram, we must make sure the material we use to teach it is based off of findings in previous research.

Throughout past research (Baum, 2005; Novick & Catley, 2007, 2009; O'Hara, 1997), there has been a common call for students to be taught a foundation in phylogenetics prior to taking higher-level biology classes. By gaining a foundation, students will be able to better learn higher level tree-thinking skills. Novick and Catley (2010) suggested that the introduction of this

material earlier in the education process, either in middle or high school, would lay a foundation on which later tree thinking could be built. However, incorrect prior knowledge can also interfere with performance on tree-thinking problems. Thus, it is also important that this introduction, however early or late it may be, is taught correctly.

O'Hara (1997) also supported teaching tree thinking earlier in the education process and said "just as beginning students in geography need to be taught how to read maps, so beginning students in biology should be taught how to read trees and to understand what trees communicate". O'Hara did a wonderful job of illustrating the necessity of understanding cladograms to biologists. Maps are one of the pillars of geography just as cladograms are one of the pillars of biology. Baum (2005) elaborated on this quote and stressed that cladograms are the fundamentals of the evolutionary theory and that students must have a better understanding of them in order to understand evolutionary biology.

Previous Research Towards Effective Teaching of Macroevolution

To effectively teach tree thinking in high school or middle school, it is necessary to understand what concepts have not been learned or are misunderstood at the college level. If we can identify errors made later in evolutionary biology education, we can begin to teach these concepts earlier, or at the very least lay the ground work so acquisition is easier later in the education process.

While putting together a program called EvoBeaker to teach evolutionary biology, Meir, Perry, Herron, and Kingsolver (2007) found that there was no literature on the misconceptions of college students on macroevolution such as there is for microevolution. To compensate for this, they created a list of misconceptions that Herron and Kingsolver believed the students would have and turned the list into multiple-choice problems testing for these misconceptions. All of

the students had taken some college level biology; however, some were in upper-level biology classes and some had only taken lower-level biology classes.

There were four misconceptions identified: time moves left to right; taxa on branches close together are more related than to those on branches farther away; the fewer the number of nodes between two taxa, the more closely related they are; and if there is a straight line between two taxa, it connects the ancestor to its descendent. Three skills were also identified: being able to read traits from the diagram; being able to identify traits shared with common ancestors; and given set of species, being able to construct the diagram. Although they found that students in the upper level courses of biology showed fewer misconceptions than the lower level students (for reading traits from a tree, straight line equals no change, and incorrect mapping of time), there was no difference between the two groups for deducing ancestral traits, reconstructing trees, tip proximity, and node counting. These misconceptions have also been replicated in studies done by Novick and Catley (2007, 2010).

With this collective call for phylogenetics to be taught prior to upper level biology courses, researchers began to look at students in classes prior to college. Passmore and Stewart (2001) created a curriculum, with the aid of high school teachers and college professors, which would be introduced at the high school biology level. Passmore and Stewart looked to find supporting material that would reinforce what was being taught. They used a case approach in which they gave “realistic contexts” which allowed the students to connect the material to a “body of known information” making the learning of such a difficult topic less abstract. Although Passmore and Stewart taught microevolution (natural selection), and not macroevolution as I did, their study gives support to supplementing a curriculum with a realistic

metaphor or context. Within the curriculum I wrote, subjects were presented with the real world context of the material they were learning.

Like Passmore and Stewart, Goldsmith looked to find material that would support what was being taught. Goldsmith though looked at college level students. Goldsmith (2003) created the method called the Great Clade Race in which phylogenetics is taught using a race metaphor. Goldsmith explained that the vocabulary and principles of phylogenetics are not introduced until the students understand the race metaphor. This suggests that once the students understand the underlying concepts, this foundation will enable understanding and reasoning using higher-level concepts. He reported that this way of teaching phylogenetics was successful for both biology majors and non-biology majors, although no detailed presentation of the supporting data was provided. This study suggested that the curriculum for teaching phylogenetics earlier must begin by teaching the underlying concepts, and then later incorporating the vocabulary and principles of phylogenetics.

Although many have called for phylogenetics to be taught in high school and college as a foundation for future higher-level biology classes, no research has been done to look at a prototype of the curriculum that needs to be taught. By creating a prototype for a curriculum to be taught in a college-level biology course, and later to be revised for high school, I hoped that students with both stronger and weaker biology backgrounds would be able to understand and perform better, when compared to baseline performance, on cladogram test problems. Having been given instruction of all underlying principles, subjects' performance should increase on problems that have been proven to be difficult for even the stronger-background subjects (e.g., Novick & Catley, 2010).

Overview of the Curriculum

The prototype curriculum I created was founded on past research on the misconceptions students have about cladistics. The specific concepts or skills taught were those that had been previously found to be difficult for students, even those with a stronger biology background (Meir et al., 2007; Novick & Catley, 2007, 2010). These concepts were evolutionary distance, polytomies, clades, and inferences made using the cladogram. All of these concepts are explained in greater detail later in this paper. Additionally, the curriculum taught subjects about skills that are commonly used in real world applications but may not appear as frequently in textbooks or previous research. These skills looked at manipulating a larger cladogram to create smaller versions that include only specific taxa. Conversely, the skills also looked at the concept of creating a larger cladogram from two or three smaller ones.

The primary goal of the curriculum was to increase students' ability to understand and reason with cladograms. For this reason, all cladograms used in the curriculum were of the tree format, which has been proven to be easier for subjects to understand (Novick & Catley, 2007, 2010). Thus, I decided to present the instruction, and all of the subsequent test problems, in this format. Half of the subjects received the instruction and half did not.

In addition, I examined the effectiveness of my curriculum for those who had taken more courses for biology majors and those who had taken fewer courses. I predicted that with instruction, subjects' performance on previously identified difficult skills would improve. In addition, I wanted to see the curriculum be effective for those with a variety of biology backgrounds.

Method

Subjects

Subjects of this study were 127 undergraduate students of Vanderbilt University. Subjects were voluntarily recruited for pay through the university's psychology research sign-up system. Subjects were asked prior to the study if they were currently enrolled or had taken the second semester of the biology course Introduction to Biological Sciences B (BSCI 110B) or a comparable course at another university, and if they were enrolled in or had taken the biology sciences course Evolution. Subjects were excluded from participating if they were currently enrolled in BSCI 110B in order to enforce a greater difference between the stronger and weaker background subjects. Students were also excluded if they were enrolled or had taken Evolution as they had already been taught the material being tested. The data from one subject who participated in the study were excluded from the analyses because I discovered, after her participation, that she had taken the Evolution class, BSCI 205. The data from another subject were excluded from the analyses because she did not complete the experiment within the allotted 2 hrs. Finally, the data from a third subject were excluded from the analyses due to inaccurate reporting of her academic background (N=124).

Subjects in this study completed a background questionnaire in which they were asked whether they were enrolled in or had taken any of the 10 listed biology classes and 3 listed historical geology classes at the university, or comparable courses at another university, and the grade they had received. The students who had completed BSCI 110A and 110B, or comparable courses at another university, were assigned to the stronger biology background group ($n = 63$). Subjects who had not completed these courses were assigned to the weaker biology background group ($n = 61$). These two subject groups differ greatly in the number of potentially relevant classes taken. The stronger background students on average had taken 2.20 ($sd = 0.43$) such classes. The weaker background students on average had taken .41 classes ($sd = .62$).

Design

A second factor of my study was whether or not the subjects received the instructional condition ($n = 62$) or the no-instruction (control) condition ($n = 62$). There were 31 stronger background subjects and 31 weaker background subjects in the instructional condition. There were 32 stronger background subjects and 30 weaker background subjects in the control condition. Subjects in the instructional condition were given the instructional curriculum, which was created to teach the foundation of tree thinking and corrected errors identified in previous research.

Subjects in the control condition were not given any instruction on tree thinking. Instead, they completed three individual differences tests that took approximately the same amount of time as the instructional booklet in the instructional condition. The three tasks were a subset of the science reasoning section of the ACT and two spatial reasoning tasks. Each of these is part of a separate study.

Procedure

The instructional booklet was self-paced and took about 30 minutes. The measures for the control condition were timed and took approximately the same amount of time as the instructional booklet and practice problems. Then, all subjects were given two test problem booklets and the Conceptual Inventory of Natural Selection (CINS). Finally, they were given the background questionnaire. The second half of the study was also self-paced and took roughly an hour and a half.

The second test booklet is part of another study looking at superficial similarity and how this affects subject's comprehension of relationships among taxon compared to evolutionary relatedness. The Conceptual Inventory of Natural Selection (CINS) also is part of another study.

Instructional Condition

All of the measures were completed using paper test booklets. Subjects were supplied pencils. The instructional condition consisted of two booklets. The first booklet was 16 pages that educated subjects on three levels of cladogram skills and reasoning. The second booklet contained two practice tests for the material taught in the first booklet.

Curriculum. The first sections (cladogram structure and clade; nested hierarchical structure) are considered foundation instruction on tree thinking. The cladogram structure section taught students how to construct and read a cladogram. Subjects were given instruction in the following areas: how time (past and present) is read on a cladogram; most recent common ancestors; evolutionary relatedness; and how a cladogram is based on a 3-taxon statement, the principle that two taxa are more closely related to each other than either is to a third taxon.

The clades and nested hierarchical structure section instructed subjects on what constitutes a clade, a group of taxa that includes the most recent common ancestor and all of its descendants. The section also instructed subjects on what constitutes a sister group, a clade that consists of only two taxa, and how clades can be nested within each other.

The third section, a deeper understanding of evolutionary trees, taught a more global understanding of how the individual cladogram fits into the larger tree of life. While the previous sections taught a more local understanding of cladograms, this section covered reasoning about polytomies, a group of three or more taxa at the same level in a cladograms, and common misconceptions about evolutionary distance/relatedness.

To reason about the tree of life using this more global understanding, it is important for a biologist to consider subsets of the larger tree. One of the ways to do this is to make a larger cladogram smaller by pruning off taxa not being considered with the subset. A second way is to

collapse taxa into a larger group. An example of this, from the instructional condition, would be collapsing bears and felines into their larger group *mammals*. A third way to do this is to merge several smaller cladograms into a single cladogram. The instructional condition ended with a section that instructed the subjects about the real-world importance of proper tree-thinking education.

The prototype of the curriculum was written in order to teach the basics of tree thinking. In order to do, there were certain key features I added to enhance the instructions. The first was the use of figures. Embedded within the text were figures that helped to depict the concept being taught. The first figure was a simple cladogram with only three taxa and characteristics. This is shown in Figure 3a.

Each cladogram that followed also had certain characteristics marked. Subjects were taught that specific shared characteristics, known as synapomorphies, are used to define a group of taxa that includes the most recent common ancestor of the group and all of the descendants, or a clade. All of the following cladograms in the instructional booklet, except that demonstrating polytomies, were built from this original cladogram. This was done so as not to confuse the subjects with too many taxa presented, and to demonstrate to the subject that all of the concepts were built upon each other. The cladograms used throughout the instruction were taken from the Tree of Life website. In order to have an accurate polytomy example, I could not use an extension of the original cladogram as it did not incorporate a polytomy.

A second key feature of the curriculum was the color-code used to help the subject understand each individual concept being taught. Each time the instruction identified a new or different structural aspect of the cladogram, that structure was bolded in a color. Figure 3b followed the introduction of a polytomy. The polytomy in the cladogram is highlighted in purple.

New colors were used for each concept and the colors did not carry over from cladogram to cladogram. I believed that using color indicators would help the subjects more easily identify the part of the cladogram structure being addressed. Because structure is such a vital part of each concept, the color helped draw the eye of the subject to the area on which I wanted them to focus.

The instructional condition was created to not only teach the basics of tree thinking, but also to specifically teach skills that have been problematic for stronger background subjects in previous research (Novick & Catley 2010). These skills include evolutionary relatedness; what constitutes a clade; reasoning with polytomies; and inference questions about character possession. Within each section of instruction, the subjects were told about common misconceptions students have when using these skills (Meir et al. 2007; Novick & Catley 2007, 2010). My instruction thus also is a refutational text. A refutational text, as explained by Mason et al. (2008), is one that “acknowledges students’ alternative conceptions about a topic, directly refutes them, and introduces scientific conceptions as viable alternatives”. Refutational texts have been proven to be superior to traditional ones in science education (Hynd, 1998; Mikkila-Erdmann, 2002; Wang & Andre, 1991).

Practice problem booklet. The practice booklet was included in the study in order to allow subjects to test what they had learned prior to the problems in the test booklets and to re-learn any material that they had not understood. In this way, the instructional condition aimed to mimic actual class instruction. Subjects were given text and example problems, tests throughout the instructional period, and a final exam.

The practice problem booklet of the instructional condition was comprised of two practice quizzes. The first quiz came after the first instructional section on cladogram structure.

Subjects were instructed to stop in the instructional booklet and turn to the practice booklet. The first quiz asked subjects true and false questions about evolutionary relatedness, yes and no questions about 3-taxon statements, and fill-in-the-blank questions about shared characters. After the quiz, subjects were instructed to turn to the next page of the instructional booklet and check their answers. If they had gotten an answer wrong, they were then asked to re-read sections that provided the relevant instruction. After the second instructional section on clades and nested hierarchical structure, subjects were once again told to stop and take a practice quiz. This second quiz asked subjects yes and no questions about what constitutes a clade and what constitutes a sister group. Subjects were also asked fill-in-the-blank questions about which character the most recent common ancestor of a clade possessed. After completion, subjects were once again asked to review their answers and re-read sections if necessary. A practice quiz was not given after the third instructional section, as these skills were not considered foundation skills but rather higher-level skills. This concept of using tests while studying to increase performance was also demonstrated in a study done by Jaffe (Jaffe 2008). This study showed that subjects who were tested while studying performed better and were able to retain information for longer when compared to subjects who simply studied the material.

Tree-Thinking Skills Assessed

The test booklet assessed how well the subjects knew the concepts taught in the instructional condition. The first test booklet was made up of questions that assessed nine different tree-thinking skills. The skills comprised two groups: the first group were skills that have been assessed as problematic in previous research for both stronger and weaker background students. These skills were evolutionary distance incorporating a polytomy, evolutionary distance incorporating levels, recognizing clades, nested clades, and inference. These test

problems were taken from the experimental materials used in previous research in the lab. The second group of problems addressed skills that do not commonly appear in textbooks or cladistic instruction. However, they are tree-thinking skills that are commonly used by biologists in studying the Tree of Life. These questions focused on the understanding that any given cladogram is simply a smaller depiction of the larger Tree of Life and can be manipulated to study different parts of the larger tree. These skills were collapsing branches on the tree into a single group branch, pruning off branches to focus on a specific subset of taxa, and merging together two smaller versions of the tree. A final skill that was tested was the ability to identify errors presented in a cladogram. Previous appropriate test questions for these final types of questions did not exist. For these questions, I wrote new problems that were used in this study for the first time.

Evolutionary distance

Incorporating a polytomy. The booklet contained two evolutionary distance questions that incorporated polytomies. The question asked subjects to reason about relatedness of three taxa of a polytomy. In addition, the subjects were asked to explain their answer. Figure 1 gives an example of this question type. Novick and Catley (2007) found that subjects had difficulty understanding the structure of cladograms when a polytomy was present. Preliminary analyses of research in progress suggest that the evolutionary distance incorporating a polytomy is difficult for students, even after instruction in phylogenetics in an evolution class.

Incorporating levels. The booklet contained 4 evolutionary distance questions that incorporated levels of most recent common ancestry. These questions were of two types. The first type asked subjects to reason which of two taxa had the closest evolutionary relationship to a specified taxon. All of the taxa were located at different levels of the cladogram. Previous

research showed that subjects from both backgrounds performed poorly on these question types. Stronger background subjects had a mean of 0.39 and weaker background subjects had a mean of 0.29 (Novick and Catley 2010). The second question type, a follow-up to the first in a subset of cases, asked subjects to reason if their answer would change if 5 additional characters were added at a specific point in the cladogram. Figure 4 gives examples of both question types.

Clades. As “clade” was a term learned in the instruction condition but also used in the test booklets, subjects in both conditions were given a definition sheet to use throughout the test which defined the word “clade”. This allowed for the subjects in the control condition to also be able to understand and use the term. There were two types of clade questions: recognizing clades and nested clades.

Recognizing clades. The booklet contained 4 recognizing clades questions. In one problem of this type (from Novick, Catley, & Funk, 2010), subjects were shown a cladogram and asked which subset of taxa is a valid biological group. Subjects were then given three answers from which to choose. Figure 5 gives an example of this type of recognizing clades question. In a second problem type (from Novick & Catley, 2010), subjects were asked whether the bracketed taxa comprised a clade. If the subject said “no”, they were asked what taxon/ taxa would need to be added or removed in order for it to be a clade. With both types of problems, subjects were asked to explain their answer. Previous research using the bracketed taxa question (Novick & Catley, 2010) showed that stronger background subjects were able to recognize and justify what comprised a clade with a mean of 0.64. In contrast, weaker background subjects had a mean of 0.49. I hoped with specific instruction detailing what a clade is, students would be better able to identify and explain a clade.

Nested clades. The booklet contained two nested clades questions. Subjects were shown a cladogram with a clade already marked. The subjects were then asked to mark the remaining clades on the diagram and write the number of clades. Figure 1 gives an example of a nested clades question. These questions tested subjects' understanding of what comprises a clade and how clades can be nested within each other. Preliminary analyses of research in progress suggest that the nesting of clades is difficult for students, even after instruction in phylogenetics in an evolution class.

Inference. The booklet contained two inference questions. These questions asked subjects to make judgments about what taxa are most likely to share a characteristic with a specified taxon and explain their choice. Figure 5 gives an example of an inference question. These questions tested subjects' understanding that inferences made about shared characters should be based on a shared most recent common ancestor. Previous research found that without instruction subjects had a mean of 0.51 on these types of questions (Novick & Catley, 2010). I hoped that with instruction, students' performance on these questions would improve.

Subsets of the tree of life.

Pruning. The booklet contained three pruning questions. Two questions presented an original cladogram and asked the subject to redraw that cladogram by pruning off certain taxa. Figure 6 gives an example of this type of pruning question. A third question was a multiple choice question that asked what cladograms could be created by pruning an original cladogram. This question came from Baum et al. (2005). However, I translated the original question, which contained cladograms in the ladder format, to tree format. These questions tested the student's understanding that cladograms can be manipulated, such as pruning off taxa, and the relationships among the remaining taxa can be preserved.

Collapsing. The booklet contained two collapsing questions. Each question presented an original cladogram and asked the subject to collapse multiple taxa into a single biological group. The same cladogram was used to ask both a pruning and collapsing question on separate pages. Figure 6 gives an example of a collapsing question. These questions also tested the understanding that you can manipulate a cladogram and preserve the relationships among the taxa.

Merging. The booklet contained four merging questions. Three questions presented two or three cladograms and asked the subjects to merge them into a single cladogram. Figure 7 gives an example of this type of merging question. The subjects were told that two (or three) researchers studied different taxa, some of which overlapped. They were then told that a third (fourth) researcher was interested in studying all of the presented taxa. The subject was then asked to draw a single cladogram that depicted the evolutionary relationships among all of the given taxa. A fourth question was a multiple-choice question that asked what new cladogram could be made by merging the three existing cladograms. This question came from Baum et al. (2005). I used it in its original format. These questions tested the skill of merging smaller cladograms together to get a larger cladogram. The questions also tested the understanding that a cladogram is simply a smaller subset of the greater tree of life and that you can merge two cladograms together to see a bigger part of the picture.

Identifying errors. The identification of errors question was written to test whether a subject would be able to identify errors when presented with an incorrect cladogram. Previous research had found that asking students to identify errors is a good way to measure their understanding of a concept (e.g., Whitley, Novick, & Fisher, 2006). This question is presented in Figure 8. The subjects were told that Pat was given a cladogram (shown) and then was asked to

prune and collapse taxa. Subjects were then shown what Pat drew and were told to identify the several errors that Pat made. The three mistakes that Pat made were mistakes I believed subjects might make themselves when performing these manipulations to the cladogram. The first error was an incorrect collapsing error. Pat should have collapsed the taxa into *Anglerfish* but instead he/ she put a bracket over the cryptic anglerfish and blackspot anglerfish and labeled that *Anglerfish*. The second error is a pruning error. When pruning spiny dreamer, Pat created a polytomy involving whalehead dreamer, smooth dreamer, and rounded batfish/ slantbrow batfish. There should be a more recent common ancestor of whalehead dreamer and smooth dreamer that is not shared by the other two taxa. The third error is a second type of pruning error. When pruning monkfish, Pat leaves a bend in the branch where the taxon was removed.

Results

Overview

Twenty-nine questions were given scores of 0 or 1 for accuracy. Fourteen questions had written explanations and were coded with 0, 0.5, or 1 for evidence quality. A subject's explanation was given a score of 1 if he or she talked about most recent common ancestry in the way appropriate to the question. A subject's explanation was given a score of 0.5 if they talked about evolutionary relatedness, recent common ancestry, or the presence of some of the descendants. A subject's explanation was given a score of 0 if they talked about any other concept to explain their answer. Examples of answers for each composite score are discussed in each score's sub-section. The explanation questions were coded by two people independently, who agreed on 90% of the scores. Discrepancies were resolved by discussion. A composite score was created for each of nine tree-thinking skills. Each of these scores was analyzed by a 2x2 between-subjects ANOVA. The results of all of the ANOVAs are given in Table 1.

Skills Assessed**Evolutionary distance.**

Incorporating a polytomy. The polytomy evolutionary distance score was averaged across four scores. The questions were scored two ways: accuracy and explanation of answer. Subjects were given full credit for explanations that were more sophisticated. To get a score of 1, subjects had to say the taxa shared a most/more recent common ancestor, closest common ancestor, more common ancestor, most recent common branch, taxa are in the same clade, they are a sister group or the pattern of most recent common ancestry doesn't change. One example of this type of explanation is "all share same most-recent ancestor". Subjects were given partial credit if their explanation was less sophisticated. To get a score of 0.5, subjects had to say they all have the same recent ancestor, they share a recent common ancestor, they have a close(r) common ancestor, share more ancestors, share direct ancestor, the taxa are most/more closely related evolutionarily, X & Y are equally distantly related to Z, this group has the same evolutionary history, or these taxa are part of a polytomy. One example of this type of explanation is "Walking sticks & cockroaches directly share a common ancestor whereas crickets have a closer ancestral relationship w/ cicadas". Subjects were given no credit if their explanation was anything else. An example of this is, "Walking sticks and cockroaches are only 1 more characteristic removed, crickets are 3".

The ANOVA yielded only a significant effect of condition. As predicted, students in the instructional condition did much better than those in the control condition. The mean for the control condition was 0.20 ($sd = 0.28$). The mean for the instructional condition was 0.63 ($sd = 0.33$).

Incorporating levels. The evolutionary distance questions incorporating levels score was averaged across 8 scores. The questions were scored two ways: accuracy and explanation of answer. Subjects' explanations for these questions were scored the same way as the evolutionary distance questions involving a polytomy. One example of a full credit explanation is "The sea urchin and camel share a more recent common ancestor.". One example of a partial-credit explanation is "It has the most recent evolutionary ties". An example of a zero credit explanation is "lobsters & urchins are only two branches away from each other while camels and urchins are a lot further".

The ANOVA yielded only a significant effect of condition. As I predicted, students in the instructional condition did much better than those in the control condition. The mean for the control condition was 0.22 ($sd = 0.23$). The mean for the instructional condition was 0.68 ($sd = 0.33$).

Clades.

Recognizing clades. The recognizing clades score was averaged across 9 scores. The questions were scored two ways: accuracy and explanation of answer. Subjects were given full credit for explanations that were more sophisticated. To get a score of 1, subjects had say that the taxa comprise a clade (monophyletic group), the taxa share a most recent common ancestor and all the descendants of that ancestor are include, contains all the taxa/organisms/etc. from this ancestor, or does not include all descendants of the most recent common ancestor. One example of this type of explanation is "his definition accounts for all the given taxa of a MRCA". Subjects were given partial credit if their explanation was less sophisticated. To get a score of 0.5, subjects had to say that only some (not all) descendants are included, taxon X is (other taxa are) also a descendant, other taxa also have character C, only part of the clade is included, all of the

descendants are included, they all have the same recent ancestor, they share a recent common ancestor, they have a close(r) common ancestor, share more ancestors, share direct ancestor, the taxa are most/more closely related evolutionarily, X & Y are equally distantly related to Z, this group has the same evolutionary history, or these taxa are part of a polytomy. One example of this type of explanation is “share close common ancestor”. Subjects were given no credit if their explanation was anything else. An example of this is, “these 3 are closest on the diagram, and share the same reptilian characteristics”.

The ANOVA yielded a significant effect of condition. The mean of the control condition was 0.55 ($sd = 0.24$). The mean of the instructional condition was 0.76 ($sd = 0.23$). As I predicted, students in the instructional condition did better than those in the control condition. The ANOVA also yielded a significant effect of biology background. The mean of the weaker background subjects was 0.57 ($sd = 0.28$). The mean of the stronger background subjects was 0.73 ($sd = 0.21$). Stronger background subjects did better than those with a weaker biology background. There was an interaction between condition and biology background. This is depicted in Figure 9. Weaker background subjects in the control condition had an average of 0.41 ($sd = 0.19$) compared to those in the instructional condition who had an average of 0.73 ($sd = 0.26$). Stronger background subjects in the control condition had an average of 0.68 ($sd = 0.21$) compared to those in the instructional condition who had an average of 0.79 ($sd = 0.20$). The interaction shows that the performance gap between the weaker and stronger background subjects was smaller with instruction. Put another way, the instruction was more effective for the weaker background subjects whose scores in the control condition were much lower than those of the stronger background subjects.

Nested Clades. The nested clades score was averaged across four scores. The questions were of two types: number of clades the subject marked on the clade and the number of clades the subject reported. As explained before, subjects were asked to mark clades and then record the number they had marked. Each question was given an accuracy score of 0 or 1. Discrepancies occurred between these two numbers. To account for this, I took an average of the four numbers reported.

The ANOVA yielded a significant effect of condition. The mean of the control condition was 0.44 ($sd = 0.45$). The mean of the instructional condition was 0.73 ($sd = 0.36$). As I predicted, students in the instructional condition did much better than those in the control condition. The ANOVA also yielded a significant effect of biology background. The mean for the weaker background subjects was 0.47 ($sd = 0.44$). The mean for the stronger background subjects was 0.71 ($sd = 0.39$). Stronger background subjects did much better than those with a weaker biology background. There was no interaction.

Inference. The inference score was averaged across 4 scores. These questions were scored two ways: accuracy and explanation of answer. Subjects were given full credit for explanations that were more sophisticated. To get a score of 1, subjects had to say the taxa share a most/more recent common ancestor, closest common ancestor, more common ancestor, most recent common branch, taxa are in the same clade, they are a sister group or the pattern of most recent common ancestry doesn't change. One example of this type of explanation is "Birds and lizards have an MRCA than birds and porcupines". Subjects were given partial credit if their explanation was less sophisticated. To get a score of 0.5, subjects had to say they all have the same recent ancestor, they share a recent common ancestor, they have a close(r) common ancestor, share more ancestors, share direct ancestor, the taxa are most/more closely related

evolutionarily, X & Y are equally distantly related to Z, this group has the same evolutionary history, or these taxa are part of a polytomy. One example of this type of explanation is “closer / shorter evolutionary distance”. Subjects were given no credit if their explanation was anything else. An example of this is, “b/c lizards are a nearer branch to birds than are porcupines”.

The ANOVA yielded a significant effect of condition. The mean of the control condition was 0.61 ($sd = 0.25$). The mean of the instructional condition was 0.75 ($sd = 0.26$). As predicted, students in the instructional condition did better than those in the control condition. The ANOVA also yielded a significant effect of biology background. The mean of the weaker background subjects was 0.63 ($sd = 0.29$). The mean of the stronger background subjects was 0.74 ($sd = 0.23$). Stronger background subjects did better than those with a weaker biology background. There was no interaction.

Subsets of the tree of life.

Pruning. The pruning score was averaged across 3 scores. The subjects were given accuracy scores for following the directions and redrawing the cladogram correctly. The grand mean was 0.86 ($sd = 0.23$).

Collapsing. The collapsing score was averaged across 3 scores. The subjects were given accuracy scores for following the directions and redrawing the cladogram correctly.

The ANOVA yielded only a marginally significant effect of biology background. The mean for the weaker background subjects was 0.82 ($sd = 0.32$). The mean for the stronger background subjects was 0.91 ($sd = 0.23$). Stronger background subjects did marginally better than those with a weaker biology background.

Merging. The merging score was averaged across 4 scores. 3 of these questions were given scores for accuracy of following the directions and redrawing the cladogram correctly. 1 question was scored on accuracy of the multiple-choice question.

The ANOVA yielded only a significant effect of biology background. The mean for the weaker background subjects was 0.71 ($sd = 0.26$). The mean for the stronger background subjects was 0.83 ($sd = 0.19$). Stronger background subjects did much better than those with a weaker biology background.

Identifying errors. The identifying errors score was averaged across 6 scores. These questions were scored two ways: accuracy of identification of the error and quality of explanation of how to correct the error. The correct explanation to fix the first error was to collapse the taxa into *Anglerfish*. Subjects were given full credit for responses that explain how to fix this, “instead of the [sister bracket drawn] for anglerfish it should just a be a straight line”. Subjects were given partial credit if their explanation was only partially correct or explained what the error was rather than how to fix it. An example of a partial credit response for the first error was “No need for brackets and double bar”. Subjects were given no credit if they incorrectly explained how to fix the first error or didn’t give an explanation, “the names of the two taxa should be written”.

The correct explanation to fix the second error was that there should be a more recent common ancestor of whalehead dreamer and smooth dreamer that is not shared by the other two taxa. Subjects were given full credit for responses that explain how to fix this, “Whalehead dreamer and smooth dreamer share a common ancestor. They compose a sister clade and should be branched separately.” Subjects were given partial credit if their explanation was only partially correct or explained what the error was rather than how to fix it. An example of a partial credit

response for the second error was “dreamers should be grouped”. Subjects were given no credit if they incorrectly explained how to fix the second error or didn’t give an explanation, “smooth dreamer is not at the same level as whalehead dreamer”.

The correct explanation to fix the third error was that there should not be a bend in the branch where the taxon was pruned off. Subjects were given full credit for responses that explain how to fix this, “There should be no “step” in the cladogram. There should be one branch (stem) from the MRCA w/ the triggerfish.” Subjects were given partial credit if their explanation was only partially correct or explained what the error was rather than how to fix it. An example of a partial credit response for the third error was “he should not have left a kink in the diagram”. Subjects were given no credit if they incorrectly explained how to fix the third error or didn’t give an explanation, “she should add the monkfish, spriny dreamer, and longnose seabat taxa.”

The ANOVA yielded only a significant effect of condition. The mean for the control condition was 0.67 ($sd = 0.24$). The mean for the instructional condition was 0.77 ($sd = 0.24$). As I predicted, students in the instructional condition did better than those in the control condition.

Discussion

Past research has suggested that students must be taught a foundation in phylogenetics prior to learning higher level tree-thinking skills in order for them to have an ability to truly understand and interpret cladograms, (Baum, 2005; Novick & Catley, 2007, 2009; O’Hara, 1997). This study supports this call for prior-foundation learning and demonstrated the effectiveness of teaching the foundation of tree thinking not only for those with a weaker background but also for those with a stronger biology background.

Effectiveness of Instruction

When creating this prototype I had two main goals: the first was to increase performance on questions testing specific tree-thinking skills that had been seen as problematic in the past even for students with a stronger biology background. Those in the instructional condition performed significantly better on all of these skills: evolutionary distance incorporating levels, evolutionary distance incorporating a polytomy, recognizing clades, nested clades, and inference. Across these five skills, the control condition had an average score of 40% compared to the instructional condition that had an average score of 71%. Across all test questions, those in the control condition had an average score of 57% compared to those in the instructional condition with an average score of 76%. This showed that the curriculum was effective in increasing performance on the previously identified difficult skills. The curriculum was also effective in teaching one of the skills that had not been previously tested, identifying errors. However, the overall averages show that the curriculum may not have been effective in teaching each individual skill.

The second goal was for the curriculum to increase performance for those with a weaker prior biology background as well as for those with a stronger biology background. There was a main effect of condition for six dependent variables and in all but one case there was no interaction with biology background. The one interaction showed that the instruction was better for the weaker background subjects, whose scores started lower. Our instruction thus accomplished this second goal.

Analysis of Instruction and Test Problems

Although the instruction did prove effective in overall performance, an overall average score of 76% indicates that there are areas that can still be improved. Three areas of lower performance for subjects in the instructional condition were evolutionary distance incorporating

a polytomy, evolutionary distance incorporating levels, and nested clades. Although performance was higher on these three tree-thinking skills than in previous research, the instruction on these concepts will need to be strengthened in future versions of my curriculum.

Another area for future improvement is the subset of the tree of life section of the test problems. For these skills, there was no main effect of condition. The lack of condition effect could reflect problems with the instruction, the test questions, or both. I believe the problem lies in the test questions. The problems had to be created from scratch, other than the two multiple-choice questions (Baum et al. 2005), and had not been tested on previous subjects as the other questions had. If my first draft of problems could be answered by hierarchical knowledge, instead of knowledge of this subset of the Tree of Life skills, then this would account for the lack of significant performance difference between conditions.

I believe that the high level of performance in the control condition can also be attributed to a college students' general understanding of hierarchical relationships. Previous research done by Novick (2006) showed that college students have a clear understanding of hierarchical relationships. This would explain why those in the control condition were able to perform so highly on these problems without instruction. The average score for the weaker background subjects in the control condition was 0.78. The stronger background subjects in the control condition had an average score for these problems of 0.86.

Research done by Novick (2006) also found that those in upper level computer science and engineering classes performed better on certain hierarchical properties. This might help explain the biology background difference found with the merging problems and marginally with the collapsing problems. Although all of the subjects may have had a strong understanding of hierarchical relationships, which led to high performance on the test problems, perhaps those in

the stronger background also have had either more experience or training on hierarchical levels. Because I have identified that my first draft of problems may have tested hierarchical knowledge more than these implicit concepts, I have worked on new test problems that I hope will better assess the pruning, collapsing, and merging tree-thinking skills.

A final area of improvement is the material provided for those with a weaker biology background. The apparent effect of biology background, even after instruction, showed that there is still information that needs to be included in the instruction to help those with a weaker biology background. This knowledge will help as the curriculum continues to be strengthened.

Future Implications

I hope this prototype curriculum will be able to serve as a stepping-stone on the path of early instruction of tree thinking. This curriculum is currently being implemented in a college-level introductory biology course for biology majors at another university. In the future, I want to teach the curriculum at the high school level because this is where the early education must begin. As I apply this curriculum in different settings and for different groups of students, I will continue to change and advance the curriculum.

The increased performance in the instructional condition showed that with instruction, students can perform better on these tree-thinking tasks and avoid the common misconceptions previous students have had. This instructional booklet provides a complement to existing instruction that focuses on microevolution. By using the booklet as a supplement to a current textbook, students will be taught both sides of evolution education, micro and macro, that are called for by Catley (2006).

References

- Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the Conceptual Inventory of Natural Selection. *Journal of Research in Science Teaching*, 39, 952-978.
- Baum, D.A., Smith, S.D., & Donovan, S.S. (2005). The tree thinking challenge. *Science*, 310, 979-980.
- Catley, K. M. (2006). Darwin's Missing Link: A novel paradigm for evolution education. *Science Education*. 90, 767-783.
- Catley, K. M., & Novick, L. R. (2008). Seeing the wood for the trees: an analysis of evolutionary diagrams in biology textbooks. *BioScience*, 58, 976-987.
- Catley, K. M., Novick, L. R., & Funk, D. J. (2010). The Promise and Challenges of Introducing Tree Thinking into Evolution Education. In K. Rosengren, E. M. Evans, S. Brem, & G. Sinatra (EDS.), *Evolution Challenges: Integrating Research and Practice in Teaching and Learning about Evolution*. accepted pending revision.
- Goldsmith, D.W. (2003) The great clade race: presenting cladistic thinking to biology majors & general science students. *The American Biology Teacher*, 65, 679-682.
- Hickman, C.P., Roberts, L.S., Keen, S. L., Larson, A., & Eisenhour, D. J. (2007). *Animal Diversity*. New York : McGraw Hill.
- Hickman, C.P., Roberts, L.S., Keen, S. L., Larson, A., & I'Anson, H. (2004). *Zoology*. New York : McGraw Hill.
- Jaffe, E. (2008). Will That Be on the Test? *Observer*, 21.
- Meir, E., Perry, J., Herron, J. C., & Kingsolver, J. (2007). College students' misconceptions about evolutionary trees. *The American Biology Teacher*, 69, 71-76.

- Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: a comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching*, *45*, 1131-1160.
- Novick, L.R., & Catley, K. M. (2007). Understanding phylogenies in biology: the influence of a Gestalt perceptual principle. *Journal of Experimental Psychology: Applied*, *13*, 197-223.
- Novick, L. R., & Catley, K. M. (2010). Understanding the Tree of Life: Exploring Cladogram-Based Tree-Thinking Skills in College Students. Invited resubmission under review.
- Novick, L. R., Catley, K. M., & Funk, D. J. (2010). *Effects of Taxon Familiarity on Tree Thinking*. Manuscript in preparation.
- O'Hara, R. J. (1997). Population Thinking and Tree Thinking in Systematics. *Zoologica Scripta*, *26*, 323-329.
- Passmore, C., & Stewart, J. (2002). A modeling approach to teaching evolutionary biology in high schools. *Journal of Research in Science Teaching*, *39*, 185-204.
- Perry, J., Meir, E., Herron, J. C., Maruca, S., & Derek S. (2008). Evaluating two approaches to helping college students understand evolutionary trees through diagramming tasks. *Life Science Education*, *7*, 193-201.
- Sandvik, H. (2008). Tree thinking cannot taken for granted: challenges for teaching phylogenetics. *Theory in Biosciences*, *127*, 45-51.
- Sandoval, W. A., & Reiser, B. J. (2003). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*, *88*, 345-372.

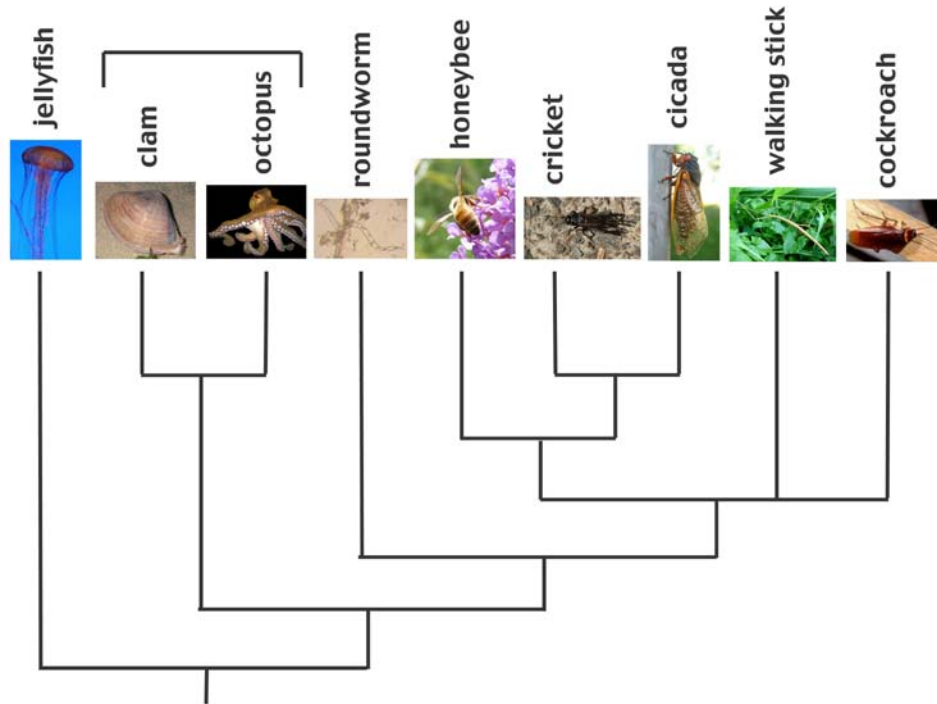
Table 1

The results of all of the 2x2 between subject ANOVAs.

Measure	Condition	Biology Background	Interaction
Evolut. Dist.Polytomy	$F=61.46, p < .001,$ partial $\eta^2 = .34$	$F=2.04, p > .15,$ partial $\eta^2 = .02$	$F=1.28, p > .25,$ partial $\eta^2 = .01$
Evolut. Dist. Levels	$F=80.81, p < .001,$ partial $\eta^2 = .40$	$F=1.28, p > .25,$ partial $\eta^2 = .01$	$F=0.03, p > .85,$ partial $\eta^2 = .00$
Recognizing Clades	$F=29.71, p < .001,$ partial $\eta^2 = .20$	$F=18.02, p < .001,$ partial $\eta^2 = .13$	$F=7.98, p < .01,$ partial $\eta^2 = .06$
Nested Clades	$F=17.96, p < .001,$ partial $\eta^2 = .13$	$F=12.25, p < .01,$ partial $\eta^2 = .09$	$F=0.91, p > .30,$ partial $\eta^2 = .01$
Inference	$F=10.87, p < .01,$ partial $\eta^2 = .08$	$F=6.46, p < .05,$ partial $\eta^2 = .05$	$F=0.14, p > .70,$ partial $\eta^2 = .00$
Pruning	$F=0.62, p > .40,$ partial $\eta^2 = .01$	$F=2.07, p > .15,$ partial $\eta^2 = .02$	$F=2.12, p > .10,$ partial $\eta^2 = .02$
Collapsing	$F=1.45, p > .20,$ partial $\eta^2 = .01$	$F=3.65, p < .10,$ partial $\eta^2 = .03$	$F=2.51, p > .10,$ partial $\eta^2 = .02$
Merging	$F=0.72, p > .35,$ partial $\eta^2 = .01$	$F=7.93, p < .01,$ partial $\eta^2 = .06$	$F=0.51, p > .45,$ partial $\eta^2 = .00$
Identifying Errors	$F=5.62, p < .05,$ partial $\eta^2 = .05$	$F=0.49, p > .45,$ partial $\eta^2 = .00$	$F=0.71, p > .40,$ partial $\eta^2 = .01$

Note: all $F(1,120)$.

The students in a basic biology class are learning about evolutionary relationships among taxa. According to biologists, the following cladogram provides this information about the indicated taxa, which are various invertebrates. Use this cladogram to answer the questions on this page.



Example of an evolutionary distance question that incorporates a polytomy:

1. Which of the following three statements (A, B, or C) is best supported by the scientific evidence: (Answer: C)
 - A. Walking sticks are more closely related to cockroaches than to crickets.
 - B. Walking sticks are more closely related to crickets than to cockroaches.
 - C. Cockroaches, walking sticks, and crickets are all equally closely related to each other.

Explain your answer:

Example of a nested clades question:

2. Two taxa that constitute a clade are enclosed in brackets in the diagram. How many clades are there in this diagram (including the one already marked)? _____ (Answer:7)
 Mark each additional clade with a bracket as shown in the example.

Figure 1. Example of an evolutionary distance question that incorporates a polytomy and a nested clades question.

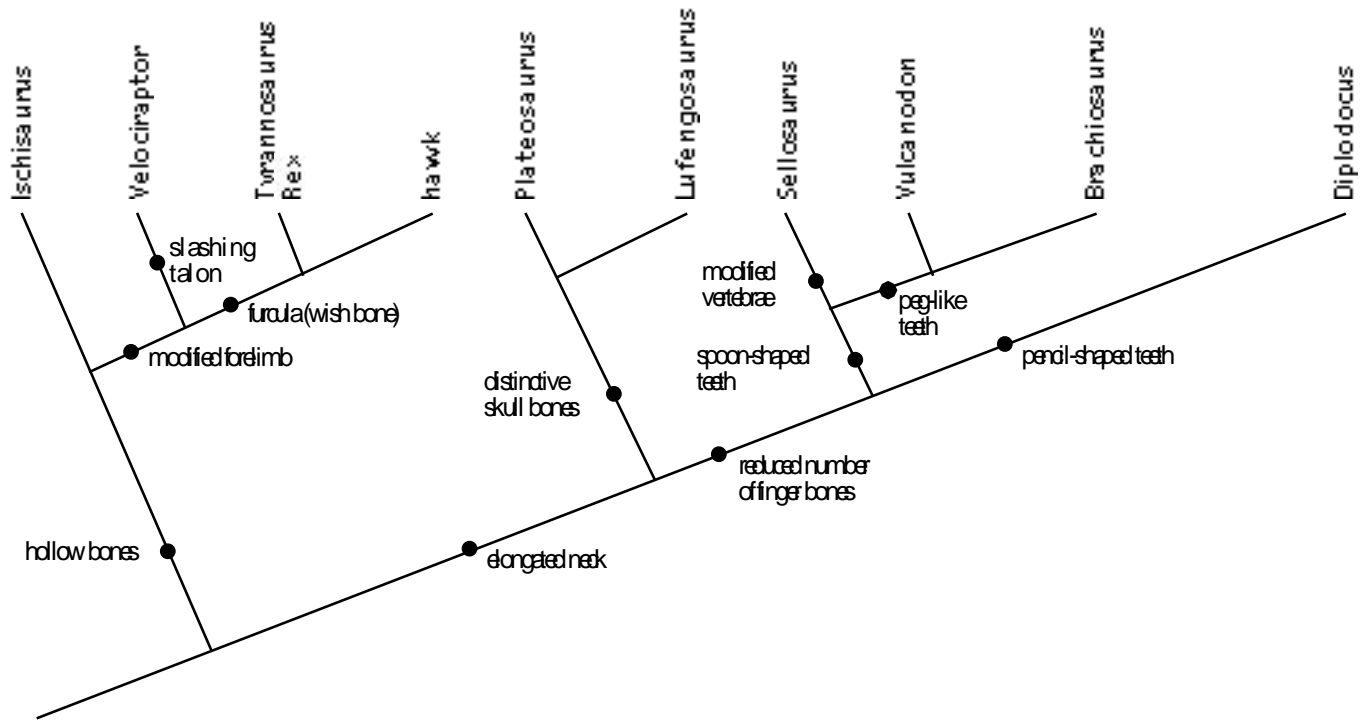
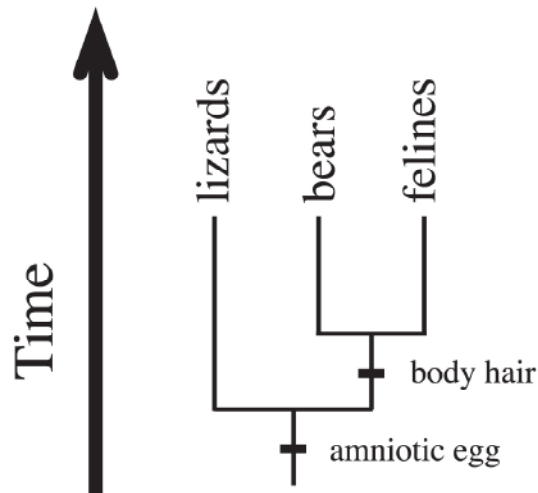
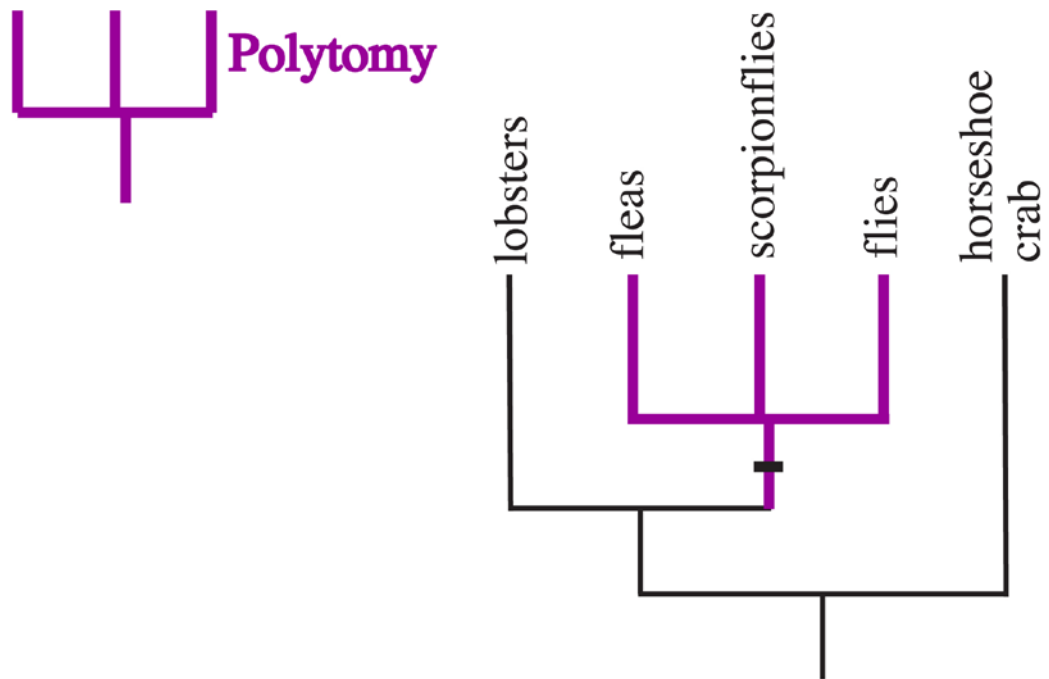


Figure 2. A cladogram in ladder format.



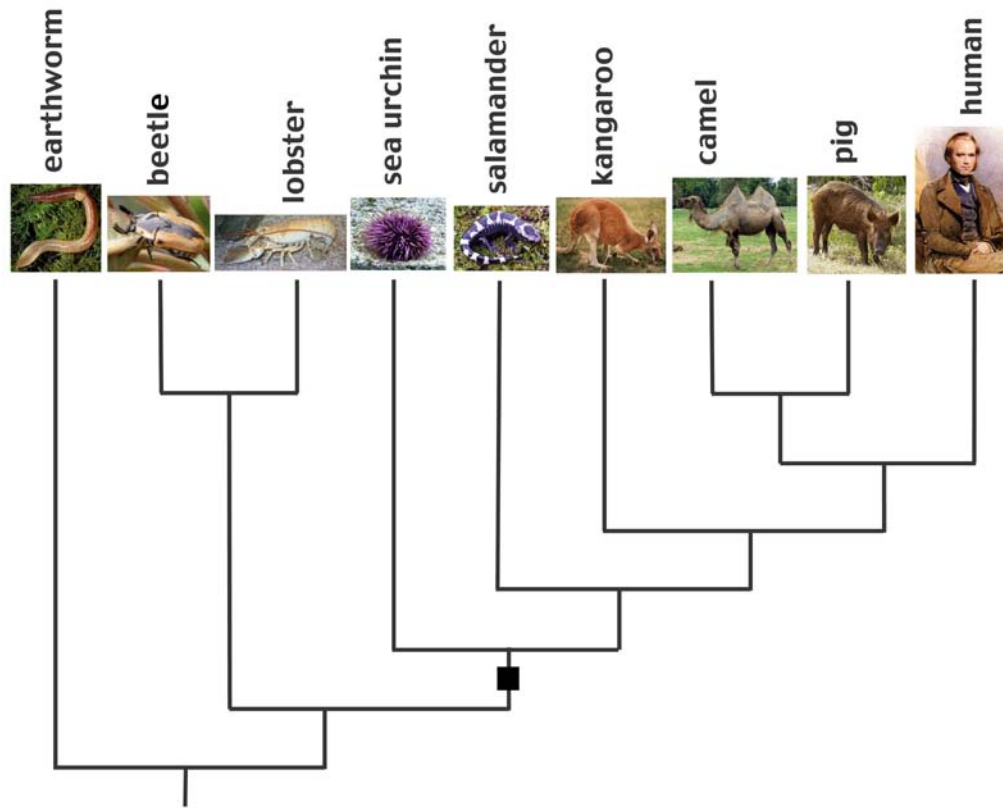
(a) The first figure used in the instructional condition. The figure demonstrates a simply three taxa cladogram and shared characters.



(b) An example of one of the figures in the instructional condition that used color to emphasise instruction.

Figure 3. Examples of the cladograms used in the instructional condition.

The students in a basic biology class are learning about evolutionary relationships among taxa. According to biologists, the following cladogram provides this information about the indicated taxa, which are various animals. Use this cladogram to answer the questions on this page.



1. Which taxon—lobster or camel—is the closest evolutionary relation to the sea urchin?
 _____ (Answer: camel)

Explain your answer:

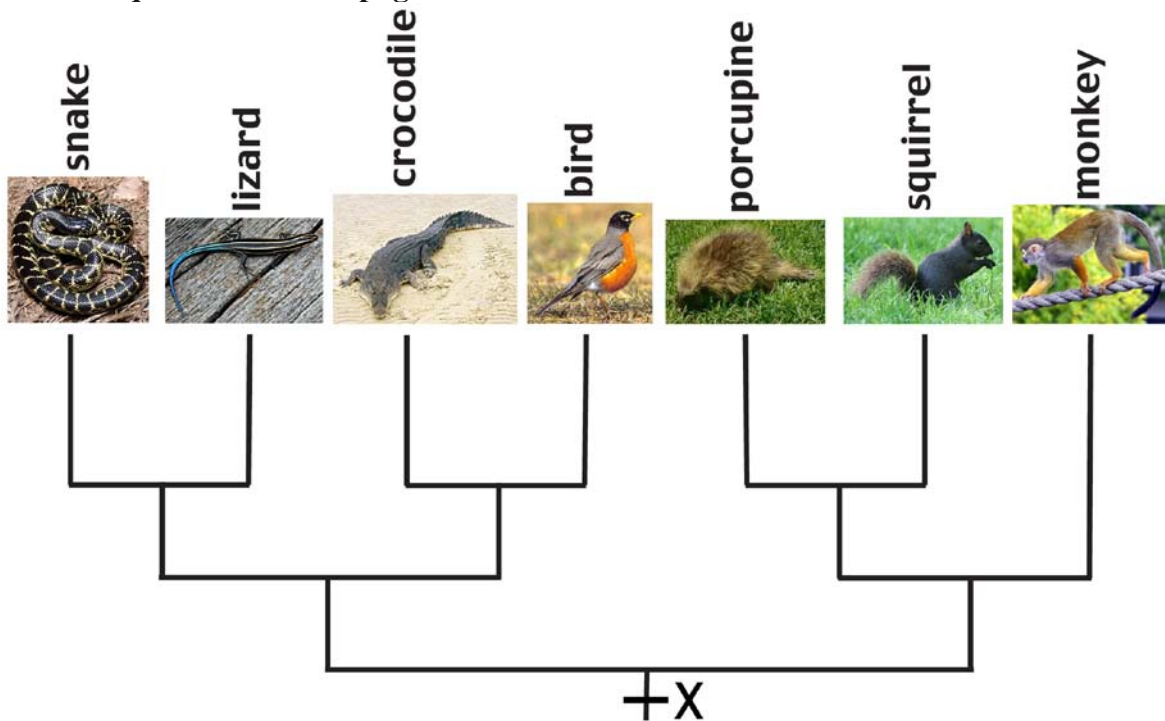
2. Suppose that, given new discoveries, biologists now realize that the black square on the diagram should be interpreted as representing 5 novel characters rather than just 1 as was thought when the diagram on this page was constructed. Given this new information, would your response to part (a) of this question change:

yes **no** (circle one) (Answer: no)

Explain your answer:

Figure 4. Example of evolutionary distance questions that incorporate levels.

The students in a basic biology class are learning about evolutionary relationships among taxa. According to biologists, the following cladogram provides this information about the indicated taxa, which are various animals. The students understand that all of the taxa shown in this diagram share a common ancestor marked by the X. Use this cladogram to answer the questions on this page.



Example of a recognizing clades question:

1. The following students disagree about which taxa should be considered reptiles. Which student's definition of reptiles is best supported by evolutionary evidence?

_____ (Answer: Taylor or Robin)

Taylor: *reptiles = snakes + lizards*

Robin: *reptiles = snakes + lizards + crocodiles*

Jordan: *reptiles = snakes + lizards + crocodiles + birds*

Explain your answer:

Example of an inference question:

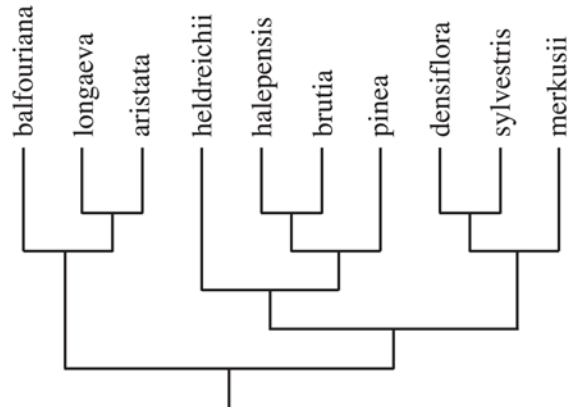
2. Given that birds have a large hole in their palate, are lizards or porcupines more likely to share this character?

(Answer: lizards)

Explain your answer:

Figure 5. Example of a recognizing clades question and an inference question.

Consider the following cladogram, which shows evolutionary relationships among pine trees.



1. Redraw the cladogram shown above without the following two taxa:

(a) *pinea*

(b) *sylvestris*

Make sure the cladogram you draw preserves all of the relationships among the remaining taxa.



Answer:

1. Redraw the cladogram shown above incorporating the following two revisions:

(a) put together the three taxa *balfouriana*, *longaeava*, and *aristata* into a single group labeled *Balfourianae*.

(b) put together the four taxa *heldreichii*, *halepensis*, *brutia*, *pinea* into a single group labeled *Pinaster*.

Make sure the cladogram you draw preserves all of the relationships among the remaining taxa.

Answer:

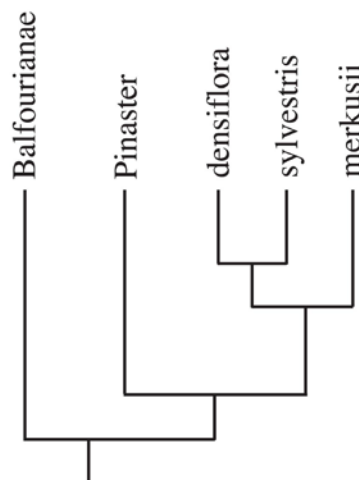
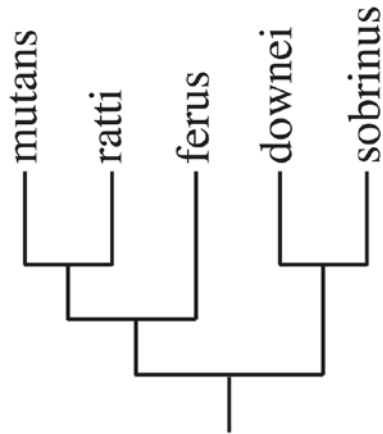


Figure 6. Example of a pruning question and a collapsing question.

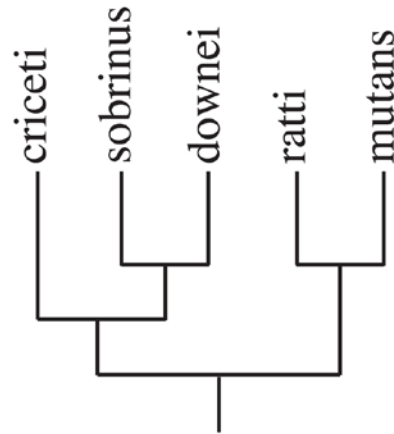
Researcher 3 is interested in the evolutionary relationships among the *Streptococcus* bacteria taxa that were studied by Researchers 1 and 2. Those two researchers studied some of the same taxa but some different taxa as well, as described below.

Researcher 1 studied the evolutionary relationships among 5 *Streptococcus* bacteria taxa. On the basis of this research, Researcher 1 drew the following cladogram:



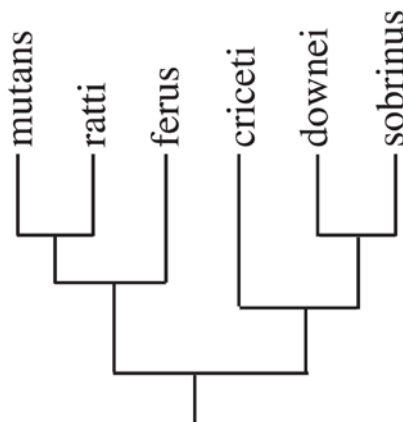
Researcher 1

Researcher 2 also studied the evolutionary relationships among 5 *Streptococcus* bacteria taxa. On the basis of this research, Researcher 2 drew the following cladogram:



Researcher 2

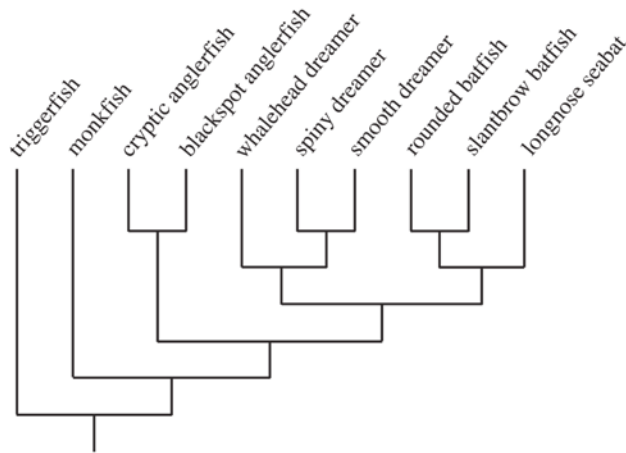
1. Draw a single cladogram for Researcher 3 that depicts the evolutionary relationships among all 6 *Streptococcus* bacteria taxa studied by Researchers 1 and 2. Make sure the cladogram you draw preserves all of the relationships depicted in the two smaller cladograms.



Answer:

Figure 7. Example of a merging question.

Consider the following cladogram, which shows evolutionary relationships among fish.

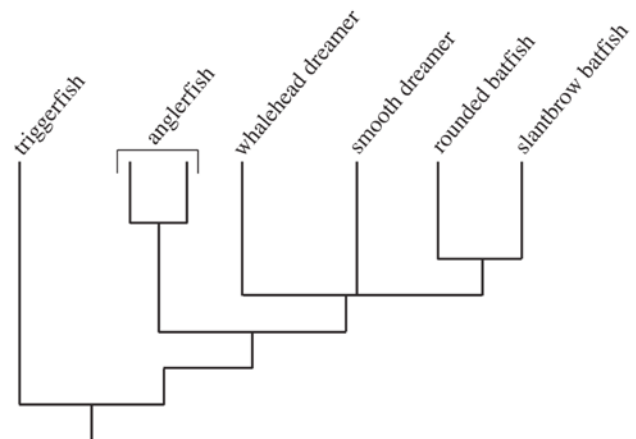


1. Pat was asked to redraw this cladogram with the revisions indicated below on the left. What Pat drew in response to these instructions is shown below on the right.

(a) without the following three taxa:

- i. monkfish
- ii. spiny dreamer
- iii. longnose seabat

(b) put together the two taxa cryptic anglerfish and blackspot anglerfish into a single group labeled anglerfish



There are several errors in this drawing. Circle each error on Pat's drawing. Then in the space below explain how each error should be corrected.

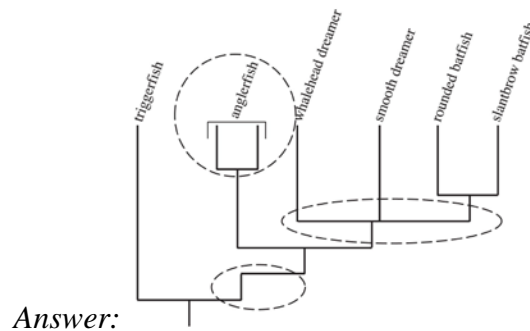


Figure 8. Example of the identification error question.

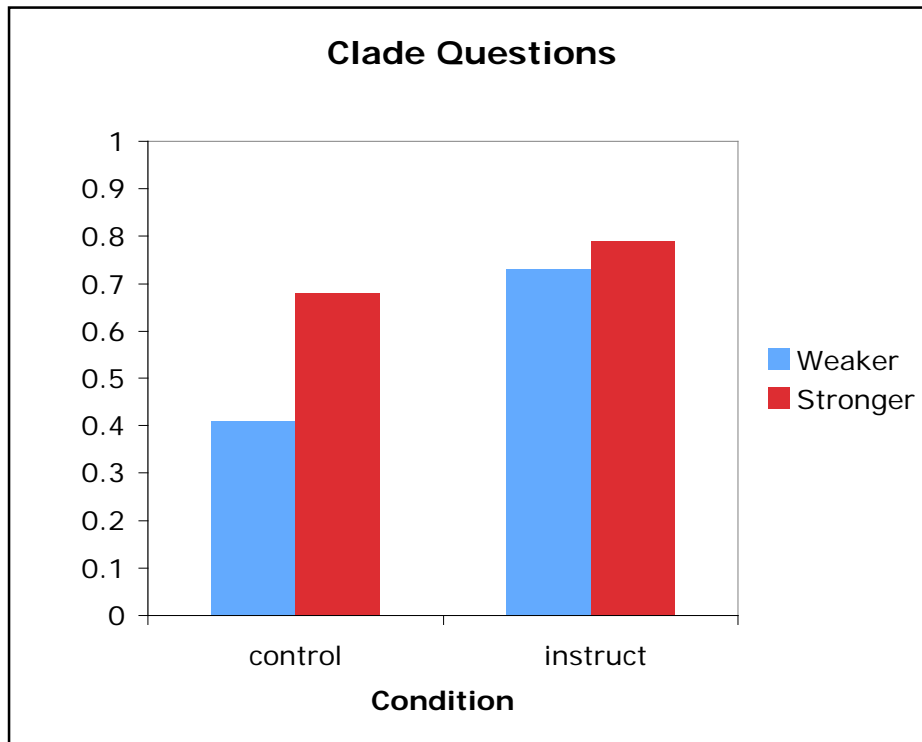


Figure 9. Interaction between condition and biology background for the nested clades questions.