EFFECTS OF PHYLOGENETIC TREE STYLE ON STUDENT COMPREHENSION

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Jonathan Andrew Dees

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Biological Sciences

May 2017

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Effects of Phylogenetic Tree Style on Student Comprehension

By

Jonathan Andrew Dees

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Dr. Jennifer Momsen
Chair

Dr. Julia Bowsher

Dr. Jeffrey Boyer

Dr. Lisa Montplaisir

Dr. Jane Schuh

Approved:

May 10, 2017	Dr. Wendy Reed
Date	Department Chair

ABSTRACT

Phylogenetic trees are powerful tools of evolutionary biology that have become prominent across the life sciences. Consequently, learning to interpret and reason from phylogenetic trees is now an essential component of biology education. However, students often struggle to understand these diagrams, even after explicit instruction. One factor that has been observed to affect student understanding of phylogenetic trees is style (i.e., diagonal or bracket). The goal of this dissertation research was to systematically explore effects of style on student interpretations and construction of phylogenetic trees in the context of an introductory biology course. Before instruction, students were significantly more accurate with bracket phylogenetic trees for a variety of interpretation and construction tasks. Explicit instruction that balanced the use of diagonal and bracket phylogenetic trees mitigated some, but not all, style effects. After instruction, students were significantly more accurate for interpretation tasks involving taxa relatedness and construction exercises when using the bracket style. Based on this dissertation research and prior studies on style effects, I advocate for introductory biology instructors to use only the bracket style. Future research should examine causes of style effects and variables other than style to inform the development of research-based instruction that best supports student understanding of phylogenetic trees.

ACKNOWLEDGEMENTS

Thank you to my supervisory committee for their guidance on this dissertation, especially Dr. Jennifer Momsen as chair. Special thanks to the graduate students in the Momsen lab and my undergraduate research assistants, Caitlin Bussard, Danielle Freiermuth, and Robert Zastre. This research complied with all Institutional Review Board regulations (protocol #SM12217) and was supported by a STEM Education Ph.D. Program fellowship from North Dakota State University and grants from the National Science Foundation (DRL-1420321 and DUE-1156974).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIX TABLES

LIST OF APPENDIX FIGURES

INTRODUCTION

Phylogenetic trees are often viewed simply as visual representations of evolution, but this perspective obscures their capacity to be a powerful framework for understanding biology. These diagrams are the main tool used by biologists to assess evidence of evolution and also provide an efficient method for organizing our knowledge of biological diversity (Baum et al., 2005; Novick and Catley, 2007; Thanukos, 2009; Wiley, 2010). Further, phylogenetic trees facilitate reasoning about biological phenomena from an evolutionary point of view (O'Hara, 1988; Gregory, 2008), and "nothing in biology makes sense except in the light of evolution" (Dobzhansky, 1973). With phylogenetic trees as a framework, biologists can investigate the current patterns of life on Earth, test hypotheses regarding the evolutionary processes that produced those patterns, and infer how those patterns will change in the future as a result of evolution (Novick and Catley, 2013).

Because evolution is a unifying theory and fundamental concept of biology (Dobzhansky, 1964; American Association for the Advancement of Science, 2011), and due to improvements in phylogenetic inference methods and DNA sequencing technology (Baum et al., 2005; Omland et al., 2008), phylogenetic trees have become increasingly important across the life sciences (Baum and Offner, 2008). As a consequence, learning to interpret and reason from phylogenetic trees is now an essential component of biology education (O'Hara, 1997; Lents et al., 2010; Novick and Catley, 2016). However, students often struggle to understand these diagrams (Meir et al., 2007; Halverson, 2011; Catley et al., 2013; Novick and Catley, 2013), even after instruction (Phillips et al., 2012; Smith et al., 2013; Dees et al., 2014). Although evidence is limited, student difficulties with phylogenetic trees have been attributed to a number of factors, including principles of visual cognition (Novick and Catley, 2007, 2013), visual scanning and processing biases (Novick et al., 2012), and misinterpretations of evolution in general (Gregory, 2008; Omland et al., 2008).

Through observations of students struggling to interpret and construct phylogenetic trees during introductory biology courses and an extensive review of the literature, I became interested in researching variables that potentially influence student comprehension. One such variable that can easily be controlled by instructors and has been observed to impact student understanding of phylogenetic trees is style (i.e., diagonal or bracket; Baum and Offner, 2008). Although previous research has shown that students often demonstrate better understanding of bracket phylogenetic trees compared to the diagonal style (Novick and Catley, 2007, 2013), these studies had important limitations. Both investigations used voluntary surveys for data collection, and students may not take surveys as seriously as coursework that affects their academic standing (Sundberg, 2002). In addition, neither study included construction tasks, which are common instructional activities for phylogenetic trees (e.g., Goldsmith, 2003; Burks and Boles, 2007; Lents et al., 2010; Eddy et al., 2013; Bokor et al., 2014; Lampert and Mook, 2015). The goal of this dissertation research was to address these limitations by systematically investigating effects of style on student interpretations and construction of phylogenetic trees in the context of an introductory biology course.

Unfortunately, an extensive review of the literature uncovered few research-based tools to support a systematic investigation of style effects. Therefore, my MS thesis (published as Dees et al., 2014) developed novel assessment items and a rubric for measuring student understanding of taxa relatedness on phylogenetic trees. Similarly, the first study of this dissertation (published as Dees and Momsen, 2016) developed novel assessment items and a rubric for measuring accuracy of student-constructed phylogenetic trees. Subsequently, the second study (Dees et al., accepted) took the first direct step toward fulfilling the research goal by developing isomorphic assessment items for equivalent diagonal and bracket phylogenetic trees. These assessment items were given to introductory biology students, and the research-based tools from previous studies were used to successfully investigate style effects. However, the second study concerned only pre-instructional data and was limited by uncontrolled variables. The third and final study of this dissertation used the tools and knowledge acquired from all previous investigations to satisfy the research goal by collecting data throughout an introductory biology course and by controlling as many variables as was feasible. The results of this dissertation research inform instruction on the critical concept of evolution, as well as future research on student understanding of phylogenetic trees.

STUDY 1. STUDENT CONSTRUCTION OF PHYLOGENETIC TREES IN AN INTRODUCTORY BIOLOGY COURSE[1]

Abstract

Phylogenetic trees have become increasingly essential across biology disciplines. As a consequence, learning about phylogenetic trees has become an important component of biology education and an area of interest for biology education research. Construction tasks, in which students generate phylogenetic trees from some type of data, are often used for instruction. However, the impact of these exercises on student learning is uncertain, in part due to our fragmented knowledge of what students construct during the tasks. The goal of this project was to develop a more robust method for describing student-generated phylogenetic trees, which will support future studies that attempt to link construction tasks with student learning. Through iterative examination of data from an introductory biology course, we developed a method for describing student-generated phylogenetic trees in terms of style, conventionality, and accuracy. Students used the diagonal style more often than the bracket style for construction tasks. The majority of phylogenetic trees were constructed conventionally, and variable orientation of branches was the most common unconventional feature. In addition, the majority of phylogenetic trees were generated correctly (no errors) or adequately (minor errors only) in terms of accuracy. Suggesting extant taxa are descended from other extant taxa was the most common major error, while empty branches and extra nodes were common minor errors. The method we developed to describe student-constructed phylogenetic trees uncovered several trends that warrant further

---

[1] The material in this chapter was co-authored by Jonathan Dees and Dr. Jennifer Momsen and is published as Dees and Momsen (2016). The publication is reproduced here with permission from the copyright holder. Jonathan Dees designed the assessment items, contributed to data collection, completed the data analyses, drafted this chapter, and contributed to revisions of this chapter. Dr. Jennifer Momsen contributed to data collection and revisions of this chapter.

investigation. For example, while diagonal and bracket phylogenetic trees contain equivalent information, student preference for using the diagonal style could impact comprehension. In addition, despite a lack of explicit instruction, students generated phylogenetic trees that were largely conventional and accurate. Surprisingly, accuracy and conventionality were dependent on each other. Our method for describing phylogenetic trees constructed by students is based on data from one introductory biology course at one institution, and the results are limited. We encourage researchers to use our method as a baseline for developing a more generalizable tool, which will support future studies that attempt to link construction tasks with student learning.

Introduction

Phylogenetic trees are visual representations that depict hypothesized evolutionary relationships among nested groups of taxa (Novick and Catley, 2007; Baum and Offner, 2008). These tools are used primarily by evolutionary biologists to evaluate evidence for evolution (Baum et al., 2005), but phylogenetic trees have also become increasingly essential in nearly all disciplines of biology (Omland et al., 2008). Consequently, learning about phylogenetic trees has become an important component of biology education and an area of interest for research.

Undergraduates in the sciences should develop competence with visual representations in general (National Research Council, 2012). However, “tree-thinking” skills are particularly important for students due to the subject matter of phylogenetic trees. Evolution is a unifying theory in biology (Dobzhansky, 1973) and a core concept for biological literacy (American Association for the Advancement of Science, 2011). As conceptual models, phylogenetic trees offer insights into patterns and processes of evolution and provide powerful scaffolding for learning about biology (Novick and Catley, 2007). However, the utility of phylogenetic trees is tempered by widespread misinterpretations among biology students (Meir et al., 2007; Novick and Catley, 2013; Dees et al., 2014) that potentially create obstacles to understanding evolution (Meir et al., 2007; Gregory, 2008). The importance of phylogenetic trees for biologists and lack of interpretation skills among students necessitate continued research to address this discrepancy.

Some of the most common instructional activities concerning phylogenetic trees are construction tasks, in which students build phylogenetic trees from provided or self-generated data. Such tasks assume that constructing phylogenetic trees will improve interpretation skills, but research exploring this relationship is limited and conflicting. Eddy et al. (2013) observed that scaffolded construction tasks significantly improved student interpretations of phylogenetic trees. However, Halverson (2011) concluded students must develop interpretation skills before construction abilities. Thus, effects of construction tasks on student learning remain uncertain.

One reason that such effects are uncertain could be that what students construct during the tasks is largely unknown. Halverson (2011) only characterized representations from students as valid phylogenetic trees or one of several alternatives (e.g., dichotomous keys, flow charts, food webs, pictures, and lists), while the conflicting study by Eddy et al. (2013) did not describe representations created by students. A third study, Young et al. (2013), was limited to measuring the prevalence of basic phylogenetic tree characteristics (e.g., single common ancestor, branches, and hierarchy) in representations generated by students before and after instructional activities.

Overall, descriptions of student-constructed phylogenetic trees are fragmented, and the impact of construction exercises on student learning is unresolved. The goal of this study was to develop a more robust method for describing phylogenetic trees generated by students, which will support future research that attempts to link construction exercises with student learning. Specifically, a method for describing student-constructed phylogenetic trees in terms of style, conventionality, and accuracy emerged through answering the following research questions:

1) Which style of phylogenetic tree (diagonal or bracket) do introductory biology students prefer to construct?
2) How conventionally do introductory biology students construct phylogenetic trees, and what are the common deviations?
3) How accurately do introductory biology students construct phylogenetic trees, and what are the common errors?

Methods

This investigation was conducted in the context of an introductory biology course for science and related majors at a large, public university in the midwestern United States. The large-enrollment course (n=88) served students at various stages in their academic programs (24% freshmen, 33% sophomores, 18% juniors, and 25% seniors) and was comprised of three units: evolution (first six weeks), form and function (next five weeks), and ecology (last five weeks). Students often collaborated in self-selected groups of three or four individuals during instructional activities and assessments (Johnson et al., 1998; Smith, 2000), including exams with individual and group components (Cortright et al., 2003). All classes were observed, and instructional materials and assessments were collected to document instruction.

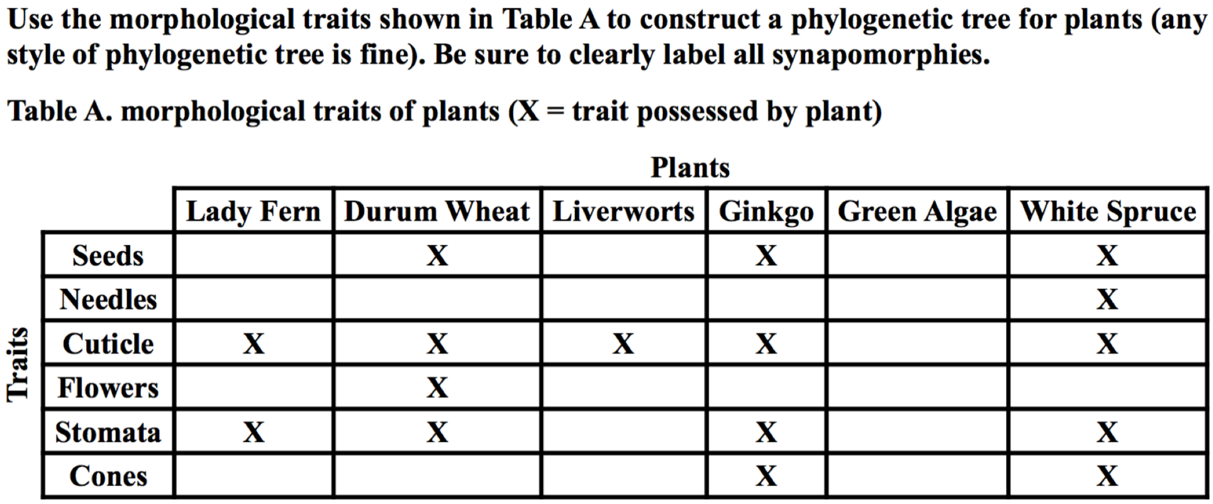
Phylogenetic Tree Instruction

Phylogenetic trees were first introduced during the evolution unit through reading in the textbook (Freeman, 2011), individual and group reading quizzes, and a series of multiple-choice questions presented by the instructor and answered by students using letter cards (Freeman et al., 2007). These tasks familiarized students with basic characteristics of phylogenetic trees, such as nodes and monophyletic groups, and introduced the critical concept of taxa relatedness (Novick and Catley, 2013; Dees et al., 2014). Responses to letter card questions were ungraded but public, which allowed students to view answers from neighbors in preparation for collaborative learning activities. Correct answers using appropriate reasoning were established through group and class discussions, and by students iteratively responding to the same or similar letter card questions if necessary. All phylogenetic trees used during the course were cladograms, in which only branch patterns contain reliable information (Gregory, 2008). The instructor briefly presented examples of phylograms (branches scaled for degree of divergence) and chronograms (branches scaled for time), but students were never asked to reason from them during the course.

After the phylogenetic tree introduction, students completed a group homework featuring a diagonal phylogenetic tree of chordates accompanied by several interpretation questions. The prompts specifically concerned trait possession, synapomorphies, most recent common ancestry, monophyletic groups, taxa relatedness, and convergent evolution. Student interpretations of taxa relatedness and convergent evolution submitted by groups were exclusively incorrect (i.e., failed to include both the correct answer and correct reasoning). Responses also exhibited a wide array of inappropriate reasoning strategies (Morabito et al., 2010; Dees et al., 2014), which compelled the instructor to respond with feedback and remedial activities. Phylogenetic trees were revisited during class through additional letter card questions with subsequent discussions. It is important to note that students were not asked to construct phylogenetic trees prior to data collection.

Data Collection

The first phylogenetic tree construction task (Figure 1) appeared on the group section of the evolution unit exam during the sixth week of class. Group sections of exams were intended to be more challenging than individual components, yet not overwhelming. Thus, although students were not asked to build phylogenetic trees prior to the evolution unit exam, the phylogenetic tree resulting from the construction task is relatively simple, with no unresolved nodes or convergent evolution. Following the evolution unit exam, phylogenetic trees continued to appear throughout the course when appropriate. For example, phylogenetic trees were used in the form and function unit to help students reason about cardiovascular adaptations of animals. Two weeks before the comprehensive final exam, students completed a course review homework that included a series of interpretation questions for a diagonal phylogenetic tree of tetrapods. However, students were not asked to construct phylogenetic trees between the evolution unit exam and final exam.



Use the morphological traits shown in Table A to construct a phylogenetic tree for plants (any style of phylogenetic tree is fine). Be sure to clearly label all synapomorphies.

Table A. morphological traits of plants (X = trait possessed by plant)

	Plants					
Traits	**Lady Fern**	**Durum Wheat**	**Liverworts**	**Ginkgo**	**Green Algae**	**White Spruce**
Seeds		X		X		X
Needles						X
Cuticle	X	X	X	X		X
Flowers		X				
Stomata	X	X		X		X
Cones				X		X

Figure 1. Initial phylogenetic tree construction task from the group component of the evolution unit exam during the sixth week of class.

The second phylogenetic tree construction exercise (Figures A1-A2 in Appendix A) was placed on the individual section of the comprehensive final exam. The two versions of the task involve different taxa and traits but result in the same branch pattern, with no unresolved nodes or convergent evolution. In preparation for the subsequent group component of the final exam, two students from each group of four received version A, while the other two students received version B. For groups of three, at least one student received each version. The third phylogenetic tree construction exercise (Figure A3 in Appendix A) was created by merging both versions of the construction prompt from the individual component of the final exam into a larger and more challenging task for the group component of the final exam. The resulting phylogenetic tree does not contain unresolved nodes, but unlike the earlier construction exercises, convergent evolution is present. All phylogenetic trees constructed for the group component of the evolution unit exam (n=23), individual component of the final exam (n=77), and group component of the final exam (n=22) constitute the data to be analyzed for this investigation.

Rubric Development and Coding

Rubrics were developed to evaluate the phylogenetic trees produced by students during construction tasks in terms of style, conventionality, and accuracy. A general inductive approach was used for rubric development, which is a qualitative research method that allows categories to emerge from iterative examination of data rather than predetermined hypotheses (Thomas, 2006). Phylogenetic trees were coded for style as diagonal or bracket (Figure 2; ladder and tree formats described by Novick and Catley, 2007), and rare cases containing diagonal and bracket features were coded as the predominant style. For example, a mainly diagonal phylogenetic tree with one divergence constructed in the bracket style was coded as diagonal.

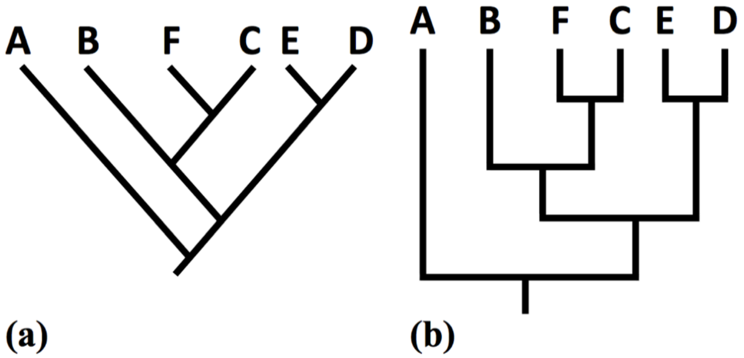
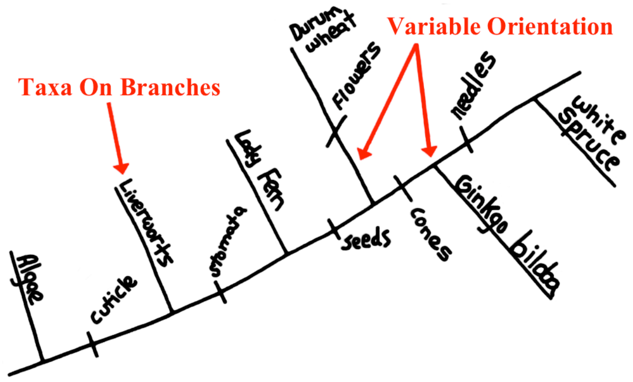
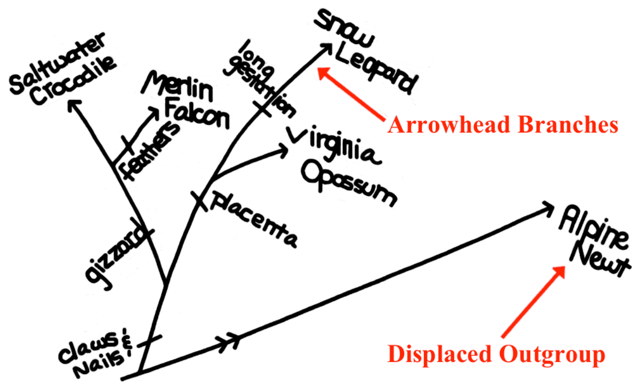


Figure 2. Two common phylogenetic tree styles with equivalent branching patterns: (a) diagonal and (b) bracket (Dees et al., 2014; adapted from Gregory, 2008).

Conventionality was used to describe features of student-generated phylogenetic trees that were not accounted for by style or accuracy. Specifically, conventionality was determined based on the presence or absence of features that are unusual but not incorrect (Table 1). For example, the outgroup is usually the left-most terminal taxon of phylogenetic trees oriented in the vertical direction, especially in high school and college textbooks (Catley and Novick, 2008; Novick et al., 2012). Placing the outgroup on the right side of vertical phylogenetic trees is unusual but equally correct (coded as displaced outgroup). Phylogenetic trees containing one or more unconventional features were coded as unconventional, while all others were coded as conventional.

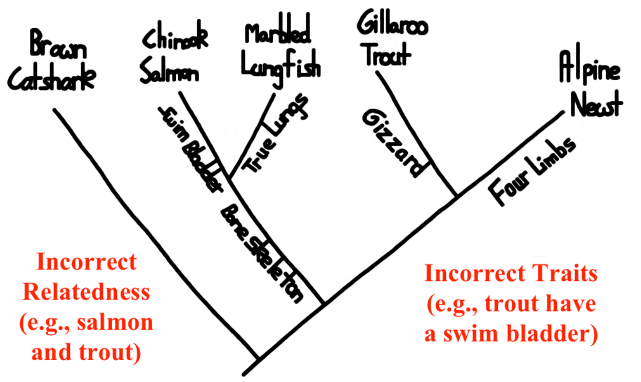
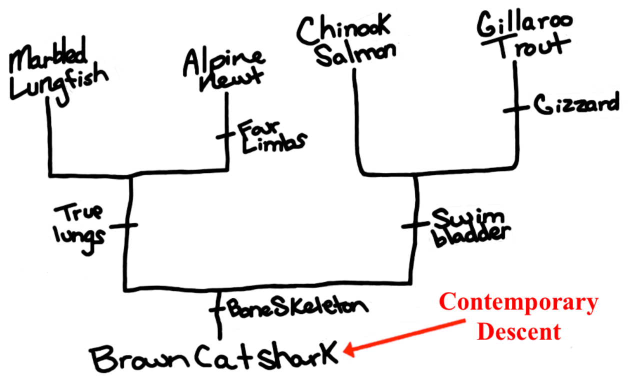
Table 1. Unconventional features observed in phylogenetic trees constructed by students.

Feature	**Description**	**Student-Generated Example**
Variable Orientation	Branches are not oriented in a consistent direction.	
Taxa On Branches	Taxa are on the branches rather than at the tips.	
Arrowhead Branches	Some or all branches are drawn with arrowheads.	
Displaced Outgroup	Outgroup is placed in an unconventional location.	

Phylogenetic trees constructed by students were assessed for accuracy based on major and minor errors. Major errors, such as incorrect placement of taxa, prevent students from correctly interpreting taxa relatedness or trait possession (Table 2). Minor errors, such as empty branches, are structural in nature and do not inhibit students from correctly interpreting taxa relatedness or trait possession (Table 3). Phylogenetic trees containing one or more major errors were coded as incorrect, while those with one or more minor errors but no major errors were coded as adequate.

Note that incorrect phylogenetic trees could also contain minor errors in addition to major errors (e.g., incorrect relative placement of taxa and empty branches). Phylogenetic trees with no major or minor errors were coded as correct. Style, conventionality, and accuracy were each evaluated by two independent raters with 96.7%-100.0% agreement (kappa of 0.93-1.00; Cohen, 1960).

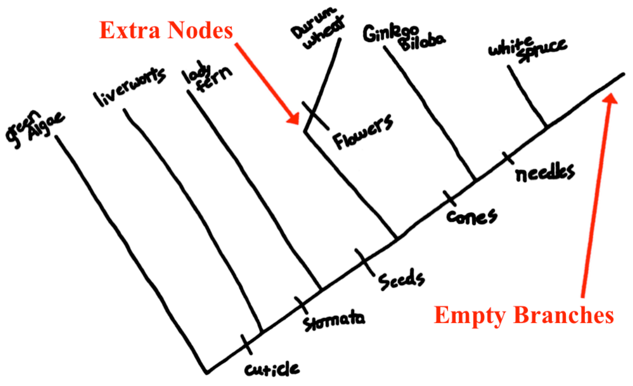
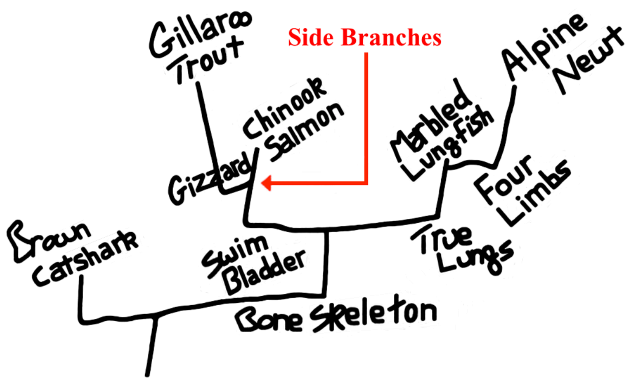
Table 2. Major errors observed in phylogenetic trees constructed by students.

Major Error	Description	Student-Generated Example
Incorrect Relatedness	Relative placement of taxa is incorrect based on provided data.	
Incorrect Traits	Traits assigned to taxa are incorrect based on provided data.	
Contemporary Descent	Representation implies extant taxa are directly descended from one or more other extant taxa.	

Statistical Analyses

Phylogenetic trees constructed for the individual component of the comprehensive final exam (only data set from individuals) were analyzed for associations between task version, style, conventionality, and accuracy using Fisher's exact tests (Fisher, 1934). The null hypothesis is that one variable of phylogenetic tree construction, such as style, is independent of a second variable, such as conventionality. An exact test for goodness-of-fit was used to analyze the distribution of diagonal and bracket phylogenetic trees from the individual component of the final exam, where the null hypothesis is an equal distribution (McDonald, 2014). Phylogenetic trees from the group component of the evolution unit exam and group section of the final exam were not analyzed for variable associations or style distribution due to small sample sizes and low statistical power.

Table 3. Minor errors observed in phylogenetic trees constructed by students.

Minor Error	**Description**	**Student-Generated Example**
Empty Branches	One or more branches are not linked to taxa.	
Extra Nodes	One or more nodes do not denote divergence of taxa (bifurcation).	
Side Branches	One or more nodes do not correspond with a fork structure (applies to bracket style only).	

Results

Phylogenetic trees generated by introductory biology students during the group section of the evolution unit exam (n=23), individual section of the final exam (n=77), and group section of the final exam (n=22) were each evaluated for style, conventionality, and accuracy. The majority of phylogenetic trees were constructed in the diagonal style, conventionally, and either correctly or adequately in terms of accuracy across all three assessments for this investigation.

Construction Style

Students created diagonal phylogenetic trees more often than bracket phylogenetic trees during all three assessments (Figure 3). Overall, 80% of phylogenetic trees were generated in the diagonal style. For the individual section of the final exam (only data obtained from individuals), the distribution of styles significantly differed from an equal distribution ($p<0.001$). In addition, style was independent of task version for the individual component of the final exam ($p=1.00$).

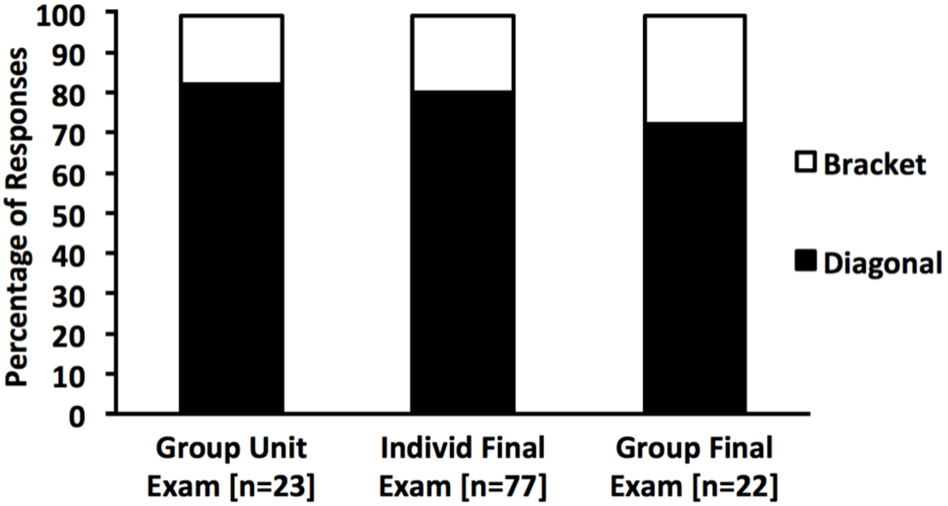


Figure 3. Styles used by students to construct phylogenetic trees.

Construction Conventionality

The majority of phylogenetic trees generated by students were coded as conventional for all assessments (Figure 4). Overall, 64% of phylogenetic trees were created conventionally. The most common unconventional features were branches with variable orientations and taxa placed on branches (Table 4). For the individual component of the final exam (only data obtained from individuals), conventionality was independent of style ($p=0.77$) and task version ($p=0.20$).

Construction Accuracy

The majority of phylogenetic trees created by students were correct (no major or minor errors) or adequate (one or more minor errors but no major errors) in terms of accuracy (Figure

5). Overall, 73% of phylogenetic trees were created correctly or adequately, including the group section of the final exam with convergent evolution (64% correct or adequate). The most common major construction error was contemporary descent (extant taxa are descended from other extant taxa), while empty branches and extra nodes were very common minor construction errors for all three assessments (Table 5). For the individual section of the final exam (only data obtained from individuals), accuracy was independent of style (p=0.77) and task version (p=0.71). Conversely, accuracy was dependent on conventionality (p=0.01), as unconventional phylogenetic trees were more likely to be incorrect compared to conventional phylogenetic trees.

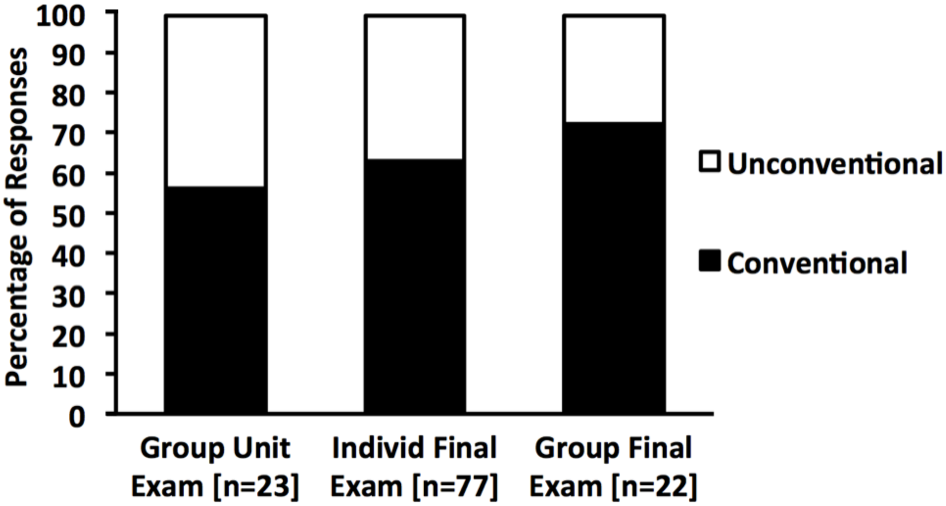


Figure 4. Conventionality of phylogenetic trees constructed by students.

Table 4. Unconventional features observed in phylogenetic trees constructed by students.

Feature	Group Unit Exam [n=23]	Individual Final Exam [n=77]	Group Final Exam [n=22]
Variable Orientation	7 (30%)	15 (19%)	4 (18%)
Taxa On Branches	3 (13%)	8 (10%)	2 (9%)
Arrowhead Branches	1 (4%)	6 (8%)	0 (0%)
Displaced Outgroup	2 (9%)	5 (6%)	1 (5%)

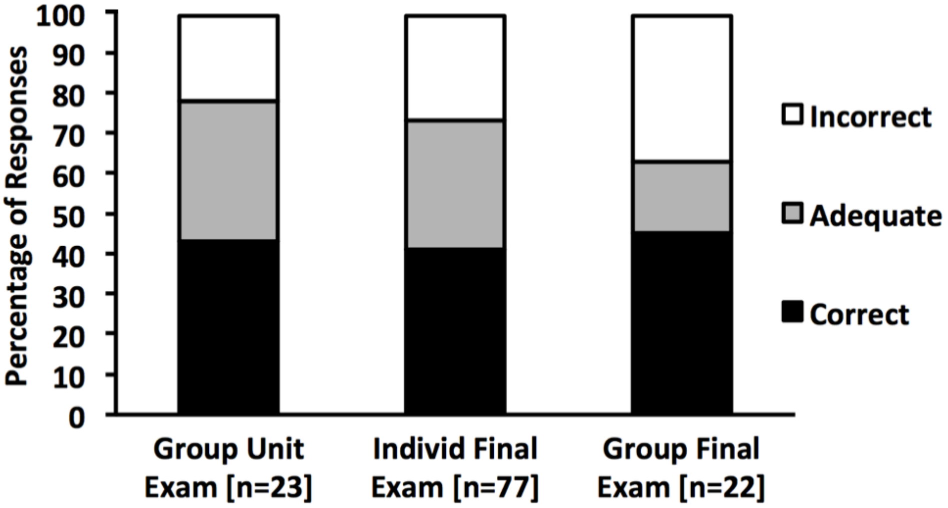


Figure 5. Accuracy of phylogenetic trees constructed by students.

Table 5. Major and minor errors observed in phylogenetic trees constructed by students.

Major Error	**Group Unit Exam [n=23]**	**Individual Final Exam [n=77]**	**Group Final Exam [n=22]**
Incorrect Relatedness	0 (0%)	10 (13%)	3 (14%)
Incorrect Traits	1 (4%)	10 (13%)	3 (14%)
Contemporary Descent	5 (22%)	12 (16%)	5 (23%)
Minor Error	**Group Unit Exam [n=23]**	**Individual Final Exam [n=77]**	**Group Final Exam [n=22]**
Empty Branches	6 (26%)	31 (40%)	5 (23%)
Extra Nodes	10 (43%)	30 (39%)	8 (36%)
Side Branches	0 (0%)	7 (9%)	1 (5%)

Discussion

Construction tasks are some of the most common instructional activities for phylogenetic trees, but the impact of these exercises on student learning is uncertain (Halverson, 2011; Eddy et al., 2013). One factor contributing to this uncertainty could be our fragmented knowledge of what students construct during the tasks (Halverson, 2011; Young et al., 2013). The goal of this project was to develop a more robust method for describing student-generated phylogenetic trees, which will support future research that attempts to link construction tasks with learning. By examining responses to construction tasks from an introductory biology course, we developed a method for describing student-generated phylogenetic trees in terms of style, conventionality, and accuracy.

Construction Style

Students showed a preference for constructing diagonal phylogenetic trees across all three assessments (Figure 3). While diagonal and bracket phylogenetic trees are equivalent in terms of information, the choice of style could influence comprehension. For example, Novick and Catley (2013) concluded that students performed significantly better with bracket phylogenetic trees on a variety of interpretation tasks, regardless of background in biology. Thus, our students favored the style that may hinder their interpretation abilities. However, we caution that the present study did not explicitly investigate how students interpret self-constructed phylogenetic trees, which is another important research topic for understanding the effects of construction tasks on learning.

Construction Conventionality

The majority of students generated conventional phylogenetic trees for each assessment (Figure 4), despite receiving no explicit instruction on how to construct phylogenetic trees from data. Therefore, many students adopted conventions on their own, presumably through repeated exposure to phylogenetic trees. Surprisingly, accuracy was dependent on conventionality, in that unconventional phylogenetic trees were more likely to be incorrect. The cause of this outcome is unknown, but we speculate that students who constructed unconventional phylogenetic trees may have had less experience with the diagrams, and thus were also more likely to generate incorrect phylogenetic trees. Lack of experience could be due to many factors, such as class absences (rare during phylogenetic tree instruction), non-participation in group instructional activities, or poor study habits. Unfortunately, we have no way of systematically investigating this result due to the group nature of instruction and unknown study habits of our students.

Construction Accuracy

The majority of phylogenetic trees were correct or adequate in terms of accuracy across assessments (Figure 5), including the group section of the final exam when convergent evolution was present. Thus, students were relatively proficient at constructing phylogenetic trees, which is notable considering the lack of explicit instruction. However, we caution that minor construction errors (Table 3), which were common during all three assessments (Table 5), are not necessarily without consequences. Major errors, such as incorrect relative placement of taxa, directly impact interpretations of trait possession and taxa relatedness, which are skills that were assessed during the course. Minor errors could influence student thinking in other ways that are more difficult to measure. For example, empty branches on phylogenetic trees could reflect a common belief that trait evolution occurs only at nodes (Baum et al., 2005). Establishing relationships between each construction error and specific misinterpretations is an important goal for future research.

Limitations

Although students constructed diagonal phylogenetic trees more frequently than bracket phylogenetic trees, this result could have been affected by the curriculum (Table A1 in Appendix A). The textbook (Freeman, 2011) contained only bracket phylogenetic trees, and instruction was also biased toward the bracket style. However, assessments (i.e., homework, reading quizzes, and exams) were skewed toward diagonal phylogenetic trees. Because assessment strongly influences learning behaviors (e.g., Cohen-Schotanus, 1999; Wormald et al., 2009), students may have been tacitly steered toward using the diagonal style. Future classroom studies involving style should control the curriculum such that both styles are equally represented in all aspects of the course.

Students were required to build one phylogenetic tree, in the style of their choice, during the individual section of the final exam (only data set from individuals). Thus, the study design for style was between-student rather than a stronger within-student approach. It is particularly an issue in this case due to the strong preference for constructing diagonal phylogenetic trees, which resulted in a smaller number of bracket phylogenetic trees for comparison. Due to this limitation, no conclusions should be drawn from this study about the effects of style on conventionality and accuracy. Future investigations should use a stronger within-student design that requires students to generate both diagonal and bracket phylogenetic trees during construction tasks.

Two major construction errors, incorrect relatedness and incorrect traits, were uncommon in phylogenetic trees constructed by students (Table 5). However, some of these errors may have been provoked by the assessment prompts, which did not state the polarity of traits. We assumed that introductory biology students would treat the provided traits as derived rather than ancestral characters (i.e., traits were gained over time). Although we did not find any evidence to suggest that students assumed the traits were ancestral, it is possible that the lack of polarity information in our prompts affected student reasoning. Future studies could protect against this possibility by explicitly providing polarity information to students before construction tasks or within prompts.

Conclusions

The impact of phylogenetic tree construction tasks on student learning is uncertain based on literature, and one factor contributing to this uncertainty could be our fragmented knowledge of what students construct during the tasks. We developed a method for describing phylogenetic trees generated by students, which will support future research that attempts to link construction tasks with student learning. However, our method is based on data from one introductory biology course at one institution, and the results likely do not reflect undergraduate biology students as a whole. Other researchers and instructors may find additional errors and unconventional features that were not present or not recognized in our data. We encourage researchers to use our method of style, conventionality, and accuracy as a baseline for developing a more generalizable tool. In addition, we urge others to use our method for research that advances the broader goal of linking construction tasks with student learning.

STUDY 2. EFFECTS OF PHYLOGENETIC TREE STYLE ON STUDENT COMPREHENSION IN AN INTRODUCTORY BIOLOGY COURSE[2]

Abstract

Phylogenetic trees are now an important component of biology education, but their utility in classrooms is compromised by widespread misinterpretations among students. One factor that may contribute to student difficulties is style, as diagonal and bracket phylogenetic trees are both commonly used in biology. Previous research using surveys found that students performed better with bracket phylogenetic trees across a variety of interpretation tasks. The present study builds on prior research by comparing how students interpret diagonal and bracket phylogenetic trees in the context of an introductory biology course and by expanding the style comparison to include construction tasks. Students performed significantly better with bracket phylogenetic trees for some, but not all, interpretation tasks. In addition, students who constructed bracket phylogenetic trees were significantly more accurate compared to those who used the diagonal style. Thus, our results reinforce previous research for interpretations, and the performance gap between styles extended to construction tasks. It remains to be seen, however, if such differences persist after instruction that balances the use of diagonal and bracket phylogenetic trees.

Introduction

Phylogenetic trees are essential in nearly all areas of biology (Baum et al., 2005; Omland et al., 2008), and consequently, learning about phylogenetic trees is also an important component of biology education (O'Hara, 1997). Phylogenetic trees are a powerful framework for thinking

---

[2] The material contained in this chapter was co-authored by Jonathan Dees, Danielle Freiermuth, and Dr. Jennifer Momsen and has been accepted for publication as Dees et al. in the American Biology Teacher. Jonathan Dees developed the assessment items, contributed to data collection and analyses, drafted this chapter, and contributed to revisions. Danielle Freiermuth contributed to data analyses. Dr. Jennifer Momsen contributed to data collection and revisions of this chapter.

and learning about biology from an evolutionary perspective by serving as visual representations, analytical tools, and concept models (Novick and Catley, 2007; Baum and Offner, 2008; Wiley, 2010; Smith et al., 2013). However, the utility of phylogenetic trees for education is diminished by widespread misinterpretations among high school, introductory, and upper-division students (Morabito et al., 2010; Halverson, 2011; Catley et al., 2013; Novick and Catley, 2013; Blacquiere and Hoese, 2016). Such misinterpretations often persist through explicit instruction (Phillips et al., 2012; Dees et al., 2014) and create or reinforce student difficulties with understanding evolution (Meir et al., 2007; Gregory, 2008). The importance of phylogenetic trees for biologists and lack of interpretation skills among students is a disparity that warrants further investigation, such that researchers and teachers can determine the best course of action for instruction.

Students come across two equivalent styles of phylogenetic tree, diagonal and bracket, in textbooks, journals, and other resources (Figure 6; Catley and Novick, 2008). Based on classroom observations, style can impact how students interpret phylogenetic trees (Baum and Offner, 2008; Halverson et al., 2011). However, only two studies have explicitly examined the effects of style on student comprehension. Novick and Catley (2007) used translation tasks to reveal perceptual differences between diagonal and bracket phylogenetic trees. Students with varying backgrounds in biology translated different visual representations of evolution, including diagonal and bracket phylogenetic trees, from one representation to another while retaining the same information. The investigators found that accuracy decreased whenever diagonal phylogenetic trees were involved in translations, and the effect was larger for students with less experience in biology. In a second study, Novick and Catley (2013) used a variety of tasks, such as identifying traits shared by taxa, recognizing monophyletic groups, and evaluating relatedness, to compare how students interpret diagonal and bracket phylogenetic trees. Accuracy was lower for the diagonal style across nearly all interpretation tasks, and this effect was often observed regardless of background in biology.

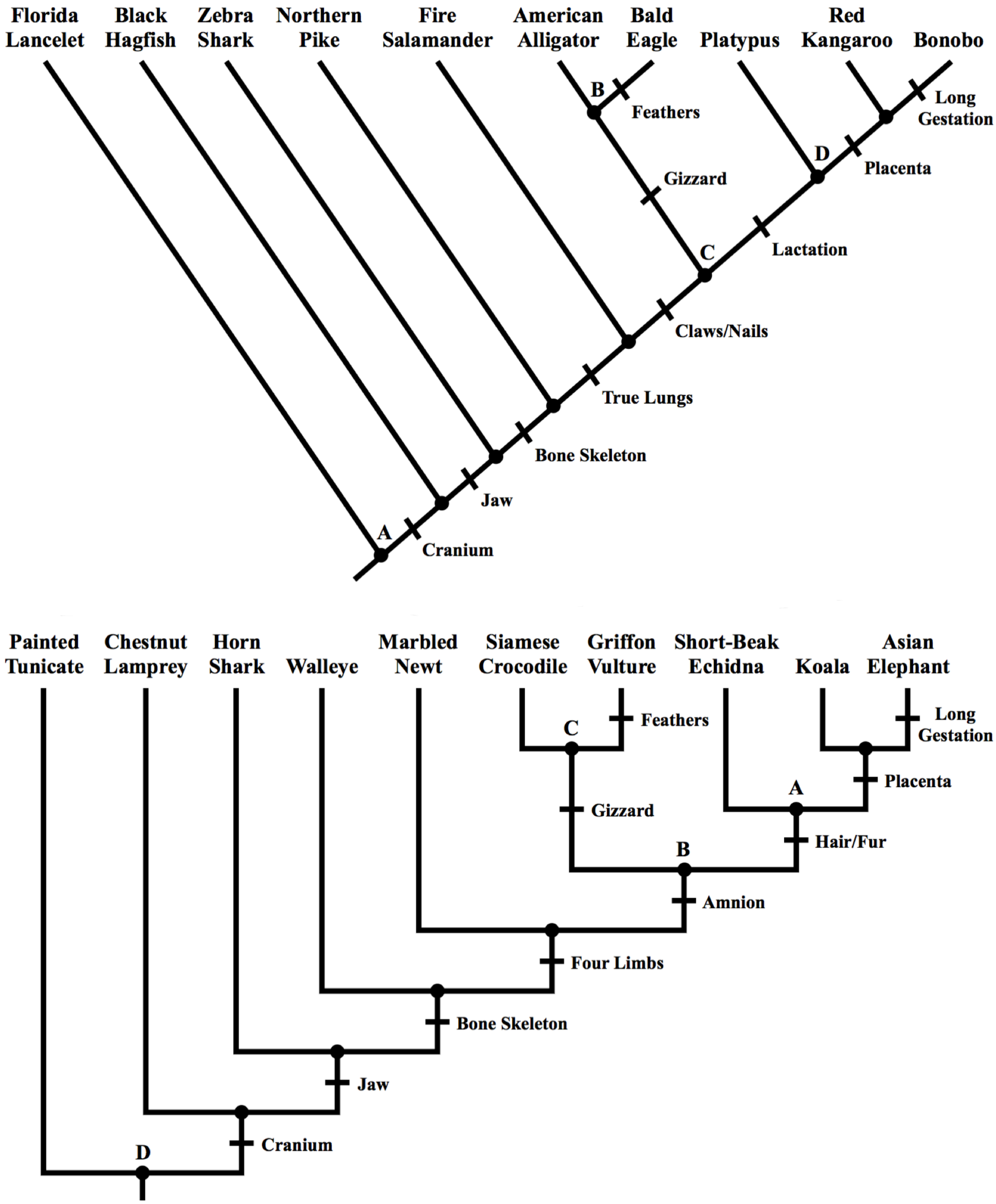


Figure 6. Diagonal (top) and bracket (bottom) phylogenetic trees with identical branch patterns.

Although the results of Novick and Catley (2007, 2013) are intriguing for instructors, the data for these innovative studies were obtained through surveys in which students had no stake in the outcome. Such data are valuable as a starting point, but students may not be motivated to take surveys that will not affect their academic standing as seriously as graded coursework (Sundberg, 2002). Thus, one goal of this study was to reinforce and build on existing, survey-based research by comparing how students interpret diagonal and bracket phylogenetic trees in the context of an introductory biology course. Specifically, we asked the following research question:

1) Do introductory biology students demonstrate differential interpretation abilities for provided phylogenetic trees based on style?

Beyond reinforcing existing research on student interpretations, we examined style effects for phylogenetic tree construction, which is of particular interest to teachers. Construction tasks, in which students build phylogenetic trees from various forms of data, are common instructional activities for phylogenetic trees (e.g., Singer et al., 2001; Goldsmith, 2003; Julius and Schoenfuss, 2006; Burks and Boles, 2007; Lents et al., 2010; Eddy et al., 2013; Bokor et al., 2014; Lampert and Mook, 2015). Although several studies examined how accurately students construct phylogenetic trees from data (Meir et al., 2007; Halverson, 2011; Hobbs et al., 2013; Young et al., 2013), only one investigation accounted for style (Dees and Momsen, 2016). However, students were allowed to choose which style to use and overwhelmingly constructed diagonal phylogenetic trees (80%), resulting in an inadequate sample of bracket phylogenetic trees for comparison. Thus, the second goal of this study was to determine if style impacts how accurately introductory biology students construct phylogenetic trees. Specifically, we asked the following research question:

2) Do introductory biology students demonstrate differential construction abilities for phylogenetic trees based on style?

To summarize, we expanded on existing research that examined effects of style on student understanding of phylogenetic trees in two ways. First, data for this study were collected through introductory biology coursework rather than through surveys. Second, we included phylogenetic tree construction in addition to interpretation tasks. Our results inform researchers and teachers who are working to develop phylogenetic tree instruction that best facilitates student learning.

Methods

Data for this study were collected in an introductory biology course (n=107) for science and related majors (Table 6) at a public university in the midwestern United States. The course served students at different stages in their academic programs (24% freshmen, 37% sophomores, 20% juniors, 19% seniors) and focused on evolution, form and function, and ecology. Instruction was learner-centered with an emphasis on collaboration (Johnson et al., 1998). Students regularly worked in assigned groups to build and evaluate conceptual models (Dauer et al., 2013), discuss clicker questions (Caldwell, 2007), and construct scientific arguments (Driver et al., 2000).

Table 6. Course enrollment by major group (n=107).

Major Group	% Students
Agricultural Sciences	25
Biological Sciences	43
Natural Resource Management	12
Pre-Professional Healthcare	10
Other Majors (Non-Science)	10

Data Collection

We developed an instrument to measure and compare student understanding of diagonal and bracket phylogenetic trees. Students interpreted two equivalent phylogenetic trees (Figure 6) in a number of ways, including identifying traits possessed by taxa, determining the most recent common ancestor of taxa, recognizing monophyletic groups, and evaluating taxa relatedness. We designed much of the instrument based on the work of Novick and Catley (2013), who suggested these interpretations are core skills associated with understanding phylogenetic trees. In addition, students were asked to use the diagonal and bracket phylogenetic trees to determine if extant taxa are descended from other extant taxa, which we refer to as contemporary descent. Earlier studies showed that students often view the straight lines of phylogenetic trees as single entities (Novick and Catley, 2007) that represent no evolutionary change (Meir et al., 2007). For example, students might assume the line from American alligators to node C in Figure 6 (top) represents American alligators rather than part of their evolutionary history. Based on this assumption, students could wrongly infer that bald eagles evolved from American alligators. Because diagonal phylogenetic trees are drawn with straight lines rather than brackets, students may demonstrate contemporary descent reasoning more often when interpreting diagonal phylogenetic trees. Thus, we designed questions that directly assess contemporary descent to test this prediction. Finally, students were asked to construct a phylogenetic tree of plants from provided morphological data in the style of their choice, such that construction accuracy could be compared across styles. All interpretation and construction tasks that were used for this investigation are available in Appendix B.

We distributed the instrument as a homework in an introductory biology course before classroom instruction on phylogenetic trees. In conjunction with the homework, students were assigned a short reading on phylogenetic trees that described their purpose and terminology (e.g., nodes and monophyletic groups). This reading from Freeman (2011) was modified to include examples of both phylogenetic tree styles. No reading was assigned from the online course textbook (Nature Education, 2013), which contained only bracket phylogenetic trees. Students were allowed to use any resource except each other and were given one week to submit the homework. Note that the order of assessment items was the same for all students: 1) diagonal phylogenetic tree and questions, 2) bracket phylogenetic tree and questions, and 3) construction task. Students were awarded full credit in the course for completing each question, but to avoid impacting motivation, the grading scheme was not disclosed prior to submission. Unfortunately, subsequent instruction on phylogenetic trees, which included a variety of interpretation and construction activities, was not controlled for style. Instructional materials collected from the course included both styles but were biased toward bracket phylogenetic trees. Thus, we were unable to collect post-instructional data that would provide an unbiased measure of style effects.

Data Coding

Responses to the trait possession, most recent common ancestor, and monophyletic group questions were coded as correct or incorrect. Questions about contemporary descent (extant taxa evolved from other extant taxa) required a yes or no answer with reasoning. Student answers and reasoning were each coded as correct or incorrect, where correct reasoning suggested extant taxa evolved from a common ancestor rather than each other (e.g., “No, bald eagles and the American alligator are in the same monophyletic group and share a recent common ancestor, but eagles did not evolve from alligators.”). Questions about taxa relatedness asked students to choose which of two taxa is more closely related to a reference taxon and provide reasoning. Answers were coded as correct or incorrect, while student reasoning was coded as correct, incorrect, or mixed using a published rubric (Dees et al., 2014). Correct reasoning referenced most recent common ancestry or monophyletic groups, whereas incorrect reasoning included comparing distances between taxa (branch tip proximity), counting nodes and synapomorphies between taxa, and using information that was not provided by phylogenetic trees (external insights). In some cases, student responses included correct and incorrect reasoning for taxa relatedness, which resulted in the mixed code.

At the end of the instrument, students were asked to construct a phylogenetic tree from provided data in the style of their choice. Phylogenetic trees constructed by students were first coded for style as diagonal or bracket (e.g., Figure 6). Rare cases (n=4) containing both diagonal and bracket features were coded as the predominant style. Student responses were coded for accuracy as correct, adequate, or incorrect using a published rubric (Dees and Momsen, 2016). Phylogenetic trees that contained major errors, such as incorrect relatedness or incorrect traits, were considered incorrect. Student responses that included only minor errors, such as empty branches or extra nodes, were coded as adequate. The distinction between major and minor errors was based on whether or not the errors impeded the ability of students to interpret taxa relatedness or trait possession. Finally, phylogenetic trees containing no major or minor errors were considered correct. Student responses to interpretation and construction tasks were coded by two independent raters with greater than 92% agreement (kappa greater than 0.89; Cohen, 1960).

Statistical Analyses

We treated responses to the isomorphic interpretation questions associated with diagonal and bracket phylogenetic trees as paired, nominal data. For dichotomous categories (e.g., correct or incorrect), we used the exact version of McNemar's test, which takes the paired nature of our data into account (McNemar, 1947; Rufibach, 2011). The null hypothesis is that an equal number of students switched in one direction (e.g., incorrect to correct) as in the other direction from one phylogenetic tree style to the other (McDonald, 2014). For trichotomous categories (e.g., correct, incorrect, or mixed), we used an extension of McNemar's test known as the Stuart-Maxwell test (Stuart, 1955; Maxwell, 1970). An exact binomial test was used to determine if students preferred to construct diagonal or bracket phylogenetic trees, where the null hypothesis is an equal number of each style (i.e., no preference). Because students constructed only one phylogenetic tree in the style of their choice, accuracy results were nominal (correct, adequate, or incorrect) but unpaired for the construction task. Thus, we used a Fisher's exact test (Fisher, 1934) to compare accuracy across styles, with the null hypothesis that accuracy is independent of style (McDonald, 2014).

Results

Data were collected through a pre-instructional homework distributed in an introductory biology course. Of 107 enrolled students, 92 (86%) responded to the interpretation tasks for both phylogenetic tree styles, which enabled within-student comparisons for accuracy across styles. In addition, 85 students (79%) completed the single construction task using the style of their choice, which facilitated a between-student comparison for accuracy across styles.

Interpretations

Students were significantly more accurate when interpreting bracket phylogenetic trees for some, but not all, interpretation tasks (Table 7). There was no difference between styles for recognizing traits possessed by taxa. However, students were more accurate in determining most recent common ancestors of taxa and identifying monophyletic groups when interpreting bracket phylogenetic trees. In addition, students were less likely to endorse contemporary descent (extant taxa are descended from other extant taxa) when interpreting bracket phylogenetic trees. Finally, there was no difference in accuracy between styles for evaluating taxa relatedness. However, we found that students were largely unable to determine taxa relatedness, regardless of phylogenetic tree style. For some insight into why performance was poor for taxa relatedness, we tabulated the specific forms of reasoning used by students (Table 8). Counting the number of synapomorphies between taxa was by far the most common approach, followed by counting nodes and comparing distances between taxa (branch tip proximity). Other forms of reasoning were relatively rare, and students generally used the same approach for diagonal and bracket phylogenetic trees.

Table 7. Number of correct student responses to each interpretation task for diagonal and bracket phylogenetic trees (percentages in parentheses; n=92 paired responses).

Interpretation Task	Diagonal Style	Bracket Style	Comparison
Trait Possession	83 (90%)	81 (88%)	p=0.791
Most Recent Common Ancestor	75 (82%)	91 (99%)	p<0.001
Monophyletic Group	63 (68%)	77 (84%)	p=0.003
Contemporary Descent Answer	65 (71%)	82 (89%)	p<0.001
Contemporary Descent Reasoning	55 (60%)	74 (80%)	p<0.001
Taxa Relatedness Answer	12 (13%)	15 (16%)	p=0.250
Taxa Relatedness Reasoning	7 (8%)#	10 (11%)#	p=0.368*

#Mixed reasoning was found in less than 5% of student responses for each style of phylogenetic tree. *P-value was calculated using a Stuart-Maxwell test due to trichotomous categories (correct, incorrect, or mixed reasoning). All other p-values were derived from exact McNemar's tests.

Table 8. Number of student responses that contained specific forms of reasoning to evaluate taxa relatedness for each phylogenetic tree style (percentages in parentheses; n=92 paired responses).

Taxa Relatedness Reasoning	Diagonal Style	Bracket Style
Most Recent Common Ancestry	6 (7%)	7 (8%)
Monophyletic Grouping	5 (5%)	6 (7%)
Counting Nodes	20 (22%)	20 (22%)
Counting Synapomorphies	46 (50%)	41 (45%)
Branch Tip Proximity	19 (21%)	16 (17%)
External Insights	5 (5%)	2 (2%)
Other Responses	8 (9%)	6 (7%)

Student responses could include multiple forms of reasoning for taxa relatedness. See Dees et al. (2014) for complete descriptions and student-generated examples of reasoning categories.

Construction

Students showed a preference for building diagonal phylogenetic trees. Of the 85 students who completed the construction task, 59 (69%) used the diagonal style (p<0.001 versus an equal number of each style). The majority of phylogenetic trees constructed by students were correct or adequate in terms of accuracy (Figure 7). However, there was a significant difference in accuracy between diagonal and bracket phylogenetic trees (p=0.002). This style effect disappeared entirely when adequate phylogenetic trees were treated as correct (p=1.00), indicating the difference was driven by the adequate category. Specifically, diagonal phylogenetic trees included considerably more minor errors, but the incidence of major errors was similar between styles (Table 9).

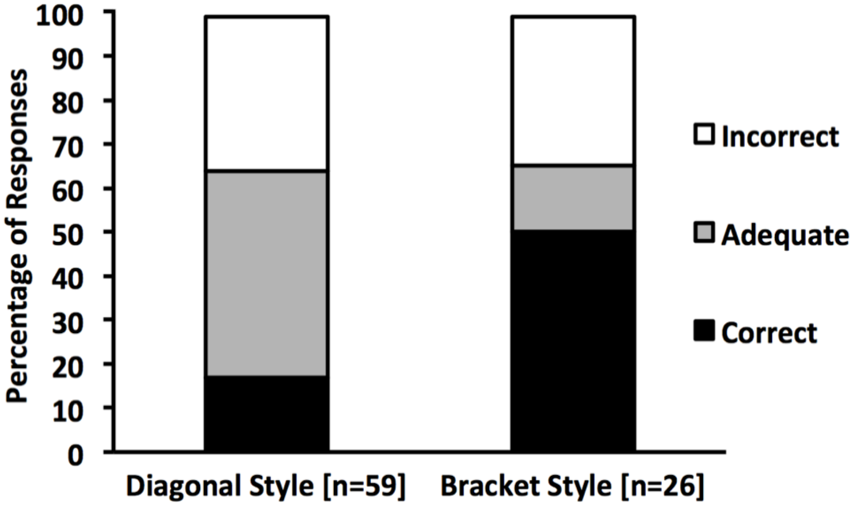


Figure 7. Accuracy of phylogenetic trees constructed by students (n=85 responses).

Table 9. Major and minor errors in phylogenetic trees constructed by students (n=85 responses).

Major Error	**Diagonal Style (n=59)**	**Bracket Style (n=26)**
Incorrect Relatedness	16 (27%)	6 (23%)
Incorrect Traits	19 (32%)	6 (23%)
Contemporary Descent	2 (3%)	2 (8%)
Minor Error	**Diagonal Style (n=59)**	**Bracket Style (n=26)**
Empty Branches	30 (51%)	3 (12%)
Extra Nodes	44 (75%)	6 (23%)
Side Branches	0 (0%)	2 (8%)

Phylogenetic trees constructed by students could include multiple major and minor errors. See Dees and Momsen (2016) for complete descriptions and student-generated examples of errors.

Discussion

We expanded on existing studies that assessed effects of style on student understanding of phylogenetic trees in two ways. First, rather than using surveys in which students had no stake in the outcome, we collected data through a pre-instructional homework in an introductory biology course for science and related majors. Second, we included a phylogenetic tree construction task in addition to a series of interpretation tasks. Our results inform researchers and teachers who are working to develop phylogenetic tree instruction that best facilitates student learning.

Interpretations

Using surveys, Novick and Catley (2013) determined that students performed significantly better with bracket phylogenetic trees across interpretation tasks. We collected data as part of an introductory biology course and used a different instrument, but our results largely agree with the previous investigation. For most interpretation tasks, students performed significantly better with bracket phylogenetic trees. However, there was no difference between styles for identifying traits possessed by taxa and evaluating taxa relatedness. For trait possession, students could answer the questions simply by reading information that was explicitly provided by the diagonal and bracket phylogenetic trees (Figure 6). In contrast, the other interpretation tasks required students to either use symbolic information provided by the phylogenetic trees (e.g., nodes and branches) or apply external knowledge (e.g., monophyletic groups and taxa relatedness). Thus, it was not surprising that students were generally able to identify traits possessed by taxa, regardless of style. For taxa relatedness, the lack of a significant difference between styles was probably due to a floor effect. The vast majority of students were unable to evaluate taxa relatedness, due in large part to using incorrect reasoning strategies, which resulted in style having no impact. Further, the poor overall performance for determining taxa relatedness aligns with previous research (Dees et al., 2014).

Construction

As part of an earlier study, Dees and Momsen (2016) found that accuracy of phylogenetic trees constructed by students was independent of style. However, few students chose to construct bracket phylogenetic trees, which resulted in an inconclusive outcome. During the present study, students again favored the diagonal style for construction, but enough students chose the bracket style to enable a meaningful comparison. Although students performed well overall, construction accuracy differed significantly by style. Specifically, there was no difference in major errors, but diagonal phylogenetic trees included considerably more minor construction errors, such as empty branches and extra nodes. While these minor errors should not affect student performance on the interpretation tasks we used for this study (e.g., trait possession and taxa relatedness), such errors could impact other interpretations. For example, empty branches could reflect the common belief that evolutionary changes occurred only at nodes (Baum et al., 2005; Meir et al., 2007; Gregory, 2008). Thus, minor errors are not necessarily inconsequential, and the prevalence of minor errors in diagonal phylogenetic trees constructed by introductory biology students is concerning.

Limitations and Future Research

We recognize this study is limited in a number of ways. First, for the single construction task, students were asked to use the style of their choice. Therefore, each student constructed one phylogenetic tree, and the accuracy comparison across styles is between-student. Future studies should use a stronger within-student approach by asking students to construct equivalent diagonal and bracket phylogenetic trees during separate tasks. Second, the order of assessment items was the same for each student: 1) diagonal phylogenetic tree with associated questions, 2) bracket phylogenetic tree with associated questions, and 3) construction task. Thus, it is possible prompt order impacted student responses. For example, some students could have performed better on questions associated with the bracket phylogenetic tree due to experience gained by attempting to interpret the diagonal phylogenetic tree. However, students also could have performed worse on questions associated with the bracket phylogenetic tree due to assessment fatigue or loss of motivation over time. Future research should control for possible order effects by systematically varying the order of assessment items. Third, we collected data from one introductory biology course at one public university, and the results may not reflect undergraduate biology students as a whole. Students at other institutions may have different academic backgrounds and motivations that could influence their performance. Thus, there is a need for further research at a variety of schools to accumulate more evidence from which we can make more robust and generalizable claims. Finally, our use of terminology may vary from that used by systematists. However, we used language consistent with undergraduate biology educators and introductory textbooks (e.g., Baum and Offner, 2008; Freeman, 2011; Novick and Catley, 2013).

Implications for Instruction

Most current introductory textbooks, in response to research similar to this study (Novick and Catley, 2007, 2013), use only bracket phylogenetic trees. However, many instructors may be unaware of this publishing decision, and many more are most likely unfamiliar with the empirical research related to style effects and phylogenetic trees. A typical introductory biology instructor uses a variety of resources to develop curricula, including but not limited to the textbook. As a result, instruction may use diagonal and bracket phylogenetic trees somewhat interchangeably. Thus, while introductory biology textbooks are consistent in their use of bracket phylogenetic trees, instruction often includes a mixture. Research on style effects, therefore, serves to inform instructors of the potential for phylogenetic tree style to impact student reasoning.

Further, research focused on the interactions of instruction with style effects is currently quite limited. For example, data for this study were collected before instruction on phylogenetic trees, and prior research on style effects either did not document instruction (Novick and Catley, 2007, 2013) or resulted in an inconclusive outcome (Dees and Momsen, 2016). Thus, the direct impact of instruction remains unknown, and we advise against making instructional decisions regarding phylogenetic tree style without additional data. It is possible, for example, that using diagonal and bracket phylogenetic trees equally during instruction could reduce or eliminate performance differences between styles. Future investigations should collect data before and after instruction on phylogenetic trees that is controlled for style. Such research is needed to determine the best course of action for teachers and researchers who are working to develop phylogenetic tree instruction that best facilitates student learning.

STUDY 3. FURTHER EFFECTS OF PHYLOGENETIC TREE STYLE ON STUDENT COMPREHENSION IN AN INTRODUCTORY BIOLOGY COURSE[3]

Abstract

Phylogenetic trees have become increasingly important across the life sciences, and as a result, learning to interpret and reason from these diagrams is an essential component of biology education. Unfortunately, students struggle to understand phylogenetic trees, even after explicit instruction. Style (i.e., diagonal or bracket) is one factor that has been observed to impact how students interpret phylogenetic trees, and the goal of this research was to investigate these style effects across an introductory biology course that integrated diagonal and bracket phylogenetic trees equally. Before instruction, students were significantly more accurate with the bracket style for a variety of interpretation and construction tasks. Instruction mitigated some, but not all, style effects. Specifically, after instruction, students were significantly more accurate for construction tasks and interpretations involving taxa relatedness when using the bracket style. Based on this study and previous research on style effects, we recommend introductory biology instructors use only the bracket style for instruction. Future studies should investigate the causes of style effects and variables other than style to inform the development of research-based instruction.

Introduction

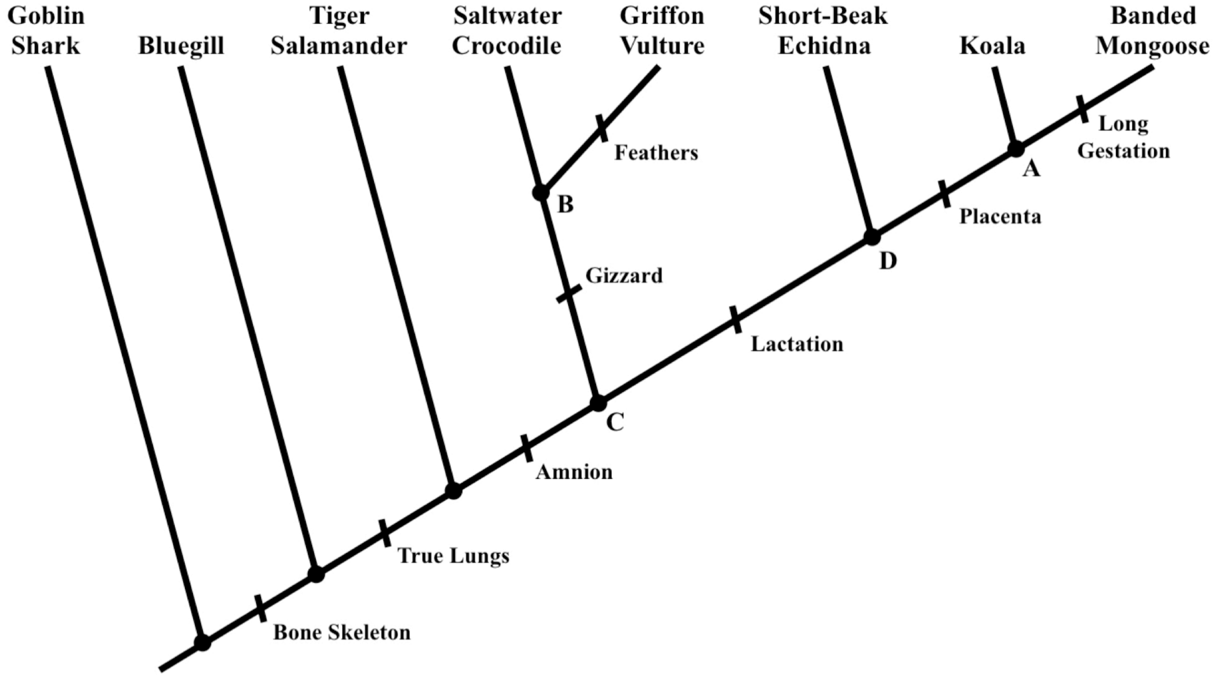
Phylogenetic trees are tools that facilitate reasoning about biological phenomena from an evolutionary perspective ("tree thinking"; O'Hara, 1988; Gregory, 2008). Although phylogenetic trees are often viewed simply as visual representations of hypothesized evolutionary relationships

---

[3] The material contained in this chapter was co-authored by Jonathan Dees, Caitlin Bussard, and Dr. Jennifer Momsen and is in review as Dees et al. at CBE–Life Sciences Education. Jonathan Dees developed the assessment items, contributed to data collection and analyses, drafted this chapter, and contributed to revisions. Caitlin Bussard contributed to data analyses. Dr. Jennifer Momsen contributed to data collection and revisions of this chapter.

among taxa, these diagrams are also the main analytical tool used by biologists to assess evidence of evolution (Baum et al., 2005; Novick and Catley, 2007). Further, phylogenetic trees provide an efficient framework to organize our growing knowledge of biological diversity (Thanukos, 2009; Wiley, 2010; Baum and Smith, 2013). Because evolution is a unifying theory and core concept in biology (Dobzhansky, 1973; American Association for the Advancement of Science, 2011; Next Generation Science Standards Lead States, 2013; College Board, 2015), and due to advancements in phylogenetic inference and DNA sequencing technologies (Omland et al., 2008), phylogenetic trees have become increasingly important across the life sciences (Baum and Offner, 2008). As a result, learning to interpret and reason from phylogenetic trees is now an essential component of biology education (O'Hara, 1997; Lents et al., 2010; Meisel, 2010; Novick and Catley, 2016).

Despite the significance of phylogenetic trees, students at all levels routinely struggle to interpret them (Meir et al., 2007; Halverson et al., 2011; Catley et al., 2013; Novick and Catley, 2013; Blacquiere and Hoese, 2016), even after explicit instruction (Phillips et al., 2012; Smith et al., 2013; Dees et al., 2014). Student difficulties with phylogenetic trees have been attributed to a number of factors, starting with abstractness. As a type of schematic diagram, phylogenetic trees present abstract information that requires learned rules and conventions for correct interpretation (Novick and Catley, 2007). In other words, understanding phylogenetic trees is not intuitive, and students must be taught how to extract information from these diagrams (Sandvik, 2008; Eddy et al., 2013). The Gestalt perceptual principles of good continuation and spatial proximity have also been shown to negatively impact students, especially for phylogenetic trees drawn in the diagonal style (Figure 8) and when interpreting taxa relatedness (Novick and Catley, 2007, 2013). Finally, student interpretations of phylogenetic trees and student conceptions of evolution are interrelated, such that each affects the other (Gregory, 2008; Omland et al., 2008). Thus, misinterpretations of phylogenetic trees impede student understanding of evolution (Meir et al., 2007), and conversely, misconceptions about evolution also lead to student difficulties with phylogenetic trees.



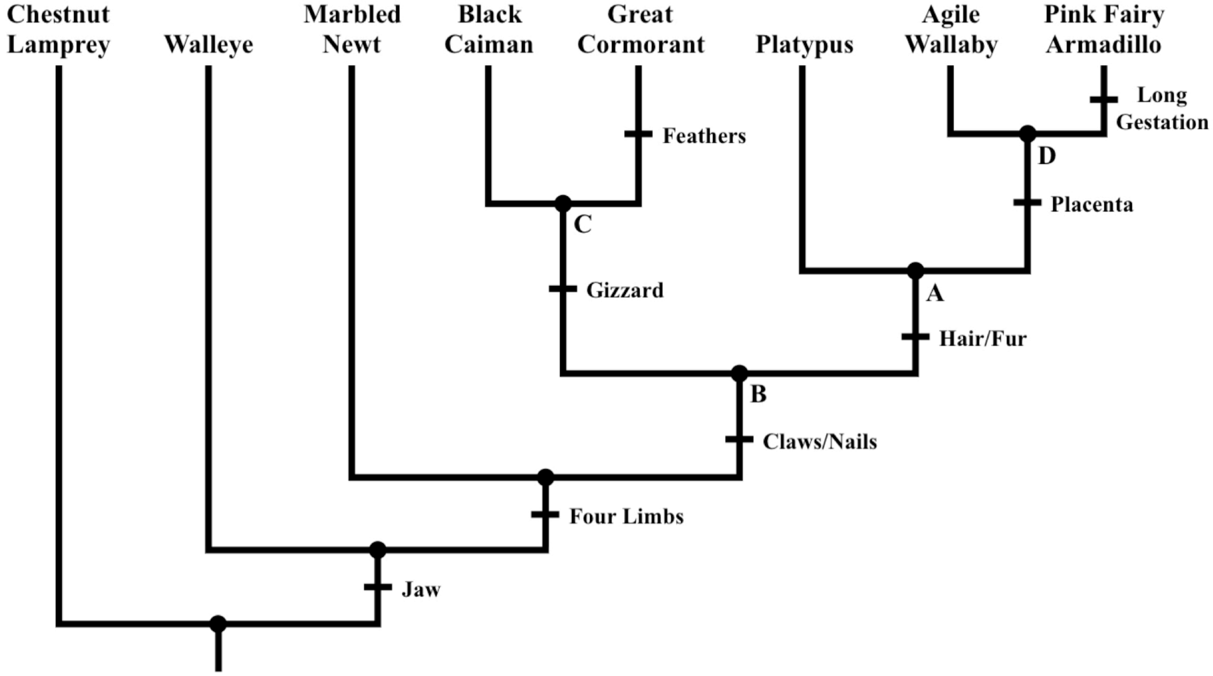


Figure 8. Equivalent diagonal (top) and bracket (bottom) phylogenetic trees, which are the same size and have the same branch pattern but involve different taxa and traits.

One factor that is controllable by instructors and has been observed to affect how students interpret phylogenetic trees is style (Baum and Offner, 2008; Halverson et al., 2011). Two styles of phylogenetic tree that contain equivalent information, diagonal and bracket, commonly appear in textbooks, journals, and other resources (Figure 8; Catley and Novick, 2008). However, to our knowledge, only three studies have explicitly examined effects of style on student understanding of phylogenetic trees. In an initial study, Novick and Catley (2007) used translation tasks to find differences in how students perceive diagonal and bracket phylogenetic trees. Students were asked to convert visual representations of evolution, including diagonal and bracket phylogenetic trees, from one representation to another while maintaining the same evolutionary relationships among taxa. Accuracy was significantly lower for translations involving diagonal phylogenetic trees, and this style effect was more pronounced for students with less experience in biology.

In a later study, Novick and Catley (2013) used interpretation tasks to further investigate effects of style on student understanding of phylogenetic trees. For example, students were asked to evaluate taxa relatedness, recognize monophyletic and non-monophyletic groups, and identify traits shared by taxa due to common ancestry. Across nearly all tasks, accuracy was significantly lower when students interpreted diagonal phylogenetic trees, and this style effect was often found regardless of background in biology. Finally, Dees et al. (accepted) examined effects of style on student interpretations and construction of phylogenetic trees by collecting data in the context of an introductory biology course. Prior to instruction on phylogenetic trees, students were asked to complete numerous interpretation tasks for both styles that were similar to those used by Novick and Catley (2013). Students also constructed a phylogenetic tree in the style of their choice from provided data. For most interpretation tasks, accuracy was again significantly lower for diagonal phylogenetic trees. Students who constructed diagonal phylogenetic trees were also significantly less accurate compared to those who used the bracket style for the construction task.

Although three studies provided multiple lines of evidence indicating students had more difficulties with diagonal phylogenetic trees compared to the bracket style, each investigation had important limitations. Novick and Catley (2007, 2013) used surveys to collect data from students who were mostly recruited as volunteers from psychology, education, and biology courses. From a motivational perspective, students may not take surveys as seriously as coursework that affects their academic standing (Sundberg, 2002). In addition, neither study included construction tasks, which are common instructional activities for phylogenetic trees (e.g., Gendron, 2000; Goldsmith, 2003; Julius and Schoenfuss, 2006; Burks and Boles, 2007; Lents et al., 2010; Eddy et al., 2013; Bokor et al., 2014; Lampert and Mook, 2015). Dees et al. (accepted) addressed these limitations by obtaining data through coursework and by examining both interpretations and construction of phylogenetic trees. However, data were only collected from introductory biology students before instruction on phylogenetic trees. Further, students were asked to construct one phylogenetic tree in the style of their choice, resulting in a between-student comparison of construction accuracy for diagonal and bracket phylogenetic trees rather than a stronger within-student comparison.

As described, Novick and Catley (2007, 2013) and Dees et al. (accepted) showed that style can impact how students interpret and construct phylogenetic trees. However, additional research is necessary to address the limitations of these investigations and determine if style effects persist after instruction that is not biased toward diagonal or bracket phylogenetic trees. Thus, the goal of the present study was to further explore style effects by gathering data that satisfied the following criteria: (1) obtained through coursework in biology, (2) included interpretations and construction of phylogenetic trees, (3) supported within-student comparisons of performance across styles, and

(4) collected before, after, and long after unbiased instruction that integrated diagonal and bracket phylogenetic trees equally. These data allowed us to address the following research questions:

1) Do introductory biology students have differential interpretation abilities for diagonal and bracket phylogenetic trees before and after instruction?
2) Do introductory biology students have differential construction abilities for diagonal and bracket phylogenetic trees before and after instruction?

Methods

Data for this study were collected during an introductory biology course for science and related majors at a large, public university in the midwestern United States. No prerequisites were required for enrollment, and the course served students (n=83) at various stages in their academic programs (30% freshmen, 41% sophomores, 18% juniors, and 11% seniors). Content started with inheritance (weeks 1-3) and progressed through evolution and biodiversity (weeks 4-8), form and function of plants and animals (weeks 9-12), and ecology (weeks 13-15). At the end of the course, students participated in review activities during class (week 16) to prepare for the comprehensive final exam (week 17). Instruction was learner-centered and emphasized collaboration (Johnson et al., 1998; Tanner et al., 2003; Armstrong et al., 2007) by having assigned groups of three or four students build and evaluate conceptual models (Dauer et al., 2013; Bray Speth et al., 2014; Long et al., 2014), discuss clicker questions (Caldwell, 2007; Freeman et al., 2007; Perez et al., 2010), and construct scientific arguments (Driver et al., 2000). Classes were observed, and instructional materials and assessments were collected to document instruction throughout the course.

Instrument Design

We developed four instruments to measure effects of phylogenetic tree style on student comprehension before (pre-instructional homework), after (post-instructional homework and unit exam), and long after instruction (review activity for the final exam). Each instrument contained a diagonal phylogenetic tree and an equivalent bracket phylogenetic tree, which were the same size and had the same branch pattern but involved different taxa and traits (e.g., Figure 8). Isomorphic interpretation tasks accompanied each diagonal and bracket phylogenetic tree such that accuracy could be compared across styles. These interpretation tasks were modified from a previous study (Dees et al., accepted) and based largely on the essential tree-thinking skills proposed by Novick and Catley (2013). Specifically, students were asked to identify the most recent common ancestor of taxa, recognize monophyletic groups, determine if extant taxa are descended from other extant taxa ("contemporary descent"; Dees et al., 2014), and evaluate taxa relatedness.

Students were also asked to construct phylogenetic trees from data, either in a specified style or in the style of their choice. The instruments that were assigned as homework included two construction tasks, one for each style, which resulted in equivalent phylogenetic trees. Because the unit exam and review activity for the final exam were completed during class and subject to time constraints, these instruments contained a single construction task that allowed students to use the style of their choice. To reduce context effects, in which student reasoning about evolution varies for different taxa and traits (Nehm and Ha, 2011), phylogenetic trees used for interpretation tasks exclusively involved animals (e.g., Figure 8) while all construction tasks involved plants. The four instruments used for this investigation are available in Appendix C.

Data Collection

Before instruction on phylogenetic trees during the seventh week, students were asked to watch a screencast that was posted on the course management system. The screencast introduced phylogenetic trees by describing their purpose and defining essential terminology (e.g., nodes and monophyletic groups), without explaining how to interpret or construct the diagrams. A diagonal phylogenetic tree and an equivalent bracket phylogenetic tree were shown side-by-side during the screencast, such that one style was not introduced first or favored over the other style. To further prevent style bias, students were never asked to read about phylogenetic trees in the textbook for the course (Urry et al., 2014), which only used the bracket style. After the screencast was posted, each student was randomly assigned either the diagonal or bracket section of the pre-instructional homework. Upon submission, students were assigned the opposite section of the homework. This distribution method was used to control for order effects, in which student responses are impacted by the sequence of assessment items (Halverson et al., 2013; Federer et al., 2015).

Instruction on phylogenetic trees began after both the diagonal and bracket sections of the pre-instructional homework were submitted by students. Similar to the earlier screencast, students were shown a pair of equivalent diagonal and bracket phylogenetic trees in a side-by-side manner during initial instruction. Subsequent instructional activities for interpretations involved one style or the other, but overall, an equal number of diagonal and bracket phylogenetic trees were used by the instructor. When instructional activities included construction tasks, students were allowed to use the style of their choice. Verification feedback (i.e., labeling responses as correct or incorrect; Marsh et al., 2012) was provided for the pre-instructional homework and submitted instructional activities. The post-instructional homework was distributed to students using the same method as the pre-instructional homework, and verification feedback was provided before the unit exam.

One week after instruction on phylogenetic trees, students completed an exam in class that assessed understanding of speciation, biodiversity, and phylogenetic trees. The section of the unit exam devoted to phylogenetic trees was structured the same as the instruments that were deployed as homework, except only one construction task was included due to time constraints. To control for order effects, each student received one of two versions of the unit exam, which varied only in the sequence of assessment items. A diagonal phylogenetic tree and associated interpretation tasks preceded an equivalent bracket phylogenetic tree and associated interpretation tasks in version A, while the order was reversed in version B. The single construction task, which allowed students to use the style of their choice, appeared after the two sets of interpretation tasks in both versions of the unit exam. Answer feedback (i.e., providing correct answers without explaining why answers are correct or incorrect; Marsh et al., 2012) was given to students one week after the unit exam in the form of a grading rubric that was posted on the course management system.

Finally, during the last week and eight weeks after the unit exam, students participated in various review activities to prepare for the comprehensive final exam. To investigate style effects long after instruction on phylogenetic trees, data had to be collected without students preparing in advance. Thus, the last instrument was deployed as one of the review activities rather than as part of the final exam. The instrument was structured the same as the section of the unit exam that was devoted to phylogenetic trees. Two versions of the instrument that varied only in the sequence of assessment items were also created and distributed in the same manner as the unit exam to control for order effects. Students completed the review activity during class without access to resources, which concluded data collection for this investigation. Although phylogenetic trees also appeared on the final exam, the associated assessment items were not designed for this study.

Data Coding

Responses to interpretation and construction tasks were coded using the methods outlined in an earlier investigation (Dees et al., accepted). Tasks that involved identifying the most recent common ancestor of taxa required a multiple-choice answer, and responses were coded as correct or incorrect. Tasks that involved recognizing a monophyletic group had multiple correct answers, and responses were again coded as correct or incorrect. Tasks that involved determining if extant taxa are descended from other extant taxa ("contemporary descent"; Dees et al., 2014) required a yes or no answer with reasoning. Answers and reasoning were each coded as correct or incorrect, where correct reasoning stated or implied that extant taxa evolved from a common ancestor rather than each other. Tasks that involved evaluating taxa relatedness required a multiple-choice answer with reasoning. Answers were coded as correct or incorrect, while a published rubric was used to code student reasoning as correct, incorrect, or mixed (Dees et al., 2014). Correct reasoning cited most recent common ancestry or monophyletic groups as criteria for determining taxa relatedness, while incorrect reasoning typically referred to the number of nodes or traits between taxa, relative distance between taxa, or information that was not provided by phylogenetic trees. Students often included multiple forms of reasoning in their responses, and in some cases, used mixed reasoning that contained both correct and incorrect criteria for evaluating taxa relatedness.

Responses to construction tasks were coded for accuracy as correct, adequate, or incorrect using a published rubric (Dees and Momsen, 2016). Phylogenetic trees that included one or more major errors, such as incorrect relatedness and incorrect traits, were considered incorrect. Student responses that included only minor errors, such as extra nodes and empty branches, were coded as adequate. Major and minor errors were differentiated based on whether or not the errors impeded students from interpreting taxa relatedness or traits possessed by taxa. Finally, phylogenetic trees with no major or minor errors were considered correct. All student responses to interpretation and construction tasks that were collected for this investigation were coded by two independent raters with greater than 94% agreement (kappa coefficient greater than 0.86; Cohen, 1960).

Statistical Analyses

For each instrument, we analyzed responses to isomorphic interpretation tasks associated with equivalent diagonal and bracket phylogenetic trees as paired, categorical data. In the case of dichotomous categories (e.g., correct or incorrect), we used an exact version of the McNemar test, which is suitable for small sample sizes, accounts for the paired nature of our data, and generates within-student comparisons of performance across styles (McNemar, 1947; Rufibach, 2011). For style effects, the null hypothesis of the McNemar test is that an equal number of students switched categories in one direction (e.g., incorrect to correct) as in the opposite direction from one style of phylogenetic tree to the other style (McDonald, 2014). In the case of trichotomous categories (e.g., correct, incorrect, or mixed), we used the Stuart-Maxwell extension of the McNemar test (Stuart, 1955; Maxwell, 1970; Sun and Yang, 2008). Order effects within each instrument and changes in student performance between instruments were investigated using the same statistics but different variables of interest (e.g., instrument as the variable rather than style of phylogenetic tree).

Responses to construction tasks were analyzed in the same manner as interpretation tasks, with the exception of data from the unit exam and review activity for the final exam. Due to time constraints, these two instruments included one construction task that allowed students to use the style of their choice rather than a construction task for each style of phylogenetic tree. Therefore, student responses to these construction tasks had to be analyzed as unpaired, categorical data. We used the Fisher exact test, which is suitable for small sample sizes and generates between-student comparisons of performance across styles. In this situation, the null hypothesis of the Fisher exact test is that accuracy was independent of the style used by students to construct phylogenetic trees (Fisher, 1934). Finally, we used the exact binomial test to determine if students chose either style significantly more than the other style for construction tasks on the unit exam and review activity for the final exam. For this scenario, the null hypothesis of the exact binomial test is that students constructed an equal number of diagonal and bracket phylogenetic trees (McDonald, 2014).

Results

Data were collected by a pre-instructional homework (n=74), post-instructional homework (n=75), unit exam (n=81), and review activity for the final exam (n=72). Some students elected not to submit their homework or attend class when the review activity was completed, resulting in smaller sample sizes compared to the unit exam. In addition, two students withdrew from the course (n=83) after the pre-instructional homework and before the unit exam. No order effects were observed for any task on any instrument (i.e., whether students received tasks for diagonal or bracket phylogenetic trees first did not significantly impact accuracy; all $p>0.26$). Accuracy increased significantly from the pre-instructional homework to the post-instructional homework for all interpretation and construction tasks across both styles (all $p<0.04$). Further, accuracy did not change significantly from the post-instructional homework to the unit exam and final exam review activity for any interpretation or construction task across both styles (all $p>0.12$).

Interpretations

Students were significantly more accurate when interpreting bracket phylogenetic trees on the pre-instructional homework for three tasks: identifying the most recent common ancestor of taxa, recognizing monophyletic groups, and determining if extant taxa are descended from other extant taxa ("contemporary descent"; Table 10). These significant differences in accuracy disappeared after instruction that used diagonal and bracket phylogenetic trees equally and did not reemerge during the unit exam or final exam review activity. For interpretations concerning contemporary descent, students were asked to provide reasoning for their answers. Although students' answers were consistently more accurate than their reasoning, the patterns of answers and reasoning were similar when comparing diagonal and bracket phylogenetic trees across all four instruments that were used to collect data for this investigation.

Table 10. Percentage of correct student responses for all interpretation tasks and instruments with comparisons of accuracy across phylogenetic tree styles.

Most Recent Common Ancestor				
Style	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=81]**	**Final Review [n=72]**
Diagonal	73%	95%	93%	92%
Bracket	86%	95%	98%	94%
Comparison	p=0.02	p=1.00	p=0.22	p=0.75
Monophyletic Group				
Style	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=81]**	**Final Review [n=72]**
Diagonal	54%	88%	93%	92%
Bracket	68%	91%	96%	93%
Comparison	p=0.04	p=0.69	p=0.38	p=1.00
Contemporary Descent: Answer				
Style	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=81]**	**Final Review [n=72]**
Diagonal	73%	97%	95%	94%
Bracket	89%	100%	99%	97%
Comparison	p<0.01	p=0.50	p=0.38	p=0.50
Contemporary Descent: Reasoning				
Style	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=81]**	**Final Review [n=72]**
Diagonal	53%	81%	75%	74%
Bracket	72%	85%	80%	78%
Comparison	p<0.01	p=0.51	p=0.39	p=0.55
Taxa Relatedness: Answer				
Style	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=81]**	**Final Review [n=72]**
Diagonal	11%	39%	49%	46%
Bracket	15%	55%	59%	60%
Comparison	p=0.58	p=0.02	p=0.04	p<0.01
Taxa Relatedness: Reasoning				
Style	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=81]**	**Final Review [n=72]**
Diagonal	5%[#]	36%[#]	42%[#]	40%[#]
Bracket	8%[#]	53%[#]	58%[#]	50%[#]
Comparison	p=0.55*	p=0.02*	p<0.01*	p=0.03*

[#] Mixed reasoning was also found in <10% of student responses. *P-values were derived from a Stuart-Maxwell test due to trichotomous categories (correct, incorrect, or mixed reasoning). All other p-values were derived from an exact version of the McNemar test.

In contrast, there was no significant difference in accuracy for evaluating taxa relatedness on the pre-instructional homework, although accuracy was very low for both diagonal and bracket phylogenetic trees (Table 10). However, students were significantly more accurate when assessing taxa relatedness on bracket phylogenetic trees following instruction, and this difference persisted through the unit exam and final exam review activity. Interpretations concerning taxa relatedness required students to provide reasoning for their answers, and the patterns of student answers and reasoning were similar across all four instruments. Specific forms of reasoning are in Table 11.

Table 11. Percentage of student responses that contained specific forms of reasoning to evaluate taxa relatedness for each instrument and style of phylogenetic tree.

Diagonal Phylogenetic Trees				
Taxa Relatedness Reasoning	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=81]**	**Final Review [n=72]**
Most Recent Common Ancestry	*14%*	*41%*	*51%*	*47%*
Monophyletic Grouping	*1%*	*0%*	*0%*	*0%*
Counting Nodes	23%	28%	31%	32%
Counting Synapomorphies	35%	16%	10%	19%
Branch Tip Proximity	36%	13%	4%	1%
External Insights	5%	5%	2%	1%
Other Responses	8%	11%	14%	11%
Bracket Phylogenetic Trees				
Taxa Relatedness Reasoning	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=81]**	**Final Review [n=72]**
Most Recent Common Ancestry	*11%*	*53%*	*64%*	*57%*
Monophyletic Grouping	*4%*	*4%*	*1%*	*3%*
Counting Nodes	28%	23%	23%	24%
Counting Synapomorphies	41%	13%	7%	19%
Branch Tip Proximity	22%	9%	5%	1%
External Insights	7%	3%	2%	1%
Other Responses	12%	5%	6%	7%

Italics indicate correct forms of reasoning for taxa relatedness. Student responses could include multiple forms of reasoning (percentages sum to greater than 100%). See Dees et al. (2014) for complete descriptions and student-generated examples of reasoning categories.

Construction

Across instruments, before and after classroom instruction, students were significantly more accurate when constructing bracket phylogenetic trees (Figure 9). However, this difference in accuracy disappeared for each instrument when adequate phylogenetic trees were considered correct, indicating the adequate category was responsible for the discrepancy between styles. In other words, diagonal phylogenetic trees contained far more minor errors, but the occurrence of major errors was similar across styles. Specific major and minor errors found in phylogenetic trees constructed by students are available in Table 12. Note that students constructed a single phylogenetic tree in the style of their choice during the unit exam and final exam review activity, and students overwhelmingly chose to use the diagonal style for both instruments (79% and 78% diagonal phylogenetic trees, respectively; $p<0.001$ versus an equal distribution by style).

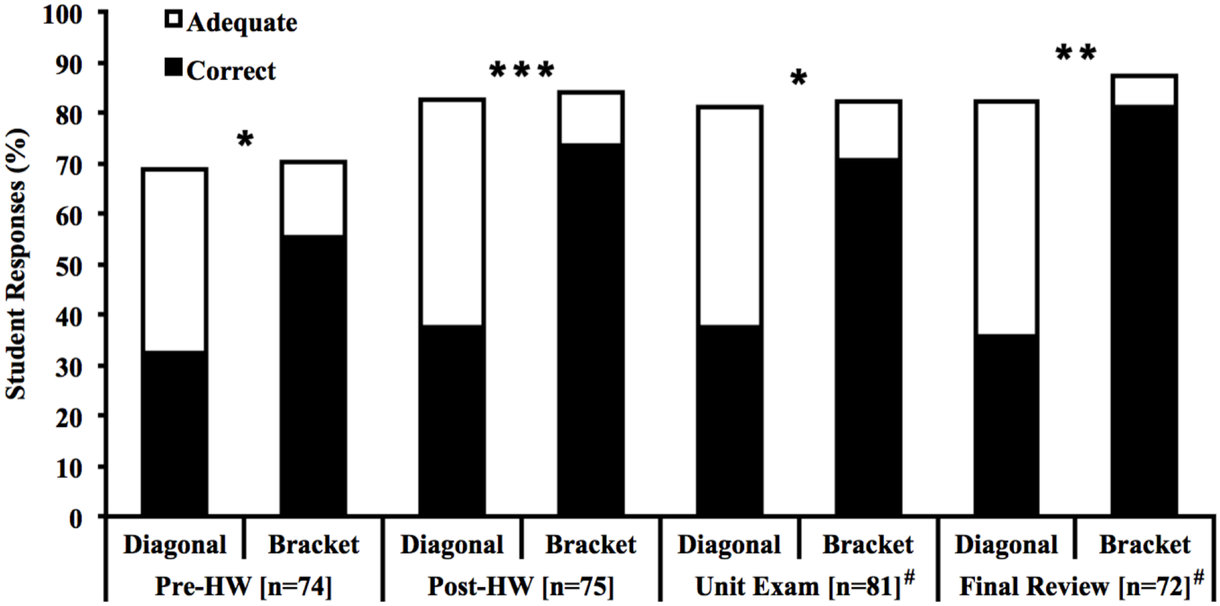


Figure 9. Accuracy of phylogenetic trees constructed by students with comparisons across styles for all instruments. #Students constructed one phylogenetic tree in the style of their choice during the unit exam (64 diagonal, 17 bracket) and final exam review activity (56 diagonal, 16 bracket), resulting in between-student rather than within-student comparisons of accuracy across styles for those instruments. *$p<0.05$, **$p<0.01$, and ***$p<0.001$.

Table 12. Percentage of phylogenetic trees constructed by students that contained specific errors for each instrument and style of phylogenetic tree.

Diagonal Phylogenetic Trees				
Major Errors	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=64][#]**	**Final Review [n=56][#]**
Incorrect Relatedness	22%	13%	11%	11%
Incorrect Traits	27%	16%	19%	18%
Contemporary Descent	3%	4%	2%	0%
Minor Errors	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=64][#]**	**Final Review [n=56][#]**
Empty Branches	42%	40%	39%	39%
Extra Nodes	61%	52%	52%	52%
Bracket Phylogenetic Trees				
Major Errors	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=17][#]**	**Final Review [n=16][#]**
Incorrect Relatedness	22%	13%	0%	6%
Incorrect Traits	24%	15%	12%	13%
Contemporary Descent	3%	3%	6%	6%
Minor Errors	**Pre-HW [n=74]**	**Post-HW [n=75]**	**Unit Exam [n=17][#]**	**Final Review [n=16][#]**
Empty Branches	9%	4%	12%	13%
Extra Nodes	19%	8%	12%	6%
Side Branches*	16%	11%	18%	13%

[#] Students constructed one phylogenetic tree in the style of their choice during the unit exam and review activity for the final exam. *Error is unique to the bracket style. Student responses could include any combination of errors or no errors (percentages do not sum to 100%). See Dees and Momsen (2016) for complete descriptions and student-generated examples of errors.

Discussion

Building from prior studies, we examined effects of style on student interpretations and construction of phylogenetic trees in the context of an introductory biology course for science and related majors. In contrast to prior research, this study supported within-student comparisons of performance across styles and included data collected before, after, and long after unbiased instruction that integrated diagonal and bracket phylogenetic trees equally. Our results indicate such instruction eliminated some, but not all, style effects that favored the bracket style, which suggests diagonal phylogenetic trees may not be suitable for introductory-level biology courses.

Interpretations

Prior to classroom instruction on phylogenetic trees, students were significantly more accurate with the bracket style for most interpretation tasks, including identifying the most recent common ancestor of taxa, recognizing monophyletic groups, and determining if extant taxa are descended from other extant taxa ("contemporary descent"). These differences in accuracy were mitigated by instruction that balanced the use of diagonal and bracket phylogenetic trees and did not reemerge during the course, which demonstrates that some style effects for interpretations were responsive to instruction that integrated diagonal and bracket phylogenetic trees equally.

In contrast, interpretations of taxa relatedness on phylogenetic trees exhibited a different pattern. Prior to instruction, there was no significant difference in accuracy across styles due to a floor effect. The vast majority of students simply did not know how to evaluate taxa relatedness, and thus, the style of phylogenetic tree did not impact student responses. Following instruction, however, students were significantly more accurate when evaluating taxa relatedness on bracket phylogenetic trees across all three post-instructional instruments. This difference included both answers and reasoning, as students used somewhat different forms of reasoning for each style of phylogenetic tree (Table 11). Note that accuracy for evaluating taxa relatedness was quite low for both styles even after instruction, which aligns with previous studies on student understanding of taxa relatedness (Phillips et al., 2012; Smith et al., 2013; Dees et al., 2014). Thus, instruction that balanced the use of diagonal and bracket phylogenetic trees was not beneficial for style effects in regards to evaluating the relatedness of taxa, and our students were typical in their struggles with evaluating taxa relatedness on phylogenetic trees in general.

Construction

The majority of diagonal and bracket phylogenetic trees created by students were correct or adequate in terms of accuracy across all four instruments. However, bracket phylogenetic trees were also significantly more accurate than diagonal phylogenetic trees across instruments due to a lower incidence of minor errors (e.g., extra nodes and empty branches; Table 12). Although the minor errors observed in student-constructed phylogenetic trees should not hinder performance on our interpretation tasks, such mistakes could be indicative of other misinterpretations. For example, extra nodes and empty branches may reflect the common belief among students that evolutionary changes occurred only at nodes (Baum et al., 2005; Meir et al., 2007; Gregory, 2008). Thus, in some cases, students may have intentionally included more minor errors when constructing diagonal phylogenetic trees, and these errors are not trivial.

Alternatively, students may be hastier when constructing diagonal phylogenetic trees and inadvertently include more minor errors. Diagonal phylogenetic trees contain about one-third the number of lines as equivalent bracket phylogenetic trees. Thus, when resulting phylogenetic trees for construction tasks are not known in advance, the diagonal style is simpler and much faster for trial-and-error approaches. We hypothesize that simplicity and speed are the primary reasons why students consistently preferred to construct diagonal phylogenetic trees when allowed to use the style of their choice during this study and two previous investigations (Dees and Momsen, 2016; Dees et al., accepted). Therefore, the speed and ease of using the diagonal style for construction tasks may have led students to inadvertently include more minor errors (i.e., sloppiness).

Implications and Future Directions

Based on this study and prior research on style effects, we join Novick and Catley (2007, 2013) in recommending introductory biology instructors use only the bracket style for instruction on phylogenetic trees and as visual representations of evolution in general. However, diagonal and bracket phylogenetic trees are both commonly used by biologists (Catley and Novick, 2008), which necessitates that biology majors gain familiarity with diagonal phylogenetic trees in their upper-division coursework. Given that significant style effects were observed for some tasks after instruction that integrated diagonal and bracket phylogenetic trees equally, it is likely these style effects will persist in upper-division courses without instructional interventions. Unfortunately, we are unaware of specific pedagogy that has successfully mitigated all style effects for student interpretations and construction of phylogenetic trees, and it is unlikely such pedagogy will be developed by instructors without first determining why these style effects exist.

One intriguing hypothesis, supported by evidence, is that students perceive diagonal and bracket phylogenetic trees differently. Novick and Catley (2007) used translation exercises (i.e., converting one visual representation of evolution to another while retaining the same information) to demonstrate that students often interpret lines of diagonal phylogenetic trees as single entities, whether accurate or not. In the diagonal phylogenetic tree of Figure 8, for example, the line from node A to koalas is a single branch. However, students may also interpret the line from node C to saltwater crocodiles as a single branch rather than two branches. In contrast, it is more apparent in the equivalent bracket phylogenetic tree of Figure 8 that two branches occur between node B and black caimans. Thus, the hierarchical structure of monophyletic groups within phylogenetic trees could be obscured by the diagonal style, and as a result, student understanding may be impeded.

Diagonal phylogenetic trees could also disproportionately encourage misinterpretations of evolution as a ladder of progress from “lower” to “higher” organisms. This hypothesis emerges from classroom observations of introductory biology students constructing a phylogenetic tree of large groups of vertebrates (e.g., amphibians and mammals) in the style of their choice. Because branches can be rotated around nodes on phylogenetic trees without changing relationships, taxa can appear in almost any order along the branch tips (Baum and Offner, 2008). However, when our students used the diagonal style to construct a phylogenetic tree of vertebrates during class, mammals almost invariably appeared in the right-most position. Conversely, we did not observe any discernable pattern when students used the bracket style, as mammals appeared in a random location along the branch tips. Therefore, it is possible diagonal phylogenetic trees reinforce the common misinterpretation of evolution as a ladder of progress toward a goal, which is generally humans and other mammals (Gregory, 2008). Consequently, students could disproportionately focus on irrelevant features when interpreting diagonal phylogenetic trees, such as the number of nodes between taxa and the proximity of branch tips (Table 11). We believe the hypothesis put forth by Novick and Catley (2007) is one critical driver of differences in student performance across phylogenetic trees styles. However, other factors likely contribute to these style effects, and future research should explore alternative hypotheses.

In addition to style, other variables may influence student understanding of phylogenetic trees. For example, equivalent phylogenetic trees can be drawn in a vertical, horizontal, or even circular orientation. To our knowledge, only one study has investigated effects of orientation on student comprehension. Phillips et al. (2012) found no significant difference in accuracy for two tasks, identifying monophyletic groups and evaluating taxa relatedness, between horizontal and vertical phylogenetic trees drawn only in the bracket style. Further, most phylogenetic trees in textbooks and other instructional resources are not scaled for time or degree of divergence (i.e., chronograms and phylograms), and it is unknown whether scaled phylogenetic trees would help or hinder student comprehension. Future research should explore variables other than style, such as orientation and scaling, that could impact student understanding of phylogenetic trees.

CONCLUSIONS

The goal of this dissertation research was to systematically investigate the effects of style on student interpretations and construction of phylogenetic trees in the context of an introductory biology course. Research-based tools to support this endeavor were designed in conjunction with my MS thesis (published as Dees et al., 2014) and the first two investigations of this dissertation (published as Dees and Momsen, 2016; Dees et al., accepted). The tools and knowledge acquired from these studies were subsequently used for the third and final investigation of this dissertation, which fulfilled the research goal by collecting data throughout an introductory biology course and by controlling as many variables as was feasible in an authentic classroom setting.

Prior to instruction on phylogenetic trees, students were significantly more accurate with the bracket style for all interpretation tasks except evaluating taxa relatedness. These differences in accuracy were mitigated by instruction that integrated diagonal and bracket phylogenetic trees equally. In contrast, student interpretations of taxa relatedness exhibited a different pattern. Prior to instruction, and in alignment with previous research (Meir et al., 2007; Halverson et al., 2011; Phillips et al., 2012; Catley et al., 2013; Smith et al., 2013; Dees et al., 2014), few students could accurately determine taxa relatedness on either style of phylogenetic tree. Following instruction, students were significantly more accurate, for both answers and reasoning, when evaluating taxa relatedness on bracket phylogenetic trees. Thus, instruction that balanced the use of diagonal and bracket phylogenetic trees was not beneficial for style effects in regards to taxa relatedness.

For all construction tasks, before and after instruction, the majority of phylogenetic trees generated by students were correct (i.e., no errors) or adequate (i.e., minor errors only). However, bracket phylogenetic trees were significantly more accurate than the diagonal style due to a much lower incidence of minor errors (e.g., extra nodes and empty branches). These minor errors could be indicative of misinterpretations associated with phylogenetic trees, such as the common belief among students that evolutionary changes occurred only at nodes (Baum et al., 2005; Meir et al., 2007; Gregory, 2008). Alternatively, because diagonal phylogenetic trees contain fewer lines and are faster to draw, students may simply be hastier when constructing diagonal phylogenetic trees and inadvertently include more minor errors. Additional research is needed to determine whether minor errors are indicative of misinterpretations or merely the result of sloppiness, and therefore, whether the difference in construction accuracy between styles is meaningful or superficial.

Based on the results of this dissertation research and prior studies on style effects (Novick and Catley, 2007, 2013), I recommend introductory biology instructors use only the bracket style for instruction on phylogenetic trees and evolution in general. However, both styles are common in professional biology (Catley and Novick, 2008), which dictates biology majors gain familiarity with diagonal phylogenetic trees at some point in their upper-division courses. Thus, the problem can be avoided at the introductory level, but eventually we have to face the music. Unfortunately, I am unaware of any pedagogy that has eliminated all style effects for student interpretations and construction of phylogenetic trees at any education level, and it is highly unlikely such pedagogy will be developed by instructors without first determining why style effects exist.

One hypothesis is that students visually perceive diagonal and bracket phylogenetic trees differently. Novick and Catley (2007) used a series of translation tasks (i.e., converting one visual representation of evolution to another while retaining the same information) to show that students often interpret the continuous lines of diagonal phylogenetic trees as single entities, regardless of accuracy. A single line on a diagonal phylogenetic tree may represent a single branch or multiple branches, and therefore, the hierarchical structure of phylogenetic trees could be obscured by the diagonal style and result in significantly more misinterpretations. Alternatively, the diagonal style may disproportionately promote misinterpretations of evolution as a ladder of progress from lower to higher organisms (i.e., great chain of being or scala naturae). This hypothesis emerges from my observations of students constructing phylogenetic trees of vertebrates in the style of their choice during class. Because branches can be rotated around nodes without changing relationships, taxa could appear in almost any order along the branch tips of phylogenetic trees. However, mammals almost invariably appeared in the right-most position when students chose to construct a diagonal phylogenetic tree. In contrast, no patterns were observed when students used the bracket style, as mammals appeared in seemingly random locations along the branch tips. Therefore, it is possible diagonal phylogenetic trees reinforce the widespread misinterpretation of evolution as a ladder of toward a goal, which is usually humans and other mammals (Gregory, 2008). As a result, students might focus on irrelevant features more frequently when interpreting diagonal phylogenetic trees, such as the spatial proximity of branch tips or the number of nodes between taxa.

I believe the perceptual hypothesis put forth by Novick and Catley (2007) is one driver of differences in student performance across phylogenetic tree styles, but other factors likely have a significant role in style effects. Future research should explore alternative hypotheses, such as the extent to which diagonal phylogenetic trees disproportionately encourage students to reason about evolution as a ladder of progress. In addition to style effects, other variables could impact student understanding of phylogenetic trees. For example, phylogenetic trees can be drawn in a variety of orientations and scaled for time or degree of divergence. Future research should explore variables other than style, such as orientation and scaling, that have the potential to influence how students interact with phylogenetic trees. Once we determine which variables affect student interpretations and construction of phylogenetic trees and why such variables are important, we can design, test, and deploy research-based instruction that best promotes student learning.

REFERENCES

American Association for the Advancement of Science (2011). Vision and Change in Undergraduate Biology Education: A Call to Action. Washington, DC.

Armstrong N, Chang SM, Brickman M (2007). Cooperative Learning in Industrial-Sized Biology Classes. CBE-Life Sciences Education 6, 163-171.

Baum DA, Offner S (2008). Phylogenies and Tree Thinking. American Biology Teacher 70, 222-229.

Baum DA, Smith SD (2013). Tree Thinking: An Introduction to Phylogenetic Biology. Greenwood Village, CO: Roberts and Company.

Baum DA, Smith SD, Donovan SS (2005). The Tree-Thinking Challenge. Science 310, 979-980.

Blacquiere LD, Hoese WJ (2016). A Valid Assessment of Students' Skill in Determining Relationships on Evolutionary Trees. Evolution: Education and Outreach 9, ar5.

Bokor JR, Landis JB, Crippen KJ (2014). High School Students' Learning and Perceptions of Phylogenetics of Flowering Plants. CBE-Life Sciences Education 13, 653-665.

Bray Speth E, Shaw N, Momsen J, Reinagel A, Le P, Taqieddin R, Long T (2014). Introductory Biology Students' Conceptual Models and Explanations of the Origin of Variation. CBE-Life Sciences Education 13, 529-539.

Burks RL, Boles LC (2007). Evolution of the Chocolate Bar: A Creative Approach to Teaching Phylogenetic Relationships within Evolutionary Biology. American Biology Teacher 69, 229-237.

Caldwell JE (2007). Clickers in the Large Classroom: Current Research and Best-Practice Tips. CBE-Life Sciences Education 6, 9-20.

Catley KM, Novick LR (2008). Seeing the Wood for the Trees: An Analysis of Evolutionary Diagrams in Biology Textbooks. BioScience 58, 976-987.

Catley KM, Phillips BC, Novick LR (2013). Snakes and Eels and Dogs! Oh, My! Evaluating High School Students' Tree-Thinking Skills: An Entry Point to Understanding Evolution. Research in Science Education 43, 2327-2348.

Cohen J (1960). A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 20, 37-46.

Cohen-Schotanus J (1999). Student Assessment and Examination Rules. Medical Teacher 21, 318-321.

College Board (2015). AP Biology: Course and Exam Description. New York, NY.

Cortright RN, Collins HL, Rodenbaugh DW, DiCarlo SE (2003). Student Retention of Course Content Is Improved by Collaborative-Group Testing. Advances in Physiology Education 27, 102-108.

Dauer JT, Momsen JL, Bray Speth E, Makohon-Moore SC, Long TM (2013). Analyzing Change in Students' Gene-to-Evolution Models in College-Level Introductory Biology. Journal of Research in Science Teaching 50, 639-659.

Dees J, Freiermuth D, Momsen JL (accepted). Effects of Phylogenetic Tree Style on Student Comprehension in an Introductory Biology Course. American Biology Teacher.

Dees J, Momsen JL (2016). Student Construction of Phylogenetic Trees in an Introductory Biology Course. Evolution: Education and Outreach 9, ar3.

Dees J, Momsen JL, Niemi J, Montplaisir L (2014). Student Interpretations of Phylogenetic Trees in an Introductory Biology Course. CBE-Life Sciences Education 13, 666-676.

Dobzhansky T (1964). Biology, Molecular and Organismic. American Zoologist 4, 443-452.

Dobzhansky T (1973). Nothing in Biology Makes Sense Except in the Light of Evolution. American Biology Teacher 35, 125-129.

Driver R, Newton P, Osborne J (2000). Establishing the Norms of Scientific Argumentation in Classrooms. Science Education 84, 287-312.

Eddy SL, Crowe AJ, Wenderoth MP, Freeman S (2013). How Should We Teach Tree Thinking? An Experimental Test of Two Hypotheses. Evolution: Education and Outreach 6, ar13.

Federer MR, Nehm RH, Opfer JE, Pearl D (2015). Using a Constructed-Response Instrument to Explore the Effects of Item Position and Item Features on the Assessment of Students' Written Scientific Explanations. Research in Science Education 45, 527-553.

Fisher RA (1934). Statistical Methods for Research Workers. Edinburgh, UK: Oliver and Boyd.

Freeman S (2011). Biological Science: Fourth Edition. San Francisco, CA: Benjamin Cummings.

Freeman S, O'Connor E, Parks JW, Cunningham M, Hurley D, Haak D, Dirks C, Wenderoth MP (2007). Prescribed Active Learning Increases Performance in Introductory Biology. CBE-Life Sciences Education 6, 132-139.

Gendron RP (2000). The Classification and Evolution of Caminalcules. American Biology Teacher 62, 570-576.

Goldsmith DW (2003). The Great Clade Race: Presenting Cladistic Thinking to Biology Majors and General Science Students. American Biology Teacher 65, 679-682.

Gregory TR (2008). Understanding Evolutionary Trees. Evolution: Education and Outreach 1, 121-137.

Halverson KL (2011). Improving Tree Thinking One Learnable Skill at a Time. Evolution: Education and Outreach 4, 95-106.

Halverson KL, Boyce CJ, Maroo JD (2013). Order Matters: Pre-Assessments and Student Generated Representations. Evolution: Education and Outreach 6, ar24.

Halverson KL, Pires CJ, Abell SK (2011). Exploring the Complexity of Tree-Thinking Expertise in an Undergraduate Systematics Course. Science Education 95, 794-823.

Hobbs FC, Johnson DJ, Kearns KD (2013). A Deliberate Practice Approach to Teaching Phylogenetic Analysis. CBE-Life Sciences Education 12, 676-686.

Johnson DW, Johnson RT, Smith KA (1998). Cooperative Learning Returns to College: What Evidence Is There That It Works? Change 30, 26-35.

Julius ML, Schoenfuss HL (2006). Phylogenetic Reconstruction as a Broadly Applicable Teaching Tool in the Biology Classroom: The Value of Data in Estimating Likely Answers. Journal of College Science Teaching 35, 40-45.

Lampert E, Mook J (2015). Modeling with Nonliving Objects to Enhance Understanding of Phylogenetic Tree Construction. American Biology Teacher 77, 587-599.

Lents NH, Cifuentes OE, Carpi A (2010). Teaching the Process of Molecular Phylogeny and Systematics: A Multi-Part Inquiry-Based Exercise. CBE-Life Sciences Education 9, 513-523.

Long TM, Dauer JT, Kostelnik KM, Momsen JL, Wyse SA, Bray Speth E, Ebert-May D (2014). Fostering Ecoliteracy Through Model-Based Instruction. Frontiers in Ecology and the Environment 12, 138-139.

Marsh EJ, Lozito JP, Umanath S, Bjork EL, Bjork RA (2012). Using Verification Feedback to Correct Errors Made on a Multiple-Choice Test. Memory 20, 645-653.

Maxwell AE (1970). Comparing the Classification of Subjects by Two Independent Judges. British Journal of Psychiatry 116, 651-655.

McDonald JH (2014). Handbook of Biological Statistics: Third Edition. Baltimore, MD: Sparky House Publishing.

McNemar Q (1947). Note on the Sampling Error of the Difference Between Correlated Proportions or Percentages. Psychometrika 12, 153-157.

Meir E, Perry J, Herron JC, Kingsolver J (2007). College Students' Misconceptions About Evolutionary Trees. American Biology Teacher 69, 71-76.

Meisel RP (2010). Teaching Tree Thinking to Undergraduate Biology Students. Evolution: Education and Outreach 3, 621-628.

Morabito NP, Catley KM, Novick LR (2010). Reasoning About Evolutionary History: Post-Secondary Students' Knowledge of Most Recent Common Ancestry and Homoplasy. Journal of Biological Education 44, 166-174.

National Research Council (2012). Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering. Washington, DC: National Academies Press.

Nature Education (2013). Principles of Biology. Cambridge, MA: Nature Publishing Group.

Nehm RH, Ha M (2011). Item Feature Effects in Evolution Assessment. Journal of Research in Science Teaching 48, 237-256.

Next Generation Science Standards Lead States (2013). Next Generation Science Standards: For States, By States. Washington, DC: National Academies Press.

Novick LR, Catley KM (2007). Understanding Phylogenies in Biology: The Influence of a Gestalt Perceptual Principle. Journal of Experimental Psychology: Applied 13, 197-223.

Novick LR, Catley KM (2013). Reasoning About Evolution's Grand Patterns: College Students' Understanding of the Tree of Life. American Educational Research Journal 50, 138-177.

Novick LR, Catley KM (2016). Fostering 21st-Century Evolutionary Reasoning: Teaching Tree Thinking to Introductory Biology Students. CBE-Life Sciences Education 15, ar66.

Novick LR, Stull AT, Catley KM (2012). Reading Phylogenetic Trees: The Effects of Tree Orientation and Text Processing on Comprehension. BioScience 62, 757-764.

O'Hara RJ (1988). Homage to Clio, or, Toward a Historical Philosophy for Evolutionary Biology. Systematic Zoology 37, 142-155.

O'Hara RJ (1997). Population Thinking and Tree Thinking in Systematics. Zoologica Scripta 26, 323-329.

Omland KE, Cook LG, Crisp MD (2008). Tree Thinking for All Biology: The Problem with Reading Phylogenies as Ladders of Progress. BioEssays 30, 854-867.

Perez KE, Strauss EA, Downey N, Galbraith A, Jeanne R, Cooper S (2010). Does Displaying the Class Results Affect Student Discussion During Peer Instruction? CBE-Life Sciences Education 9, 133-140.

Phillips BC, Novick LR, Catley KM, Funk DJ (2012). Teaching Tree Thinking to College Students: It's Not as Easy as You Think. Evolution: Education and Outreach 5, 595-602.

Rufibach K (2011). Assessment of Paired Binary Data. Skeletal Radiology 40, 1-4.

Sandvik H (2008). Tree Thinking Cannot Be Taken for Granted: Challenges for Teaching Phylogenetics. Theory in Biosciences 127, 45-51.

Singer F, Hagen JB, Sheehy RR (2001). The Comparative Method, Hypothesis Testing and Phylogenetic Analysis: An Introductory Lab. American Biology Teacher 63, 518-523.

Smith KA (2000). Going Deeper: Formal Small-Group Learning in Large Classes. New Directions for Teaching and Learning 81, 25-46.

Smith JJ, Cheruvelil KS, Auvenshine S (2013). Assessment of Student Learning Associated with Tree Thinking in an Undergraduate Introductory Organismal Biology Course. CBE-Life Sciences Education 12, 542-552.

Stuart A (1955). A Test for Homogeneity of the Marginal Distributions in a Two-Way Classification. Biometrika 42, 412-416.

Sun X, Yang Z (2008). Generalized McNemar's Test for Homogeneity of the Marginal Distributions. Proceedings of the SAS Global Forum, ar382.

Sundberg MD (2002). Assessing Student Learning. Cell Biology Education 1, 11-15.

Tanner K, Chatman LS, Allen D (2003). Approaches to Cell Biology Teaching: Cooperative Learning in the Science Classroom-Beyond Students Working in Groups. Cell Biology Education 2, 1-5.

Thanukos A (2009). A Name by Any Other Tree. Evolution: Education and Outreach 2, 303-309.

Thomas DR (2006). A General Inductive Approach for Analyzing Qualitative Evaluation Data. American Journal of Evaluation 27, 237-246.

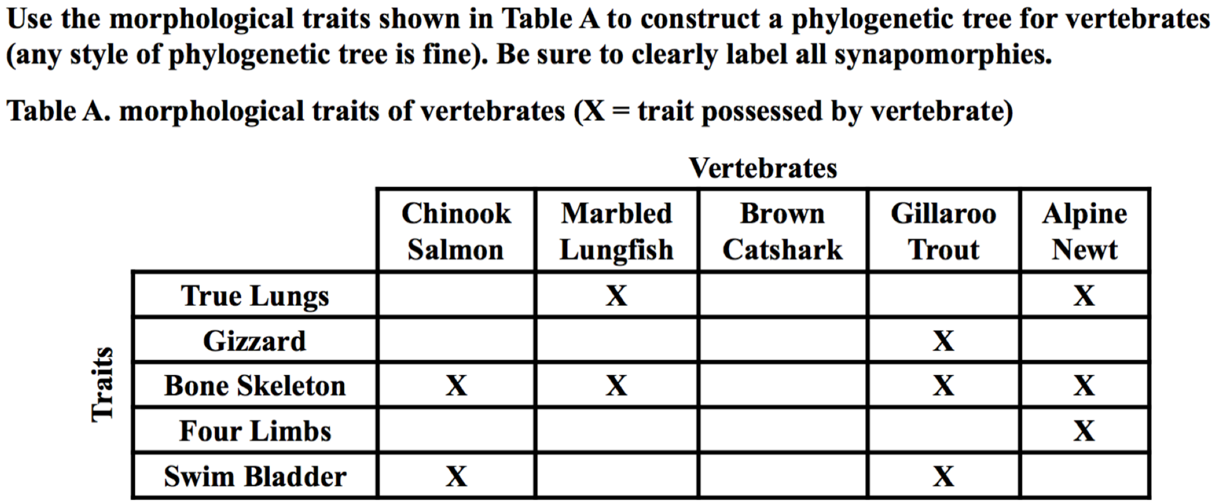
Urry LA, Cain ML, Wasserman SA, Minorsky PV, Jackson RB, Reece JB (2014). Campbell Biology in Focus. Boston, MA: Pearson Education.

Wiley EO (2010). Why Trees Are Important. Evolution: Education and Outreach 3, 499-505.

Wormald BW, Schoeman S, Somasunderam A, Penn M (2009). Assessment Drives Learning: An Unavoidable Truth? Anatomical Sciences Education 2, 199-204.

Young AK, White BT, Skurtu T (2013). Teaching Undergraduate Students to Draw Phylogenetic Trees: Performance Measures and Successes. Evolution: Education and Outreach 6, ar16.

APPENDIX A. SUPPLEMENTARY MATERIAL FOR STUDY 1

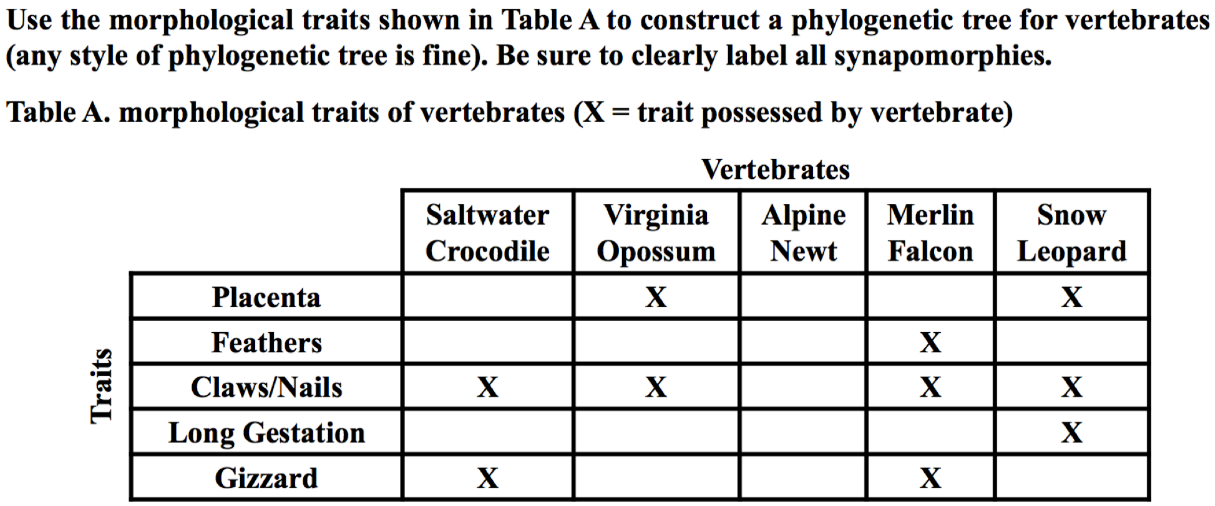


Use the morphological traits shown in Table A to construct a phylogenetic tree for vertebrates (any style of phylogenetic tree is fine). Be sure to clearly label all synapomorphies.

Table A. morphological traits of vertebrates (X = trait possessed by vertebrate)

		Vertebrates				
		Chinook Salmon	**Marbled Lungfish**	**Brown Catshark**	**Gillaroo Trout**	**Alpine Newt**
Traits	**True Lungs**		X			X
	Gizzard				X	
	Bone Skeleton	X	X		X	X
	Four Limbs					X
	Swim Bladder	X			X	

Figure A1. Version A of the phylogenetic tree construction task from the individual component of the comprehensive final exam.



Use the morphological traits shown in Table A to construct a phylogenetic tree for vertebrates (any style of phylogenetic tree is fine). Be sure to clearly label all synapomorphies.

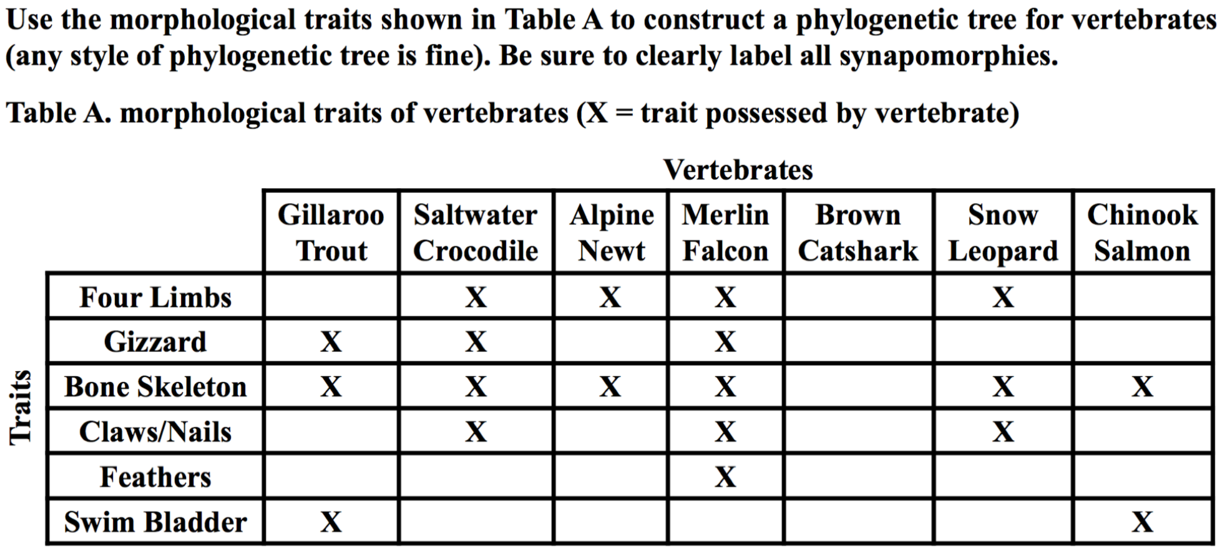
Table A. morphological traits of vertebrates (X = trait possessed by vertebrate)

		Vertebrates				
		Saltwater Crocodile	**Virginia Opossum**	**Alpine Newt**	**Merlin Falcon**	**Snow Leopard**
Traits	**Placenta**		X			X
	Feathers				X	
	Claws/Nails	X	X		X	X
	Long Gestation					X
	Gizzard	X			X	

Figure A2. Version B of the phylogenetic tree construction task from the individual component of the comprehensive final exam.

Table A1. Phylogenetic tree styles present in various components of the course curriculum.

Phylogenetic Tree Style	**Course Textbook**	**Instruction [n=17]**	**Assessments [n=13]**
Bracket	All (100%)	11 (65%)	3 (23%)
Diagonal	0 (0%)	6 (35%)	7 (54%)
Construction Task	0 (0%)	0 (0%)	3 (23%)



Use the morphological traits shown in Table A to construct a phylogenetic tree for vertebrates (any style of phylogenetic tree is fine). Be sure to clearly label all synapomorphies.

Table A. morphological traits of vertebrates (X = trait possessed by vertebrate)

Traits	Vertebrates						
	Gillaroo Trout	**Saltwater Crocodile**	**Alpine Newt**	**Merlin Falcon**	**Brown Catshark**	**Snow Leopard**	**Chinook Salmon**
Four Limbs		X	X	X		X	
Gizzard	X	X		X			
Bone Skeleton	X	X	X	X		X	X
Claws/Nails		X		X		X	
Feathers				X			
Swim Bladder	X						X

Figure A3. Phylogenetic tree construction task from the group component of the comprehensive final exam that included convergent evolution (gizzard).

APPENDIX B. SUPPLEMENTARY MATERIAL FOR STUDY 2

Pre-Instructional Homework (Diagonal Style)

The following interpretation tasks from the pre-instructional homework accompanied the diagonal phylogenetic tree of chordates in Figure 6 (top).

Trait Possession

According to the phylogenetic tree of chordates in Figure 6, which of the following traits do bald eagles possess? Draw an X inside the box below each possessed trait.

cranium	placenta	feathers	jaw	lactation	long gestation	gizzard	bone skeleton	claws or nails	true lungs

Most Recent Common Ancestor

Which node on the phylogenetic tree of chordates in Figure 6 represents the most recent common ancestor of platypuses and American alligators?

a) node A b) node B c) node C d) node D

Monophyletic Group

Draw a circle around one monophyletic group on the phylogenetic tree in Figure 6.

Contemporary Descent

According to the phylogenetic tree of chordates in Figure 6, did bald eagles evolve from American alligators? Explain your reasoning below.

Taxa Relatedness

According to the phylogenetic tree of chordates in Figure 6, how would you describe the relatedness of American alligators to fire salamanders and red kangaroos?

a) American alligators are more closely related to fire salamanders than red kangaroos.

b) American alligators are more closely related to red kangaroos than fire salamanders.

c) American alligators are equally related to fire salamanders and red kangaroos.

d) American alligators are not related to fire salamanders and red kangaroos.

Explain the reasoning for your choice below.

Pre-Instructional Homework (Bracket Style)

The following interpretation tasks from the pre-instructional homework accompanied the bracket phylogenetic tree of chordates in Figure 6 (bottom).

Trait Possession

According to the phylogenetic tree of chordates in Figure 6, which of the following traits do griffon vultures possess? Draw an X inside the box below each possessed trait.

feathers	placenta	cranium	jaw	hair or fur	bone skeleton	four limbs	long gestation	amnion	gizzard

Most Recent Common Ancestor

Which node on the phylogenetic tree of chordates in Figure 6 represents the most recent common ancestor of short-beak echidnas and Siamese crocodiles?

a) node A b) node B c) node C d) node D

Monophyletic Group

Draw a circle around one monophyletic group on the phylogenetic tree in Figure 6.

Contemporary Descent

According to the phylogenetic tree of chordates in Figure 6, did griffon vultures evolve from Siamese crocodiles? Explain your reasoning below.

Taxa Relatedness

According to the phylogenetic tree of chordates in Figure 6, how would you describe the relatedness of Siamese crocodiles to marbled newts and koalas?

a) Siamese crocodiles are more closely related to marbled newts than koalas.

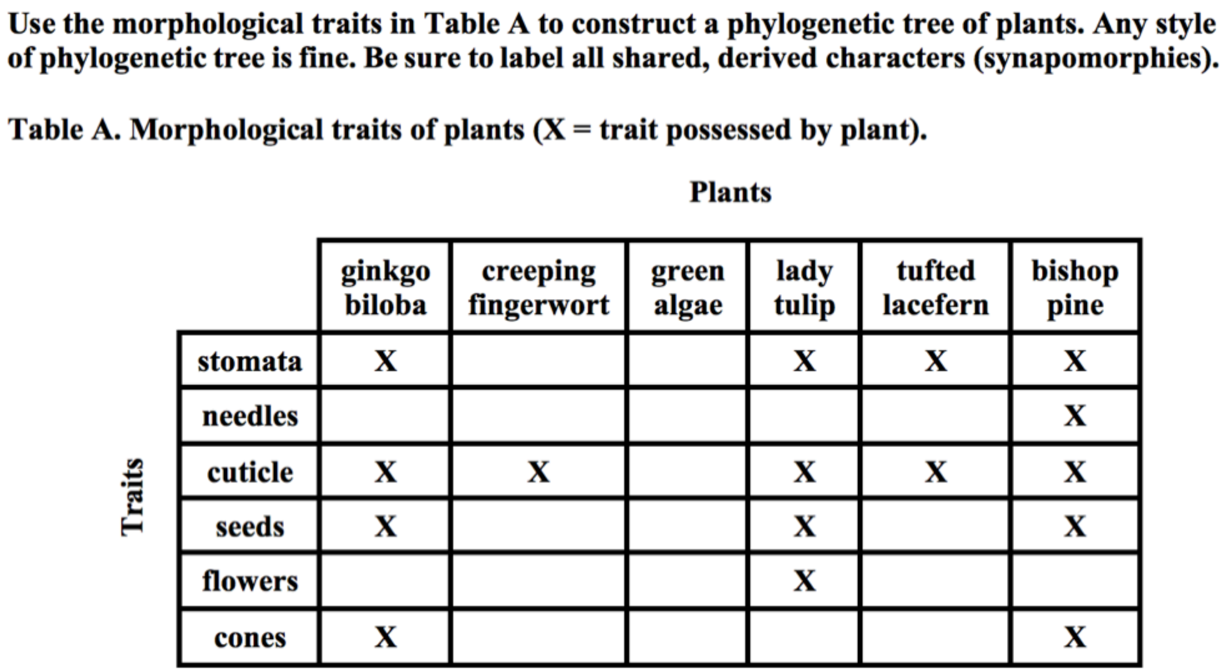
b) Siamese crocodiles are more closely related to koalas than marbled newts.

c) Siamese crocodiles are equally related to marbled newts and koalas.

d) Siamese crocodiles are not related to marbled newts and koalas.

Explain the reasoning for your choice below.

Construction Task (Optional Style)



Use the morphological traits in Table A to construct a phylogenetic tree of plants. Any style of phylogenetic tree is fine. Be sure to label all shared, derived characters (synapomorphies).

Table A. Morphological traits of plants (X = trait possessed by plant).

Plants

Traits	ginkgo biloba	creeping fingerwort	green algae	lady tulip	tufted lacefern	bishop pine
stomata	X			X	X	X
needles						X
cuticle	X	X		X	X	X
seeds	X			X		X
flowers				X		
cones	X					X

Figure B1. Construction task placed at the end of the pre-instructional homework.

APPENDIX C. SUPPLEMENTARY MATERIAL FOR STUDY 3

Instrument #1: Pre-Instructional Homework (Diagonal Style)

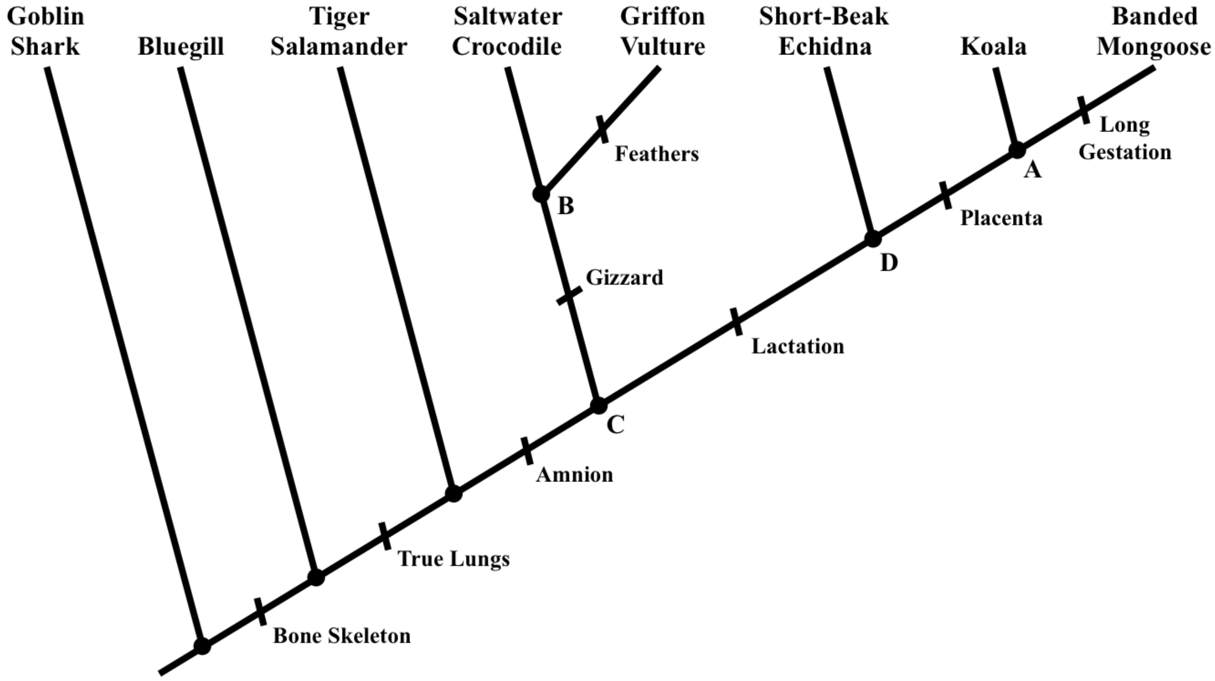


Figure C1. Phylogenetic tree from the pre-instructional homework (diagonal style).

Most Recent Common Ancestor

Which node on the phylogenetic tree in Figure C1 represents the most recent common ancestor of griffon vultures and short-beak echidnas? Select one option.

____ node A ____ node B ____ node C ____ node D

Monophyletic Group

On the phylogenetic tree in Figure C1, draw a circle around one monophyletic group that includes at least two species.

Contemporary Descent

According to the phylogenetic tree of vertebrates in Figure C1, did griffon vultures evolve from saltwater crocodiles? Select one option and explain the reasoning for your choice.

____ Yes ____ No

Taxa Relatedness

According to the phylogenetic tree of vertebrates in Figure C1, how would you describe the relatedness of saltwater crocodiles to tiger salamanders and koalas? Select one option.

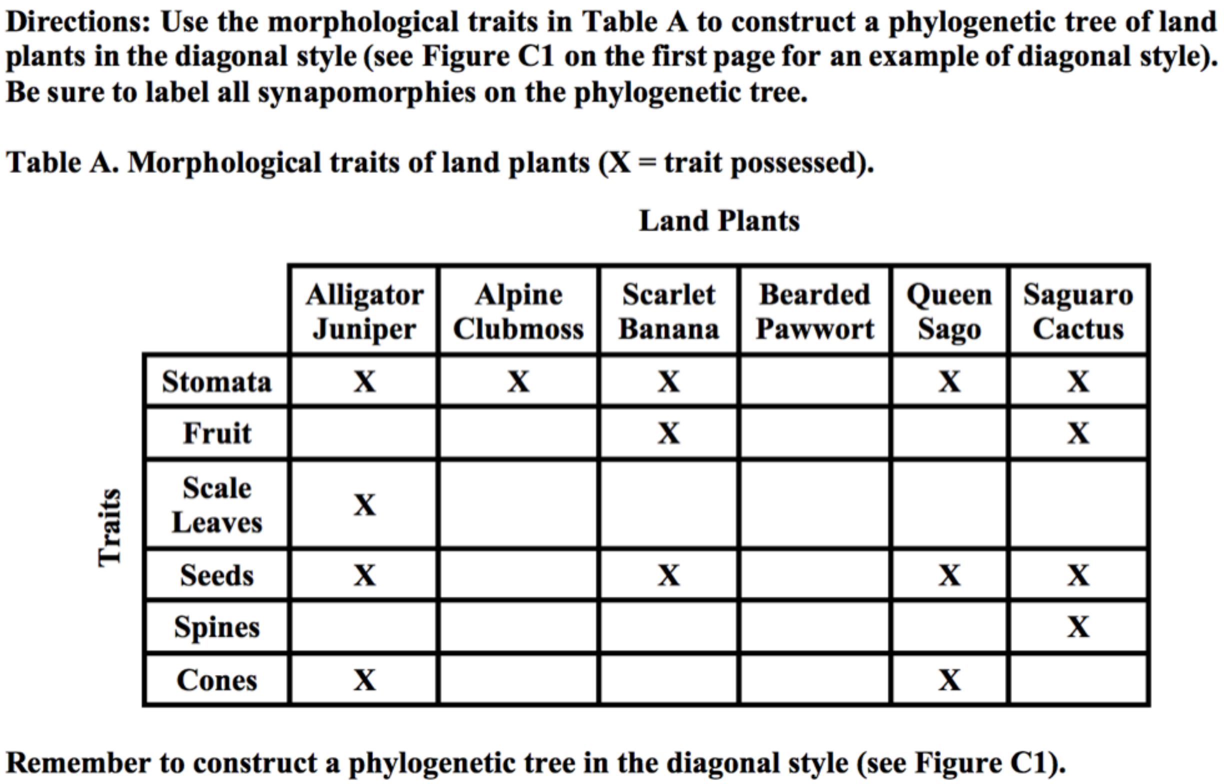
____ Saltwater crocodiles are more closely related to tiger salamanders than koalas.

____ Saltwater crocodiles are equally related to tiger salamanders and koalas.

____ Saltwater crocodiles are more closely related to koalas than tiger salamanders.

Explain the reasoning for your choice.

Construction Task



Directions: Use the morphological traits in Table A to construct a phylogenetic tree of land plants in the diagonal style (see Figure C1 on the first page for an example of diagonal style). Be sure to label all synapomorphies on the phylogenetic tree.

Table A. Morphological traits of land plants (X = trait possessed).

Land Plants

Traits	Alligator Juniper	Alpine Clubmoss	Scarlet Banana	Bearded Pawwort	Queen Sago	Saguaro Cactus
Stomata	X	X	X		X	X
Fruit			X			X
Scale Leaves	X					
Seeds	X		X		X	X
Spines						X
Cones	X				X	

Remember to construct a phylogenetic tree in the diagonal style (see Figure C1).

Figure C2. Construction task from the pre-instructional homework (diagonal style).

Instrument #1: Pre-Instructional Homework (Bracket Style)

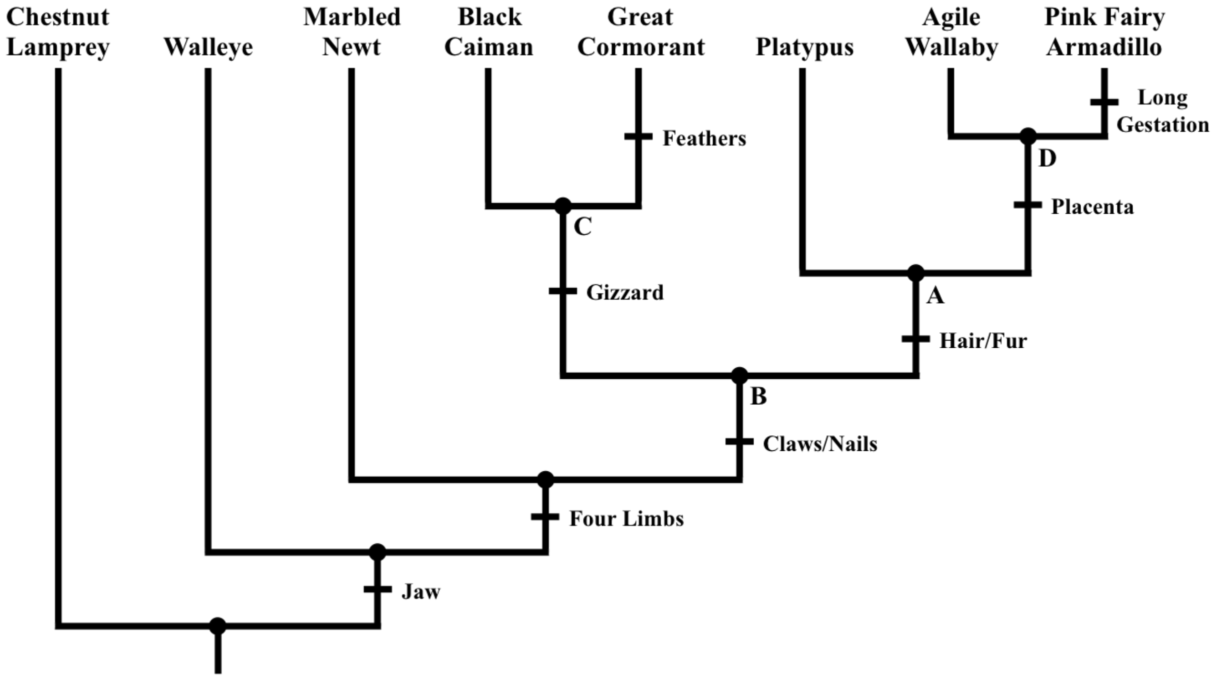


Figure C3. Phylogenetic tree from the pre-instructional homework (bracket style).

Most Recent Common Ancestor

Which node on the phylogenetic tree in Figure C3 represents the most recent common ancestor of great cormorants and platypuses? Select one option.

____ node A ____ node B ____ node C ____ node D

Monophyletic Group

On the phylogenetic tree in Figure C3, draw a circle around one monophyletic group that includes at least two species.

Contemporary Descent

According to the phylogenetic tree of vertebrates in Figure C3, did great cormorants evolve from black caimans? Select one option and explain the reasoning for your choice.

____ Yes ____ No

Taxa Relatedness

According to the phylogenetic tree of vertebrates in Figure C3, how would you describe the relatedness of black caimans to marbled newts and agile wallabies? Select one option.

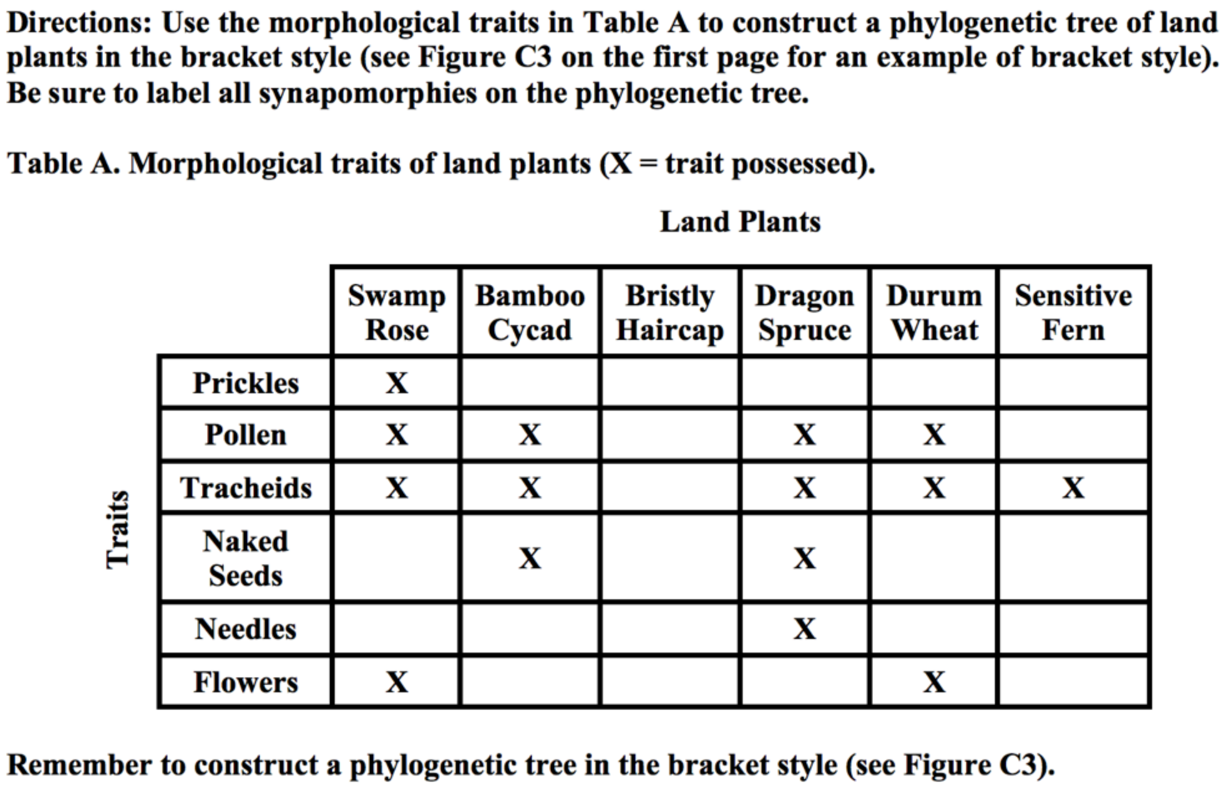
____ Black caimans are more closely related to agile wallabies than marbled newts.

____ Black caimans are equally related to marbled newts and agile wallabies.

____ Black caimans are more closely related to marbled newts than agile wallabies.

Explain the reasoning for your choice.

Construction Task



Directions: Use the morphological traits in Table A to construct a phylogenetic tree of land plants in the bracket style (see Figure C3 on the first page for an example of bracket style). Be sure to label all synapomorphies on the phylogenetic tree.

Table A. Morphological traits of land plants (X = trait possessed).

Land Plants

Traits	Swamp Rose	Bamboo Cycad	Bristly Haircap	Dragon Spruce	Durum Wheat	Sensitive Fern
Prickles	X					
Pollen	X	X		X	X	
Tracheids	X	X		X	X	X
Naked Seeds		X		X		
Needles				X		
Flowers	X				X	

Remember to construct a phylogenetic tree in the bracket style (see Figure C3).

Figure C4. Construction task from the pre-instructional homework (bracket style).

Instrument #2: Post-Instructional Homework (Diagonal Style)

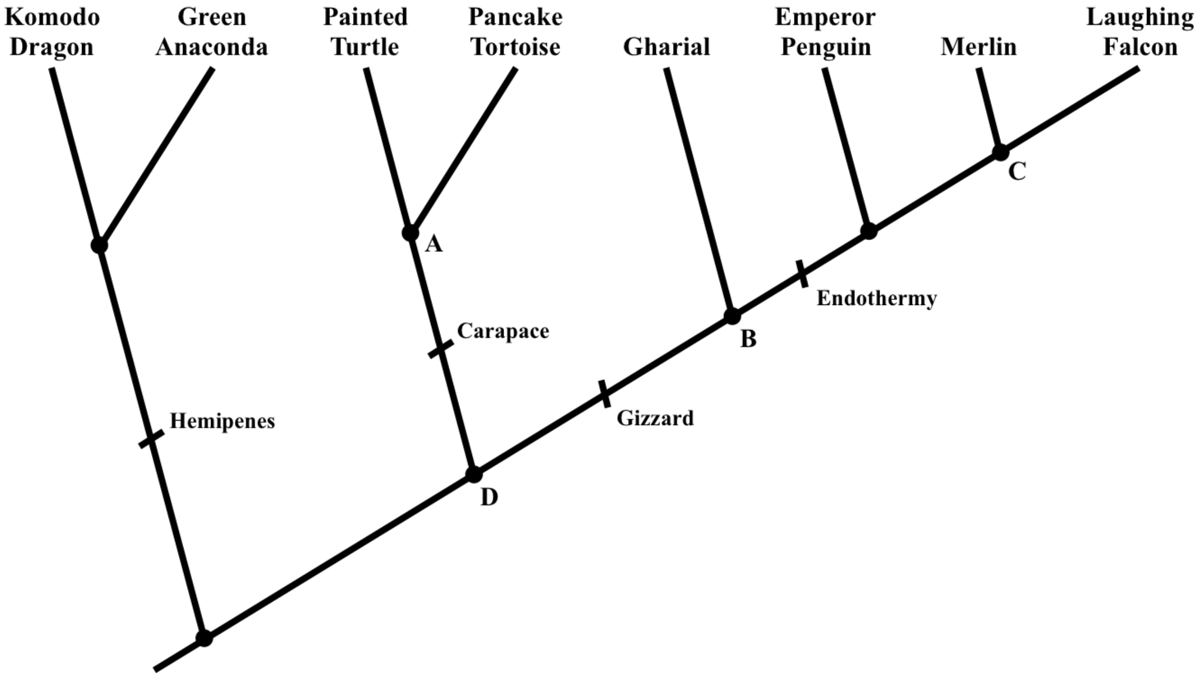


Figure C5. Phylogenetic tree from the post-instructional homework (diagonal style).

Most Recent Common Ancestor

Which node on the phylogenetic tree in Figure C5 represents the most recent common ancestor of pancake tortoises and gharials? Select one option.

____ node A ____ node B ____ node C ____ node D

Monophyletic Group

On the phylogenetic tree in Figure C5, draw a circle around one monophyletic group that includes at least two species.

Contemporary Descent

According to the phylogenetic tree of reptiles and birds in Figure C5, did pancake tortoises evolve from painted turtles? Select one option and explain the reasoning for your choice.

____ Yes ____ No

Taxa Relatedness

According to the phylogenetic tree of reptiles and birds in Figure C5, how would you describe the relatedness of painted turtles to green anacondas and merlins? Select one option.

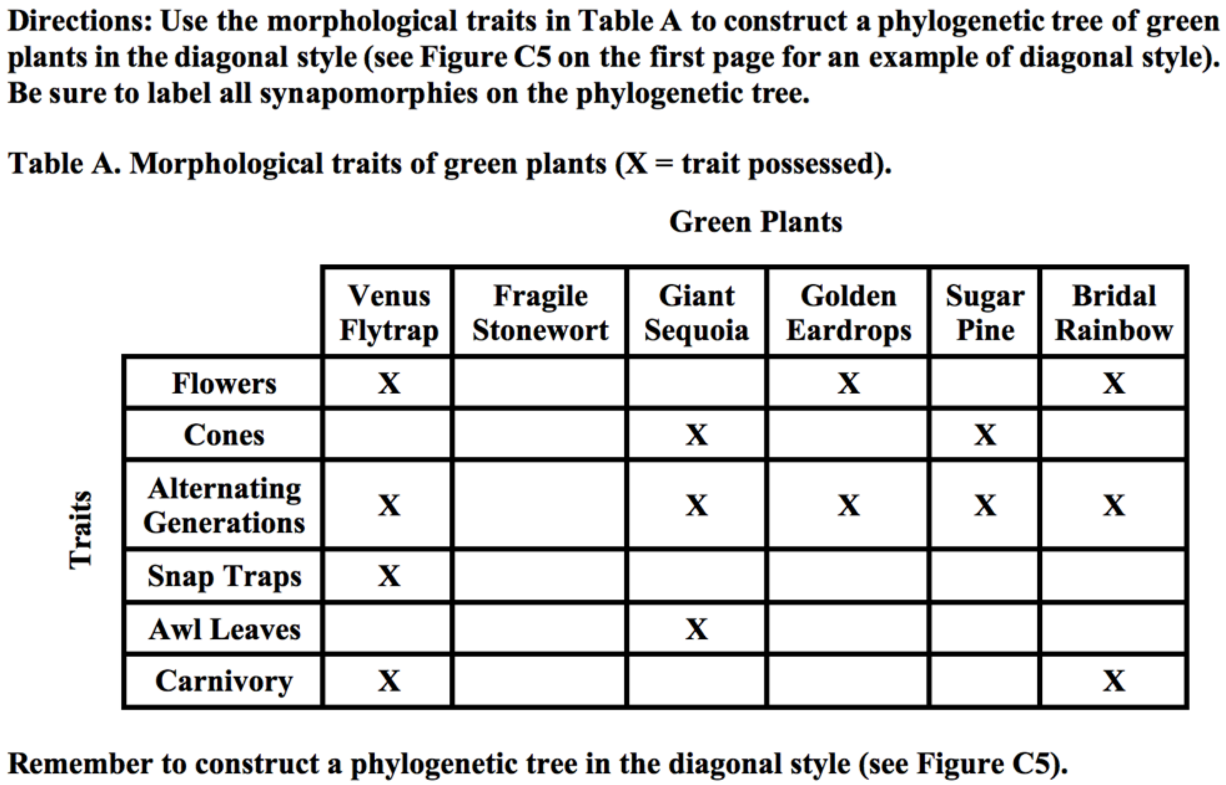
____ Painted turtles are more closely related to merlins than green anacondas.

____ Painted turtles are equally related to green anacondas and merlins.

____ Painted turtles are more closely related to green anacondas than merlins.

Explain the reasoning for your choice.

Construction Task



Directions: Use the morphological traits in Table A to construct a phylogenetic tree of green plants in the diagonal style (see Figure C5 on the first page for an example of diagonal style). Be sure to label all synapomorphies on the phylogenetic tree.

Table A. Morphological traits of green plants (X = trait possessed).

Green Plants

Traits	Venus Flytrap	Fragile Stonewort	Giant Sequoia	Golden Eardrops	Sugar Pine	Bridal Rainbow
Flowers	X			X		X
Cones			X		X	
Alternating Generations	X		X	X	X	X
Snap Traps	X					
Awl Leaves			X			
Carnivory	X					X

Remember to construct a phylogenetic tree in the diagonal style (see Figure C5).

Figure C6. Construction task from the post-instructional homework (diagonal style).

Instrument #2: Post-Instructional Homework (Bracket Style)

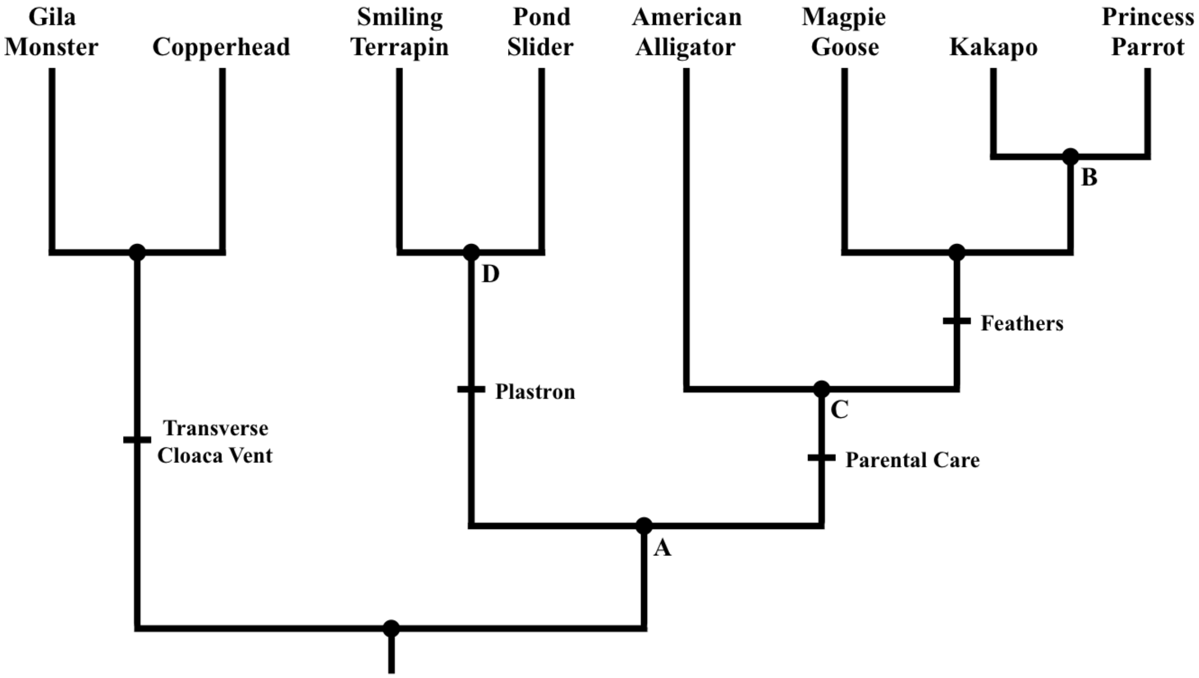


Figure C7. Phylogenetic tree from the post-instructional homework (bracket style).

Most Recent Common Ancestor

Which node on the phylogenetic tree in Figure C7 represents the most recent common ancestor of pond sliders and American alligators? Select one option.

____ node A ____ node B ____ node C ____ node D

Monophyletic Group

On the phylogenetic tree in Figure C7, draw a circle around one monophyletic group that includes at least two species.

Contemporary Descent

According to the phylogenetic tree of reptiles and birds in Figure C7, did pond sliders evolve from smiling terrapins? Select one option and explain the reasoning for your choice.

____ Yes ____ No

Taxa Relatedness

According to the phylogenetic tree of reptiles and birds in Figure C7, how would you describe the relatedness of smiling terrapins to copperheads and kakapos? Select one option.

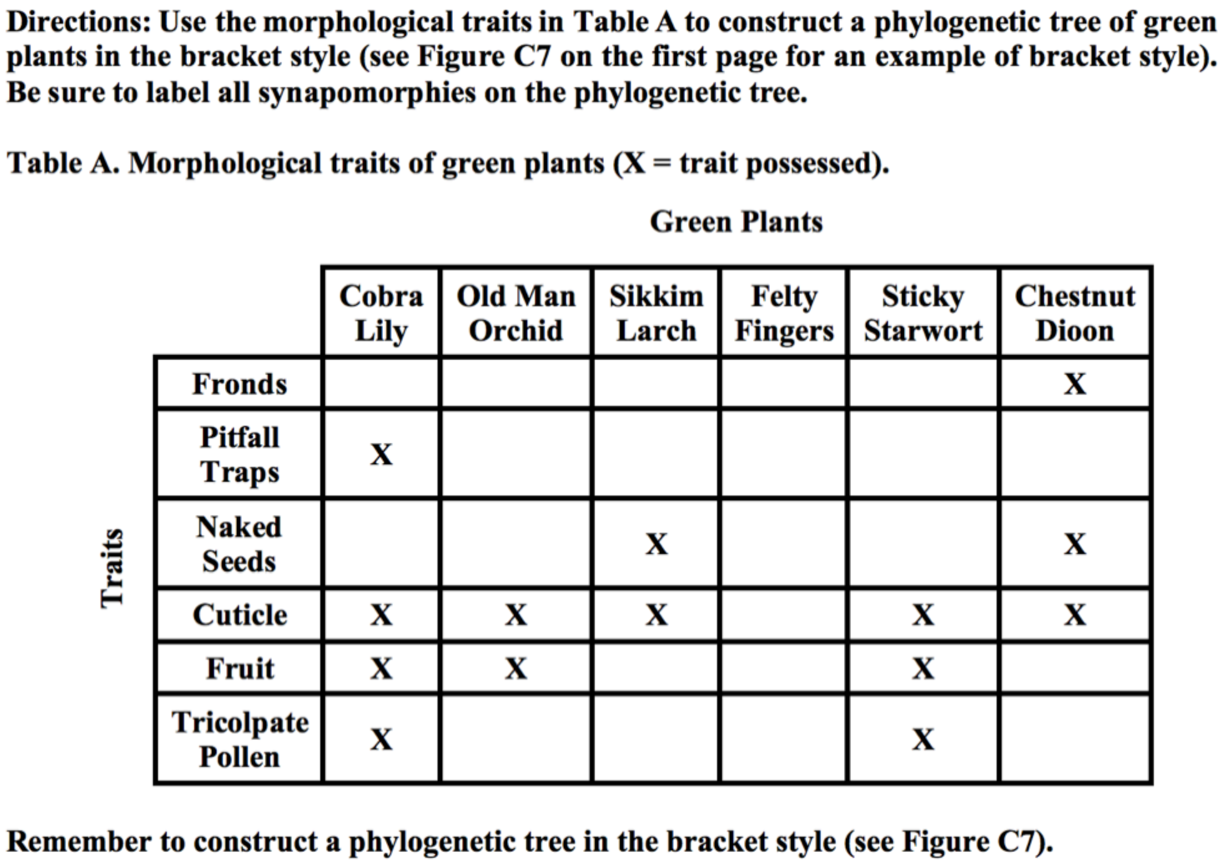
____ Smiling terrapins are more closely related to copperheads than kakapos.

____ Smiling terrapins are equally related to copperheads and kakapos.

____ Smiling terrapins are more closely related to kakapos than copperheads.

Explain the reasoning for your choice.

Construction Task



Directions: Use the morphological traits in Table A to construct a phylogenetic tree of green plants in the bracket style (see Figure C7 on the first page for an example of bracket style). Be sure to label all synapomorphies on the phylogenetic tree.

Table A. Morphological traits of green plants (X = trait possessed).

		Green Plants					
		Cobra Lily	**Old Man Orchid**	**Sikkim Larch**	**Felty Fingers**	**Sticky Starwort**	**Chestnut Dioon**
Traits	**Fronds**						X
	Pitfall Traps	X					
	Naked Seeds			X			X
	Cuticle	X	X	X		X	X
	Fruit	X	X			X	
	Tricolpate Pollen	X				X	

Remember to construct a phylogenetic tree in the bracket style (see Figure C7).

Figure C8. Construction task from the post-instructional homework (bracket style).

Instrument #3: Unit Exam (Diagonal Style)

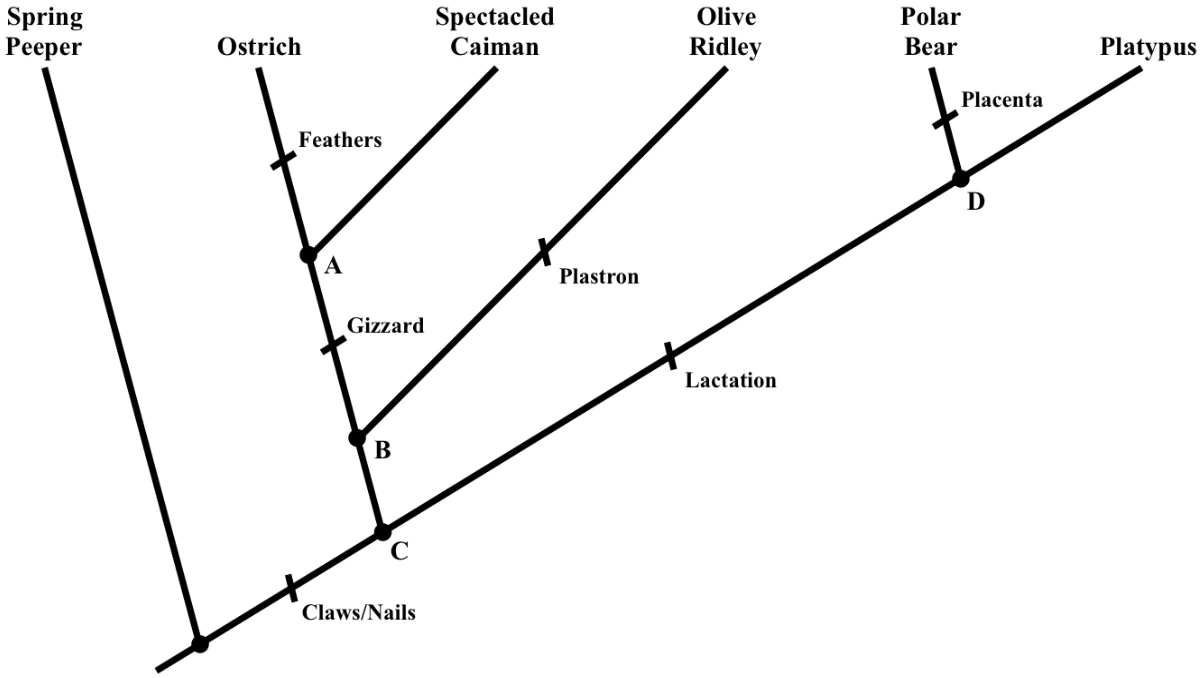


Figure C9. Phylogenetic tree from the unit exam (diagonal style).

Most Recent Common Ancestor

Which node on the phylogenetic tree in Figure C9 represents the most recent common ancestor of spectacled caimans and polar bears? Select one option.

____ node A ____ node B ____ node C ____ node D

Monophyletic Group

On the phylogenetic tree in Figure C9, draw a circle around one monophyletic group that includes at least two species.

Contemporary Descent

According to the phylogenetic tree of tetrapods in Figure C9, did olive ridleys evolve from ostriches? Select one option and explain the reasoning for your choice.

____ Yes ____ No

Taxa Relatedness

According to the phylogenetic tree of tetrapods in Figure C9, how would you describe the relatedness of polar bears to ostriches and olive ridleys? Select one option.

____ Polar bears are more closely related to ostriches than olive ridleys.

____ Polar bears are equally related to ostriches and olive ridleys.

____ Polar bears are more closely related to olive ridleys than ostriches.

Explain the reasoning for your choice.

Instrument #3: Unit Exam (Bracket Style)

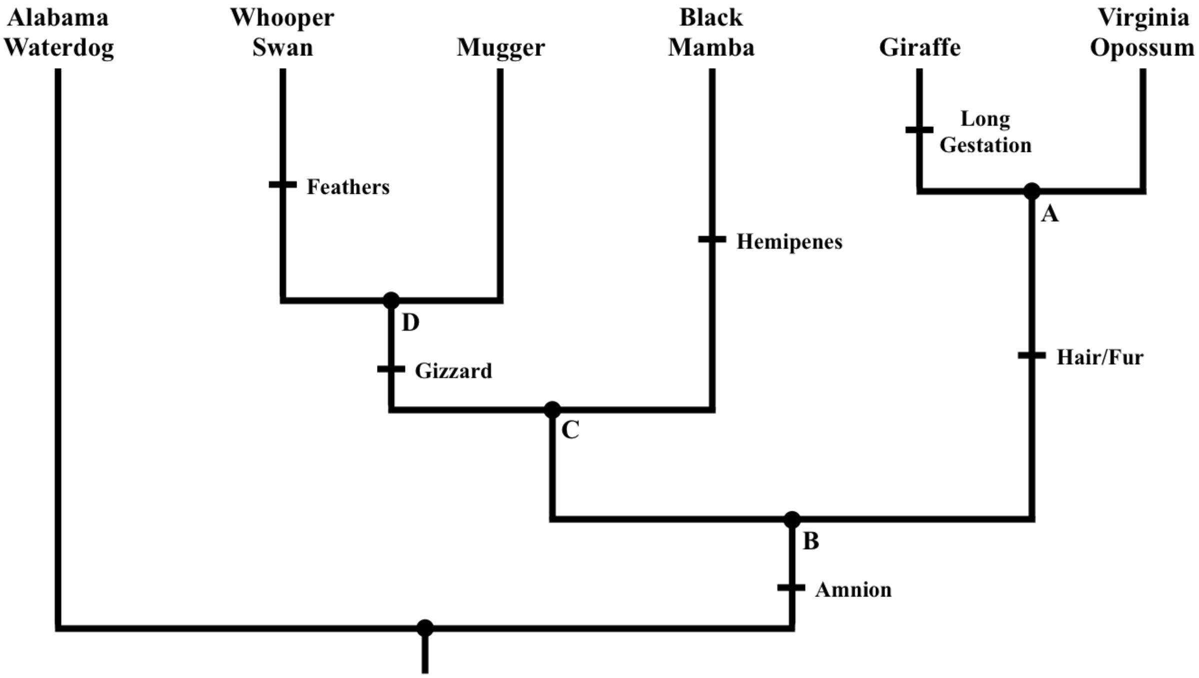


Figure C10. Phylogenetic tree from the unit exam (bracket style).

Most Recent Common Ancestor

Which node on the phylogenetic tree in Figure C10 represents the most recent common ancestor of muggers and giraffes? Select one option.

____ node A ____ node B ____ node C ____ node D

Monophyletic Group

On the phylogenetic tree in Figure C10, draw a circle around one monophyletic group that includes at least two species.

Contemporary Descent

According to the phylogenetic tree of tetrapods in Figure C10, did black mambas evolve from whooper swans? Select one option and explain the reasoning for your choice.

____ Yes ____ No

Taxa Relatedness

According to the phylogenetic tree of tetrapods in Figure C10, how would you describe the relatedness of giraffes to whooper swans and black mambas? Select one option.

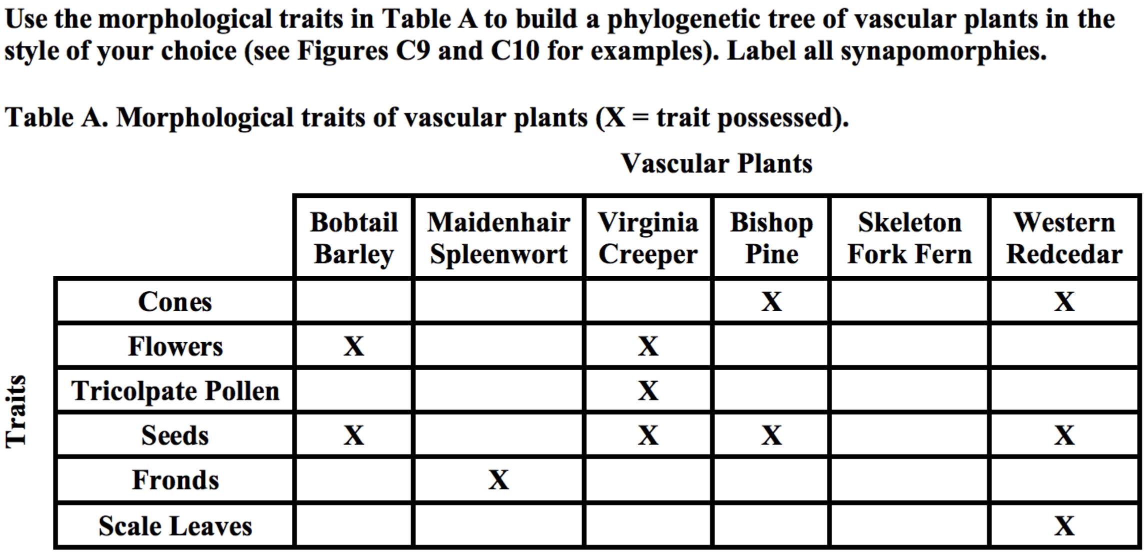
____ Giraffes are more closely related to black mambas than whooper swans.

____ Giraffes are equally related to whooper swans and black mambas.

____ Giraffes are more closely related to whooper swans than black mambas.

Explain the reasoning for your choice.

Construction Task (Optional Style)



Use the morphological traits in Table A to build a phylogenetic tree of vascular plants in the style of your choice (see Figures C9 and C10 for examples). Label all synapomorphies.

Table A. Morphological traits of vascular plants (X = trait possessed).

	Vascular Plants					
Traits	**Bobtail Barley**	**Maidenhair Spleenwort**	**Virginia Creeper**	**Bishop Pine**	**Skeleton Fork Fern**	**Western Redcedar**
Cones				X		X
Flowers	X		X			
Tricolpate Pollen			X			
Seeds	X		X	X		X
Fronds		X				
Scale Leaves						X

Figure C11. Construction task from the unit exam (optional style).

Instrument #4: Final Exam Review Activity (Diagonal Style)

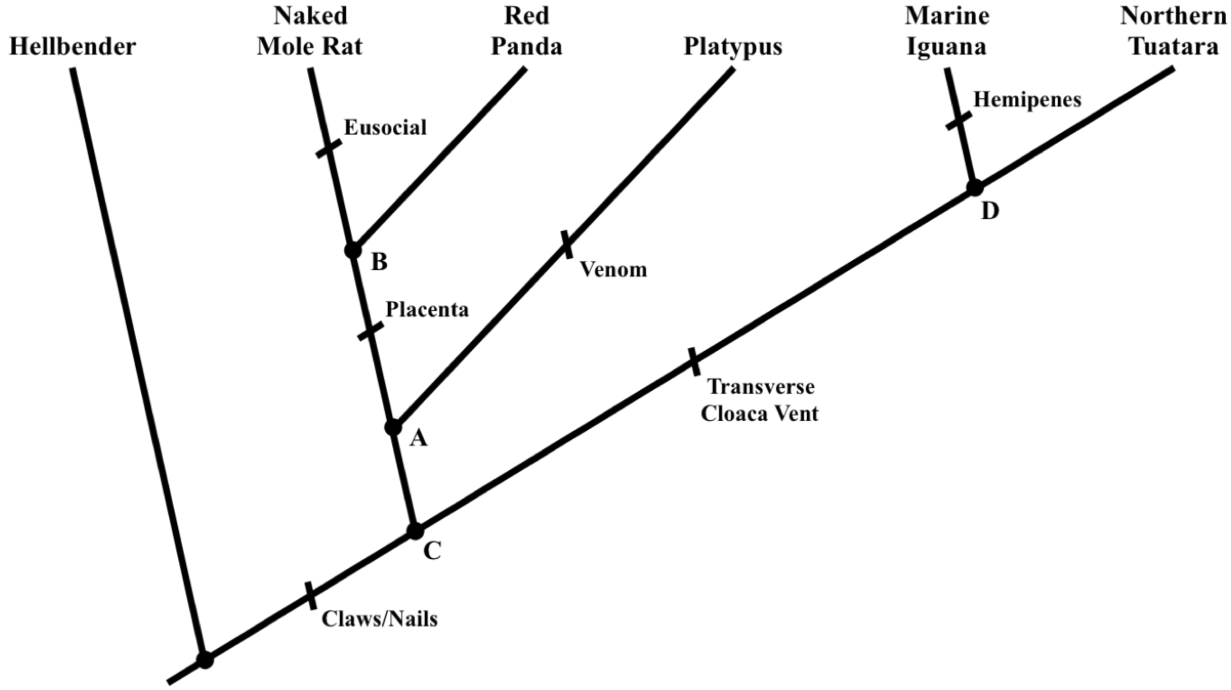


Figure C12. Phylogenetic tree from the final exam review activity (diagonal style).

Most Recent Common Ancestor

Which node on the phylogenetic tree in Figure C12 represents the most recent common ancestor of red pandas and marine iguanas? Select one option.

____ node A ____ node B ____ node C ____ node D

Monophyletic Group

On the phylogenetic tree in Figure C12, draw a circle around one monophyletic group that includes at least two species.

Contemporary Descent

According to the phylogenetic tree of tetrapods in Figure C12, did platypuses evolve from naked mole rats? Select one option and explain the reasoning for your choice.

____ Yes ____ No

Taxa Relatedness

According to the phylogenetic tree of tetrapods in Figure C12, how would you describe the relatedness of marine iguanas to naked mole rats and platypuses? Select one option.

____ Marine iguanas are more closely related to naked mole rats than platypuses.

____ Marine iguanas are equally related to naked mole rats and platypuses.

____ Marine iguanas are more closely related to platypuses than naked mole rats.

Explain the reasoning for your choice.

Instrument #4: Final Exam Review Activity (Bracket Style)

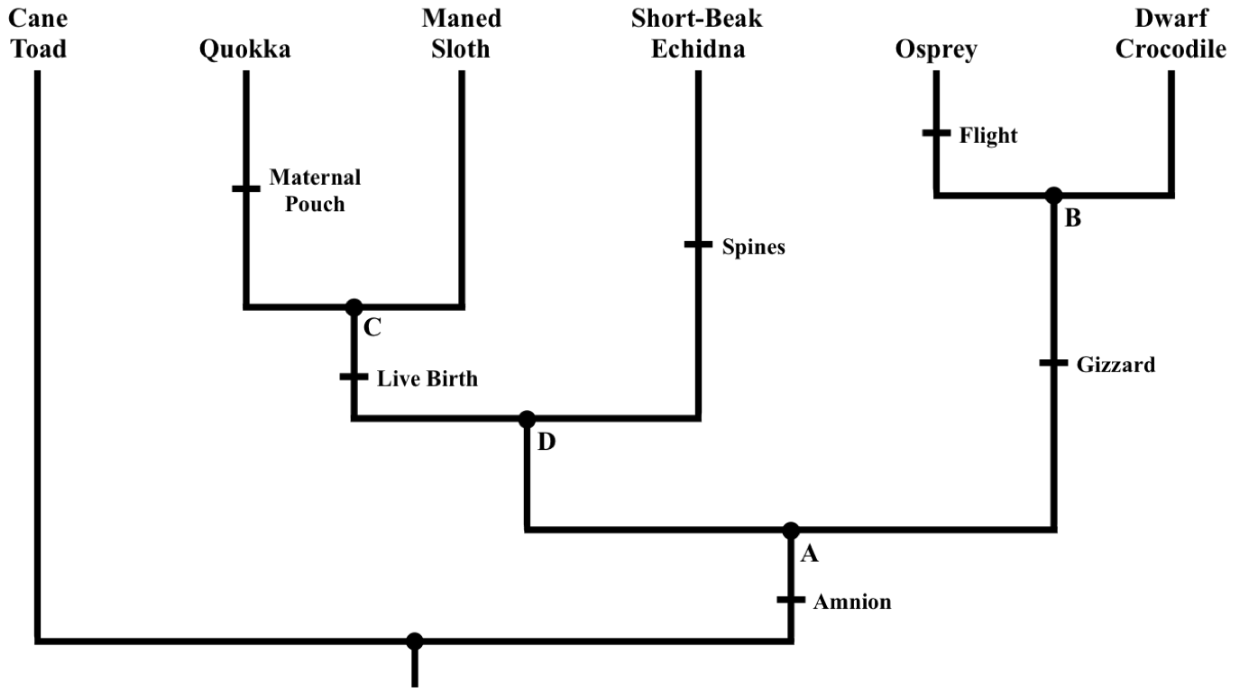


Figure C13. Phylogenetic tree from the final exam review activity (bracket style).

Most Recent Common Ancestor

Which node on the phylogenetic tree in Figure C13 represents the most recent common ancestor of maned sloths and ospreys? Select one option.

____ node A ____ node B ____ node C ____ node D

Monophyletic Group

On the phylogenetic tree in Figure C13, draw a circle around one monophyletic group that includes at least two species.

Contemporary Descent

According to the phylogenetic tree of tetrapods in Figure C13, did short-beak echidnas evolve from quokkas? Select one option and explain the reasoning for your choice.

____ Yes ____ No

Taxa Relatedness

According to the phylogenetic tree of tetrapods in Figure C13, how would you describe the relatedness of ospreys to quokkas and short-beak echidnas? Select one option.

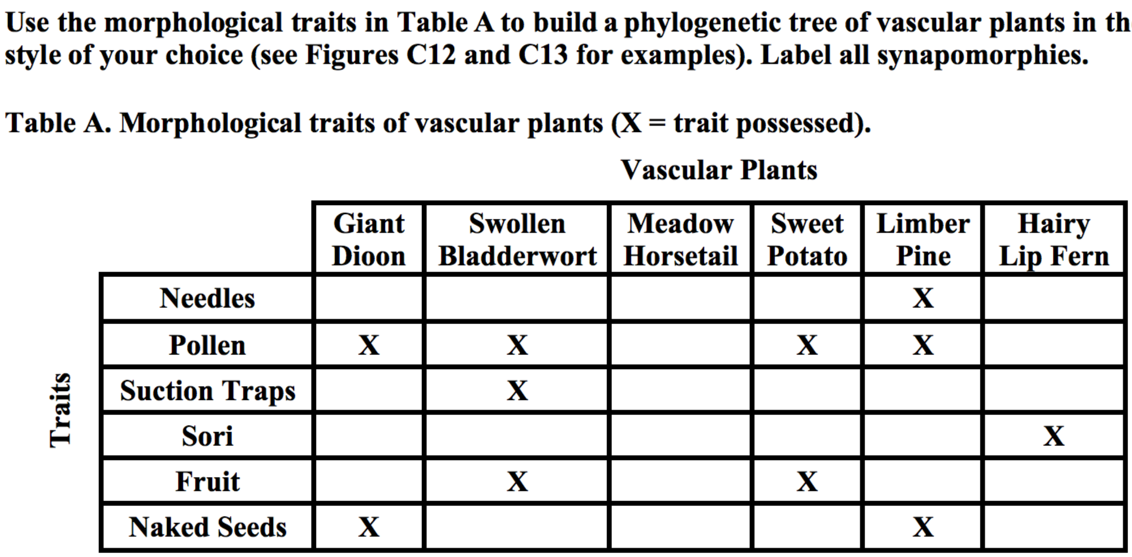
____ Ospreys are more closely related to short-beak echidnas than quokkas.

____ Ospreys are equally related to quokkas and short-beak echidnas.

____ Ospreys are more closely related to quokkas than short-beak echidnas.

Explain the reasoning for your choice.

Construction Task (Optional Style)



Use the morphological traits in Table A to build a phylogenetic tree of vascular plants in the style of your choice (see Figures C12 and C13 for examples). Label all synapomorphies.

Table A. Morphological traits of vascular plants (X = trait possessed).

Vascular Plants

Traits	Giant Dioon	Swollen Bladderwort	Meadow Horsetail	Sweet Potato	Limber Pine	Hairy Lip Fern
Needles					X	
Pollen	X	X		X	X	
Suction Traps		X				
Sori						X
Fruit		X		X		
Naked Seeds	X				X	

Figure C14. Construction task from the final exam review activity (optional style).