HOW UNDERGRADUATE STUDENTS MISREAD CLADOGRAMS

KAREN M. KORTZ[1], JOHN P. SWADDLE[2], AND DAVID E. FASTOVSKY[3]

[1]Community College of Rhode Island, Physics Department,
1762 Louisquisset Pike, Lincoln, RI 02865 USA
kkortz@ccri.edu

[2]Institute for Integrative Bird Behavior Studies, College of William and Mary,
Landrum Drive, Williamsburg, VA 23187 USA

[3]Department of Geosciences, University of Rhode Island,
9 East Alumni Ave., Kingston, RI 02881 USA

ABSTRACT.—Although phylogenetic systematics is used to reconstruct evolutionary relationships, undergraduates have a difficult time mastering its fundamental concepts. Because it is a key part of the mainstream professional thinking, we explored in what ways students misread cladograms, which are the abstract and synthetic diagrams of phylogenetic systematics. We developed a questionnaire to examine the following four hypotheses as to how introductory college-level students (n=51) read cladograms: 1) students read cladograms correctly; 2) students infer that proximity of tips equals relatedness; 3) students read cladograms as they might an evolutionary tree, reading left to right as primitive to more advanced, and perceiving organisms as branching off; and 4) students infer ancestors at the nodes. Most responses fell into one of the four hypotheses, with 55% following the scientific ('correct') hypothesis. Most students answered between six and eight of the ten questions correctly. Slightly more than half of the students generally followed the scientific hypothesis, while others applied both the scientific and proximity (hypothesis 2, above) hypotheses together. A few students followed the primitive hypothesis (hypothesis 3, above). Our recommendation is that instructors address discrepancies between the scientific and proximity hypotheses in particular. For undergraduates, generally, cladograms require focused teaching, explanation, and active-learning approaches to be successfully used to teach phylogenetic systematics.

INTRODUCTION

THE ADVENT of using phylogenetic systematic methods in the late 1970s in systematic biology radically altered perceptions of the relationships of organisms, particularly vertebrates. For the first time, it provided a testable means by which the phylogenetic relationships of organisms could be assessed, and it imposed an evolutionary classification upon the Linnaean framework (which is an approach rooted in a pre-evolutionary, static-world view) (Hennig, 1979; Eldredge and Cracraft, 1980; Nelson and Platnick, 1981; Wiley et al., 1991). The importance of phylogenetic systematics became obvious in the succeeding 20 years, and it is fair to say that its language and concepts are now ubiquitous among professionals in evolutionary biology, including paleobiology. For all its importance, however, phylogenetic methods as well as the ramifications for organismic relationships have only just begun to find their way to introductory college-level pedagogy. But, because introducing students to mainstream scientific thinking is a goal generally shared by introductory science courses, and few would argue for teaching antiquated phylogenies and phylogenetic methods, and because discussions of evolution ought to include the means by which evolutionary relationships are understood, it is appropriate that students be exposed to phylogenetic systematics at introductory levels in college science curricula.

How does phylogenetic systematics function? Most simply put, it is a system of reconstructing phylogeny based upon hierarchical distributions of shared, derived characters (synapomorphies). The distribution of the characters is commonly represented on a dichotomous branching diagram

In *Teaching Paleontology in the 21st Century, The Paleontological Society Special Publications, Volume 12, Margaret M. Yacobucci, and Rowan Lockwood (eds.).*

TABLE *1*.—A cladogram and a tree-of-life compared.

Characteristic	Cladogram	Tree-of-Life
Branching	dichotomous	not necessarily
Shows ancestor-descendent relationships	no	yes
Shows character distributions	yes	no
Shows time	shows sequence	yes
Nature of the hierarchies	Characters	taxa
Testable	yes	no

known as a cladogram, a graphic format that bears, from a pedagogical standpoint, an unfortunate resemblance to a conventional "tree of life." Cladograms, however, differ fundamentally from trees, and Table 1 highlights a few of the salient differences. Unlike trees, cladograms are constructed within a Popperian hypothetico-deductive framework, and are ultimately testable hypotheses of relationship; generally, parsimony is used to distinguish among the competing hypotheses (Table 1). Cladograms show hierarchical distributions of characters—in effect, evolutionary novelties that diagnose groups whose members, on the basis of these characters, are inferred to be more closely related to each other than to any other organisms. As the purpose of this paper is not to review phylogenetic systematic methods, readers interested in the methods are referred to basic treatments such as Cracraft (1979), Hennig (1979), Eldredge and Cracraft (1980), Nelson and Platnick (1981), Wiley et al. (1991), de Queiroz and Gauthier (1992, 1994), and Foote and Miller (2007).

Phylogenetic systematics presents significant pedagogical challenges. It is rooted in comparatively abstract reasoning and requires logical inferences that are not accessible to many undergraduate students (McConnell et al., 2005). Moreover, as noted above, cladograms are branching diagrams and look like trees, even if the methods of interpreting each are quite different. Finally, the conclusions reached via phylogenetic systematics can be counterintuitive and, significantly, contradict much that students have been taught since early grade school. For example, phylogenetic systematics approaches unequivocally demonstrate that birds and crocodiles are more closely related to each other than either is to a lizard (see below). This view, of course, undermines the term 'reptile' as it was first proposed by Linnaeus and, more significantly, as it is understood by most students. In an evolutionary sense, a classification of Reptilia that includes lizards and crocodiles must also include birds; but it does not. Just within Vertebrata, other fundamental relationships and groups are similarly susceptible to revision: if reptiles are distinguished by the evolution of amnion (e.g., Romer, 1966), then mammals ought to be reptiles as well. Phylogenetically, birds are dinosaurs, which means that dinosaurs did not really go extinct (Dingus and Rowe, 1998). Indeed, students tend to misread the fundamental diagram of phylogenetics systematics—the cladogram—and instead interpret it as the more familiar and iconic evolutionary tree (see below).

The purpose of this paper is to explore the process of learning phylogenetic systematics, and to attempt to tease out some of the sticking points in the learning process. The data come from a large, general-education course on dinosaur paleobiology. Despite consistent emphasis placed in the text, lecture, and examinations on the use of phylogenetic systematics in dinosaur paleobiology, students had difficulties grasping the concepts, even by the end of the course. Based on exam scores and the instructor's perceptions, the course was not as effective as desired with respect to a meaningful understanding of cladistic methods and outcomes. This study seeks to explore whether one of the reasons students have difficulties is because they do not read cladograms correctly. The hope is that this work will serve as a model for gaining insights into teaching complex and/or unfamiliar subjects, as well as—and experience suggests that this is the most challenging—helping students to unlearn preexisting misconceptions, substituting more appropriate interpretations in their stead.

Reading cladograms

In a cladogram, organisms are arranged at the tips of the tree (Fig. 1). Everything above a particular node (shown by open circles) is inferred to 1) be more closely related to other organisms above that node than to anything else; and 2) share a most-recent common ancestor (e.g., the clade Dinosauria contains *Stegosaurus*, *Tyranno-*

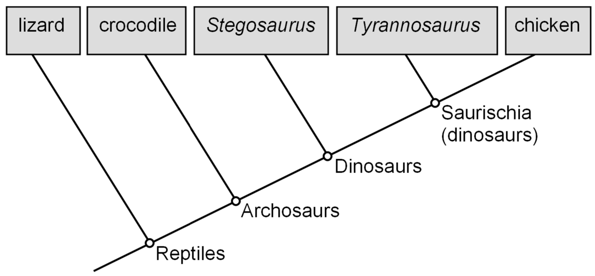


FIGURE *1*.—A cladogram representing the relationships of five taxa: *Stegosaurus*, *Tyrannosaurus*, a crocodile, a lizard, and a chicken; organisms are arranged at the tips of the branches, and clades are groups of organisms above a node (open circle) that share a most recent common ancestor; organisms within a clade are more closely related to each other than to organisms outside of that clade.

saurus, and a modern chicken, all sharing a most-recent common ancestor). The order of the branches above each node can be rearranged as long as the underlying structure depicted by the nodes remains the same.

Cladograms are synthetic diagrams that are often misinterpreted by novices because they have been socialized to read the branching of the cladogram as if it were an evolutionary tree (O'Hara, 1994, 1998; Baum et al., 2005). Deeply rooted alternative conceptions (or misconceptions) are exceptionally difficult to change, especially without direct instruction addressing the misconception (e.g., Posner et al., 1982; Hewson and Hewson, 1988; Chan et al., 1997; Guzzetti, 2000). Students reading the cladogram as if it were a 'tree,' along the tips of the branches from left to right, reinforce the long-standing perception that evolution is a linear progression from 'primitive' to 'advanced' species (e.g., O'Hara, 1988; Gould, 1989; Rudolf and Stewart, 1998; Crisp and Cook, 2005; Nee, 2005; Meisel, 2010). In Figure 1, reading along the tips might lead students to believe that 'primitive' lizards directly lead to 'advanced' chickens. In addition, when students see a long, straight, unbroken line (such as leading towards the lizard in Figure 1), they interpret it to not have changed, again leading to the belief that the organism at the tip is more primitive (in our example, the lizard) (Meir and Perry, 2007; Novick and Catley, 2007; Gregory, 2008). Another underlying factor that leads to difficulties with interpreting some organisms as more advanced is that when students see an organism with more branches leading towards it (such as the chicken in Figure 1), they view that organism as more advanced (Gregory, 2008).

Since students often do not realize that clades show evolutionary similarities, they may infer closeness of relationships between organisms that are near each other at the tips, but whose relationship is not all that close (Baum et al., 2005; Meir and Perry, 2007; Gregory, 2008). For example, in Figure 1, students may erroneously conclude that *Stegosaurus* is more closely related to the crocodile than to the chicken, instead of looking at the organization of the clades to infer relatedness. Students who read cladograms in this way (proximity of the tips indicates relatedness) fail to understand that clades can be rotated about the nodes without changing the evolutionary meaning (Meir and Perry, 2007; Gregory, 2008).
Also, living species may be mistakenly projected to occupy the internal nodes of a tree, similar to evolutionary trees, leading to the incorrect conclusion that chickens descended from crocodiles (Crisp and Cook, 2005). The correct interpretation is that chickens and crocodiles share a common ancestor, about which we can infer characteristics, although it is no longer living.

Based on the common misconceptions described above and our experiences in class, we developed the following research questions: 1) do students read cladograms correctly; and 2) if not, what are common ways in which they misread cladograms? Also guided by the misconceptions described above, we developed four hypotheses on how students read cladograms: 1) students read cladograms correctly (we term this the 'scientific' hypothesis); 2) students infer that proximity of tips equals relatedness ('proximity' hypothesis); 3) students read primitive to more advanced from left to right along the cladogram ('primitive' hypothesis); 4) students infer the existence of ancestors at the nodes ('nodes' hypothesis). We propose that, if students did not read cladograms correctly, they would apply one and/or a combination of these misconceptions (i.e. proximity, primitive, or nodes) when answering questions throughout the questionnaire.

METHODS

Study population

Students in this study were enrolled in the Evolution and Extinction of the Dinosaurs (EED), a large (~85 students), lecture-based, undergradu-

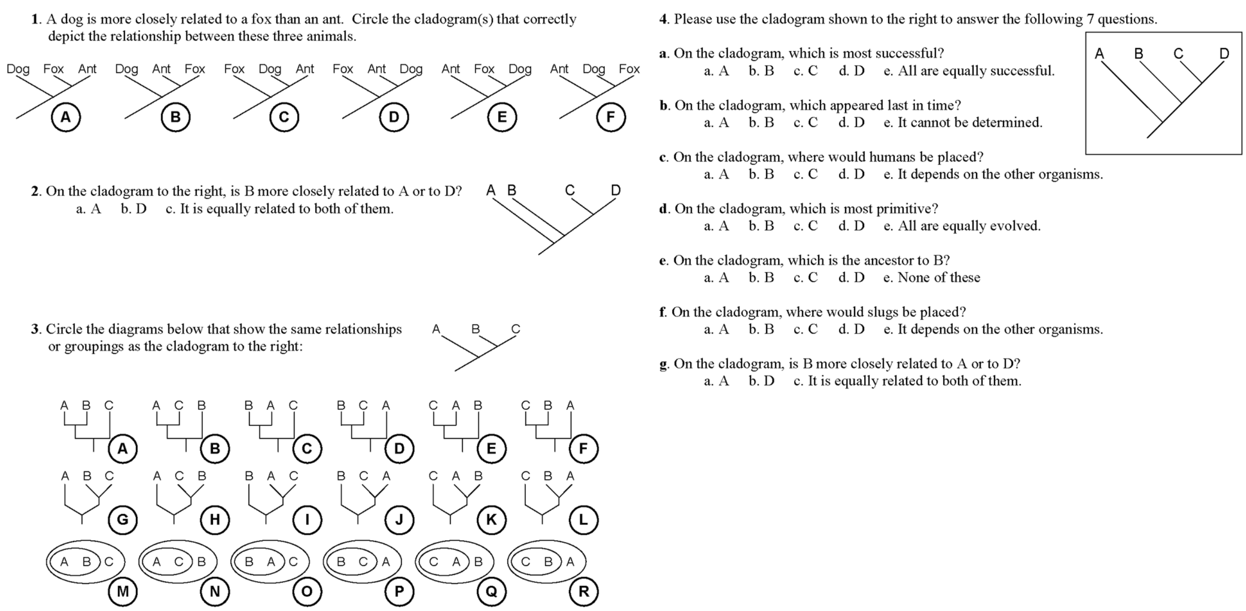


FIGURE 2.—Questionnaire used in this study; the letters in circles were not on the original questionnaire but are instead placed there to simplify discussion.

ate, introductory, general-education course offered at a state university in New England and taught in Spring 2010 by Fastovsky. The students mostly were taking the class because of interest, or to fulfill a general-education science requirement. The textbook used was Fastovsky and Weishampel (2009), the only cladistically based general-education dinosaur textbook on the market. Fifty-one students completed the questionnaire in this study, a number attributable to a combination of poor attendance and several students not completing the questionnaire. While no data exist regarding class attendance and student performance in EED, it is our qualitative impression that missed classes translate to lower grades in EED, an impression reinforced by the fact that much of what is taught and tested is not readily available from sources other than the lecture. If true, this means that the performance of the student sample upon which this study is based likely is above the class mean, highlighting the seriousness of the problem treated in this article.

The students were 33% female and 67% male, mostly early in their college education (56% first years and 24% second years), and mostly non-science majors (no geology majors, 6% biology majors, 19% other science majors, 4% education majors, and 70% other major) as declared at the beginning of the semester. Ninety-one percent of students took a biology course in high school, and 64% took a high-school earth science or geology course. In their college course work, 15% of the students had previously taken introductory geology, 4% historical geology, 19% evolution, and 19% an upper-division biology course. Although ethnicity was not asked of the students in this class, student demographics for the university as a whole are: 72.0% Caucasian, 5.1% Black, 6.4% Hispanic, 2.8% Asian, 1.0% were other ethnicities, and 12.7% did not disclose their ethnicity.

The questionnaire used in this study was distributed near the end of the semester, thirteen weeks after cladograms were first introduced, and subsequently continuously used to explain dinosaur relationships. Therefore, prior to filling out the questionnaire, students were exposed to cladograms in multiple settings during lecture and were asked to use them to answer questions on exams. Permission was attained from the Institutional Review Board (IRB) prior to distribution of the questionnaires.

Questionnaire design

We designed a ten-question, fixed-response questionnaire to assess how frequently students applied the four preconceived hypotheses (scientific, proximity, primitive, nodes) to interpret cladograms (Figure 2). Although fixed-response formats do not produce as rich responses as an open-response instrument, this format fulfilled our needs because our goal was to examine the occurrence and combination of specific misconceptions among the students tested.

To address the questionnaire's validity, the cladograms and wording were kept simple to ensure that they were addressing students' basic or fundamental cladogram-reading skills. Students were not required to know organismic relationships in order to successfully answer the questions, and language was used that was similar to language used in class. For most questions, we used letters instead of organisms at the tips so that student preconceptions of organismal relationships did not interfere with their reading of the cladograms (Gregory, 2008; Morabito et al., 2010; Bair, 2011; Novick et al., 2011). Based on feedback, future versions of the questionnaire would require some refining of wording to ensure we are testing student conceptions accurately.

Multiple questions were written to probe each hypothesis, which helped to address the questionnaire's internal reliability. The multiple questions allow comparison of student responses to questions addressing similar topics from different perspectives.

We made predictions of student responses based on our preconceived hypotheses of misconceptions (Table 2), and used these predictions during our analysis of student responses. Predictions for the proximity hypothesis expected students to identify relationships based on how close the tips are to each other, ignoring the relationships depicted by the cladogram. For example, in Questions 2 and 4g, we predicted students would respond that B is more closely related to A because they are next to each other. Predictions for the primitive hypothesis expected students to view the tips of the cladogram as a progression from more primitive to more advanced organisms. For example, we predicted that students would respond 'D' in Questions 4a and 4b, which asked which is most successful or appeared last in time. Predictions for the nodes hypothesis expected students to view organisms as being located in the nodes, such as in Question 4e, where students would answer that A is the ancestor to B.

TABLE *2*.—Predictions of responses to each question for each hypothesis.[1]See text for a description of the four hypotheses; scientific is "correct." [2]There is no prediction for this question based on the hypothesis.

Question	Scientific[1]	Proximity	Primitive	Nodes
1	EF	ACEF or AC	E	-[2]
2	B	A	-	-
3	DFGHPR	AGM or A+C, K+L, M+O	-	-
4a	E	-	D	-
4b	E	-	D or A	-
4c	E	-	D	-
4d	E	-	A	-
4e	E	-	-	A
4f	E	-	A	-
4g	B	A	-	-

Statistical analyses

To explore whether students gave answers to questions that were consistent with the same hypothesis (i.e. scientific, proximity, primitive, nodes) throughout the questionnaire, we employed two statistical approaches. First, for each hypothesis in turn, we identified whether a student's answer to the first question influenced which hypothesis they appeared to follow on subsequent questions. These comparisons were made by employing a Mann-Whitney U-test, where the two-level factor was whether the student answered the first question according to a particular hypothesis (e.g., the scientific hypothesis) or not, and the dependent variable was the number of subsequent answers that were consistent with this same hypothesis. Separate analyses were conducted for three hypotheses (scientific, proximity, primitive); we could not run this analysis for the nodes hypothesis as this was not probed by Question 1. Integral to these and subsequent statistical analyses was that we coded answers to each question according to their fit for each hypothesis (1 = scientific; 2 = proximity; 3 = primitive; 4 = nodes; 5 = other). Hence, every answer given by every student to all 10 questions was coded with these categorical responses.

With our second statistical approach, we investigated whether students could be grouped together by how they followed particular hypotheses in answering questions. To do this, we conducted a Principal Coordinates Analysis (PCO) on the coded answers to the 10 questions, where the numerical code for each hypothesis was treated as a categorical variable and the similarity matrix was based on Euclidian distances. The PCO

TABLE *3*.—Eigenvalues and percent of variance explained from PCO of categorized (by hypothesis) answers to the 10 questions in the survey.

Coordinate	Eigenvalue	Percent of variance explained
1	292.51	33.25
2	160.60	18.25
3	115.96	13.18
4	105.83	12.03
5	58.47	6.65
6	49.48	5.62
7	35.78	4.07
8	27.27	3.10
9	19.15	2.18
10	14.80	1.68

analysis generates indices, termed 'coordinates,' which represent the overall similarity of an individual student's responses to the questions. Students who answer questions similarly will have similar coordinate values. Hence, the output from a PCO can be used to generate clusters of students who answer questions in similar ways, indicating if there are general patterns of misconceptions among the students. It is necessary to answer questions about common misconceptions in a multivariate framework as a student's answer to any one question is not independent of how they answer other questions, and a student's possible misconceptions are not independent of each other (e.g., a student who does not fully comprehend and apply the scientific hypothesis may rely on a combination of other hypotheses to answer questions). Hence, the multivariate PCO (and subsequent cluster) analysis is an important step in examining whether there are multi-hypothesis patterns in how students responded to the questionnaire. PCO is a particularly useful analytic tool because it deals with categorical data in ways that traditional multivariate methods, such as Principal Components Analysis, cannot (Gower 1966), yet can generate linear metrics that are suitable for cluster analyses. The PCO analysis generated 10 coordinates with an eigenvalue greater than 1 (Table 3). We used the coordinate values (i.e. continuous metrics) for each student in a subsequent cluster analysis (using Euclidian distance to generate the similarity measure) to analyze whether students could be clustered successfully according to similarity in their questionnaire answers. All analyses were performed using PAST version 2.08b statistical software (Hammer et al., 2001) employing two-tailed tests of probability.

RESULTS

Overall, students answered the questions correctly 55% of the time (i.e. following the scientific hypothesis), although this percentage varies widely by question (Table 4). Of the 51 students who completed the questionnaire, two students (4%) answered all questions correctly, no students answered nine of ten questions correctly, two students (4%) answered all questions incorrectly, and one student answered all questions but one incorrectly. Most students answered six, seven, or eight questions correctly.

Most student responses fell into one of the four hypotheses (Table 4), although there were some outliers for some of the questions. Questions 1, 2, and 4g had a significant number of responses that are not directly explained by one of

TABLE *4*.—Percentages of responses to each question for each hypothesis.[1] See text for a description of the four hypotheses; scientific is "correct;" [2] Responses that do not fit into a hypothesis; [3] There is no prediction for this question based on the hypothesis; [4] Three students left this question blank; [5] Two students left this question blank; [6] One student left this question blank.

Question	Scientific[1]	Proximity	Primitive	Nodes	Other[2]
1	43	26	8	-[3]	24
2	24	53	-	-	24
3	61	24	-	-	16
4a[4]	67	-	27	-	0
4b	69	-	31	-	0
4c[5]	86	-	10	-	0
4d[6]	16	-	78	-	4
4e	59	-	-	37	4
4f	94	-	2	-	4
4g	31	49	-	-	20

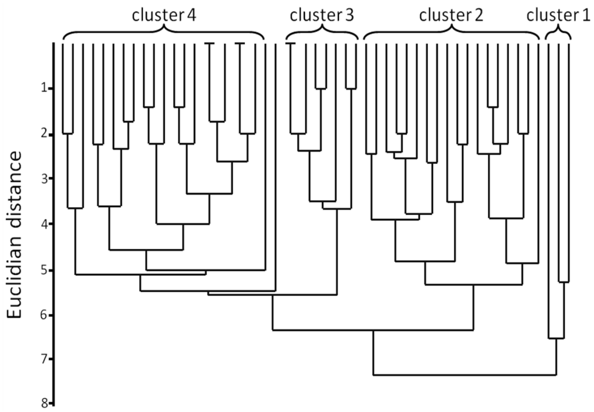


FIGURE *3*.—Output from cluster analysis (employing Euclidian distance to assess similarity) of PCO scores; every line represents a single student (N = 51); we interpreted four clusters, labeling them from right to left because of their differentiation in the distance diagram; Euclidian distance among the groups was at least twice the average internode distance within groups; one individual lay between the third and fourth clusters.

the hypotheses; hence, these responses were categorized as following an 'other' hypothesis. In Question 1, most 'other' hypothesis responses are a combination of A, C, E, and F that are not directly predicted in the hypotheses (e.g., 3 students (6%) answered A, E, F). The 'other' responses for Questions 2 and 4g are that B is equally related to A and D.

Students who gave answers to Question 1 that were consistent with the scientific hypothesis were more likely to answer subsequent questions following the same hypothesis ($U_{22,29}$ = 189.5, P = 0.013). This pattern was similar but much weaker (due to a small sample size) for students who consistently answered using the primitive hypothesis ($U_{5,46}$ = 68, P = 0.116). However, students answering Question 1 according to the proximity hypothesis were no more likely to answer other questions using this hypothesis than students who did not follow the proximity hypothesis for question 1 ($U_{12,39}$ = 217.5, P = 0.700). Overall, these analyses suggest that students tended to stick with either the scientific or the primitive hypothesis when answering questions, but students switched between hypotheses when employing the proximity hypothesis.

Using the metrics from the PCO, cluster analysis generated four discrete clusters, and one individual that lay between the third and fourth clusters (Figure 3). The first cluster (right-most cluster on Figure 3; N = 3) was typified by students who used almost every hypothesis to answer their questions. In other words, they were highly inconsistent in the way they applied any of the hypotheses to complete the survey. The second cluster (N = 18) was typified by students who used an unspecified (i.e. 'other') hypothesis to answer Question 3 and otherwise used a mixture of all four alternate hypotheses to answer the remaining questions. Cluster three (N = 8) was distinct in that all students in this group answered Questions 3, 4a and 4f according to the scientific hypothesis, otherwise they tended to answer remaining questions (especially questions 2 and 4g) mostly according to an unspecified hypothesis (i.e. 'other'). The fourth cluster (N = 21) included two individuals who answered all questions according to the scientific hypothesis. Overall, this fourth cluster was typified by students who followed the scientific hypothesis in answering Questions 3 and 4f, however, they often switched to the proximity hypothesis while answering Questions 2 and 4g.

To summarize the description of the four clusters: cluster one was typified by students who used all hypotheses to answer questions (and received the lowest scores on the questionnaire); cluster two was typified by students who used a mix of hypotheses but also consistently followed an 'other' hypothesis for Question 3. The final two clusters included students who answered questions largely consistent with the scientific hypothesis. Cluster three students switched to following an 'other' hypothesis when answering Questions 2 and 4g. Cluster four students included the highest achievers on the questionnaire (following the scientific hypothesis throughout), but this group also consistently answered Questions 2 and 4g according to the proximity hypothesis. Hence, cluster four was typified by students who occasionally switched from the scientific to the proximity hypothesis.

Taken together, the two forms of statistical testing indicate that slightly more than half of the students did tend to follow the scientific hypothesis consistently. Some students consistently applied both the scientific and proximity hypotheses together: these two hypotheses clustered together in many students' answers. However, very few students answered using the proximity hypothesis throughout the survey, so they did not consistently interpret that the proximity of tips equals relatedness. There were a few students who tended to follow the primitive hypothesis in a consistent manner across the survey, indicating that they are reading the cladogram from left to right and perceiving organisms as branching off.

DISCUSSION

These results indicate that many students frequently misread simple cladograms, even after extensive classroom instruction. The majority of students who strayed from the scientific hypothesis tended to switch to the proximity hypothesis for certain questions (Questions 2 and 4g in particular), but did not stick to this latter hypothesis in a consistent manner across other questions. Others that strayed from the scientific hypothesis fell into the 'other' category, discussed further below. A small percentage of students answered using the primitive hypothesis across several questions. An even smaller percentage answered questions in ways that were consistent with all of the hypotheses, including the 'other' category.

These response patterns indicate that most students generally followed the scientific hypothesis. However, many switched from the scientific hypothesis to the proximity hypothesis for a couple of questions, implying that they likely have a misconception about the visual structure of phylogenies. Instead of realizing that clades can be rotated about the nodes without changing the relationship depicted, they view clades that are next to each other in a cladogram as more related, independent of the underlying structure. This finding is consistent with prior studies on reading cladogram (e.g., Meir and Perry, 2007; Gregory 2008) as well as other studies that show that students tend to focus their time on reading words, but not examining diagrams that may help explain those words (e.g., Busch, 2011).

A few students answered questions in ways that were consistent with the primitive hypothesis, and these students tended to be consistent in that framework of interpretation. These students seem to be more immured in that they will consistently misinterpret cladograms, whereas students who have a proximity bias will be more likely to switch to the more appropriate scientific interpretation of cladograms. The persistence of the primitive hypothesis is not unexpected, given the myriad of fundamental visual difficulties students must overcome to read cladograms correctly. Long lines, straight lines, unbroken lines, and lines with fewer branches all conspire to lead students to believe that change is not happening, so those organisms are perceived as more primitive (Meir and Perry, 2007; Novick and Catley, 2007; Gregory, 2008).

To verify that students were not just answering haphazardly, we analyzed individual students responses to Questions 2 and 4g on the questionnaire since they were nearly identical, asking students if B was more closely related to A, D, or both equally. Forty-three of the 51 students (84%) answered the two questions the same. Of those 43 students, 24 students responded that B is more closely related to A (proximity hypothesis), 11 students correctly responded that B is more closely related to D (scientific hypothesis), and eight students responded they are equally related. The similar answers of students to both questions indicate that most students are putting thought into their responses, and even if they were guessing, there is consistent reasoning behind their guesses.

We further analyzed student responses that do not fit into any of the hypotheses (the 'other' category). Questions 1, 2, 3, and 4g are the questions where more than 5% of responses were not directly explained by one of the hypotheses (Table 4). This result can be explained in several ways: 1) students were simply guessing, 2) they were confused by the answers, or 3) they have another view not tested by the hypothesis. The similarities of students' answers to Questions 2 and 4g indicate that it is not likely that they are guessing at random. The answers to Questions 1 and 3 had many possible permutations. Question 3 in particular had the most components as compared to any other question, and it is possible that students applied more than one hypothesis while answering this question: applying one hypothesis while addressing one component of the question, and applying another while addressing another. Though we cannot identify this clearly, the many answers to this question were consistent with a blend of the scientific and proximity hypotheses. Hence, this result may give further evidence that students often confuse the scientific and proximity hypotheses while interpreting cladograms.

Recommendations for teaching cladograms

Our results indicate that students have a difficult time reading cladograms, even after consistent use in the classroom. We therefore recommend that instructors point out the common misconceptions described in this paper that students have when reading cladograms, and require students to directly confront discrepancies between their own ideas and the scientific interpretation in order to effectively reduce or eliminate the misconception (e.g., Posner et al., 1982; Hewson and Hewson, 1988; Chan et al., 1997; Guzzetti, 2000). Other researchers also advocate for direct training on reading cladograms, with explicit instruction and practice (Morabito, 2010).

Lab exercises and activities that help students to solve cladogram-interpretation problems have been developed (e.g., Wiley et al., 1991; Goldsmith, 2003; Flory et al., 2005; Cooper et al., 2006; Julius and Schoenfuss, 2006; Kalinowski et al., 2006; Burks and Boles, 2007). Most are based in modern biology (few are explicitly paleontological), and activities mostly analyze information from living organisms (including DNA), although a few use non-living things as examples (Goldsmith, 2003; Burks and Boles, 2007).

While these activities ultimately give students some experience manipulating cladograms and even a sense for the kinds of decisions that must be made by systematists, there are shortcomings associated with many of these activities. Many require extra materials, extended periods of time, such as a lab period or small discussion group, and/or small class sizes, which make them difficult to complete in a large lecture class. Most importantly, however, from the vantage point of this study, they do not directly address the particular misconceptions that beset students when working with cladograms. In short, students can complete these kinds of activities successfully, and then fail to apply their conclusions to more familiar groups. When asked about birds, for example, students revert to familiar, Linnaean-based categories, and conclude that a bird could not be a dinosaur because dinosaurs are reptiles.

When students are required to directly confront discordancies resulting from their misconceptions, they learn the scientific interpretation more effectively, as long as they are given practice (e.g., Posner et al., 1982; Hewson and Hewson, 1988; Chan et al., 1997; Guzzetti, 2000). That being the case, activities that are applied to remedy the kinds of common misconceptions engendered by cladograms should specifically target those areas of confusion highlighted by this study. Some examples are given below, and we emphasize that we recommend giving students the opportunity to participate in hands-on activities, in whose design the conflict between student preconception and scientific usage is immanent.

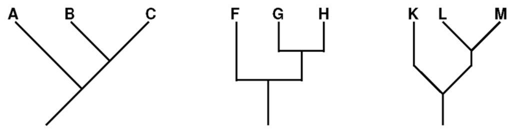
Simple activities that can be added to lecture easily are Think-Pair-Share activities and ConcepTest questions (Mazur, 1997; Crouch and Mazur, 2001; McConnell et al., 2003; McConnell et al., 2006). Both of these activities consist of questions posed during lecture with students responding individually, discussing their answer with peers, and then reporting back on their answers through a variety of means, such as a vote or an instructor-lead discussion. These questions can take as little as a few minutes of class time, but can result in dramatic change in that they require students to apply their knowledge. The questions on the questionnaire designed for this study (Figure 2) as well as questions from the Lecture Tutorials (described below) both provide a good set of discussion questions. For example, Question 2 on the questionnaire will help students confront the proximity misconception directly, and Question 4d directly addresses the primitive misconception. Initial questions should use letters or symbols instead of taxa familiar to students to encourage them to read the cladogram instead of make inferences from their prior (correct or incorrect) knowledge (Gregory, 2008; Morabito et al., 2010; Bair, 2011; Novick et al., 2011). After students are more comfortable with cladograms, organisms with relationships that students likely have misconceptions about, such as birds and dinosaurs, can be introduced to challenge the students to read the cladograms correctly.

Another possible solution is to design Lecture Tutorials, short worksheets students complete in groups during class, that address student misconceptions when learning a variety of topics (Kortz et al., 2008; Kortz and Smay, 2010). These worksheets are more involved than Think-Pair-Share activities and ConcepTest questions, and require students to delve more deeply into the topic. In the case of dinosaur phylogeny, focused Lecture Tutorials have been written to address difficulties students have when reading cladograms. An example Lecture Tutorial is included (Figure 4; additional Lecture Tutorials can be downloaded from the webpage for Fastovsky and Weishampel, 2009; http://www.cambridge.org/gb/knowledge/isbn/item5708889/?site_locale=en_GB).

Another option is web-based supplementary instruction questions that students complete out of class, where students not only give their answers but also a rating of the confidence in their answers (Brewer, 2004). This latter suggestion lets the instructor adjust lectures beforehand by giving her/him an appreciation for how deeply held any misconceptions may be. These questions also could be generated fairly simply in the same way as the Think-Pair-Share activities and ConcepTest ques-

Reading Cladograms

Part 1: Different Depictions
In paleontology, cladograms are depicted in a variety of ways, all showing the same basic idea of groups within larger groups.



1) Which letter on each of the other diagrams shares the same position as A? (circle one letter per diagram)
F G H K L M

2) For each diagram, which two letters indicate branches that share the most characteristics, so are most similar to each other? (circle two letters per diagram)
A B C F G H K L M

3) For each diagram, which letter is most distantly related to the other two letters? (circle one letter per diagram)
A B C F G H K L M

4) For the center diagram, would the same information be conveyed if you switched G and H? Explain your answer.

5) For the right diagram, would the same information be conveyed if you switched K and L? Explain your answer.

6) For the left diagram, would the same information be conveyed if you switched B and C? Explain your answer.

7) As shown above, the cladograms indicate groupings that are not about relationships. Now we'll put some animals on the cladograms, arranged in a way to show relationship. If, for example, you were to put earthworms, turtles, and dinosaurs on the diagrams, at which node would earthworms be located on each of the three diagrams? (circle one letter per diagram)
A B C F G H K L M

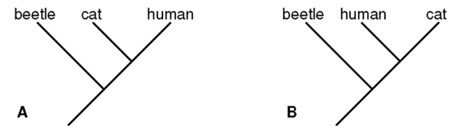
8) If you were to put earthworms, turtles, and dinosaurs on the diagrams, which of the letters below is NOT a place you could put dinosaurs? (circle all that apply)
A B G H K L

9) Check with another group of students that you have the same two letters circled for Question 8.

© KORTZ LECTURE TUTORIALS FOR DINOSAURS: A CONCISE NATURAL HISTORY
DRAFT EDITION, 2009

Reading Cladograms

Part 2: Adding Animals
Examine the cladograms below showing relationships between beetles, cats and humans.



10) Which cladogram, if any, is more correct? Explain your answer.

11) Two students are debating which cladogram is more correct.

Student 1: *I think they're both equally as correct. They both show that humans and cats are more closely related to each other than to beetles.*

Student 2: *I think Cladogram A is more correct because it shows humans at the far end. Cats are closer to beetles and humans aren't closely related to beetles, so A is better.*

With which student do you agree? Why?

12) Compare your answers to Question 10 and 11 to your answer to Question 6. Change any answers as necessary.

13) Why does the order on the top of the cladogram not necessarily indicate how closely organisms are related?

14) Draw two separate cladograms below, both showing the correct relationship between jellyfish, dinosaurs, and pigs.

© KORTZ LECTURE TUTORIALS FOR DINOSAURS: A CONCISE NATURAL HISTORY
DRAFT EDITION, 2009

FIGURE *4*.—Example Lecture Tutorial guiding students through reading cladograms; this and additional Lecture Tutorials can be downloaded as a resource from the webpage for *Dinosaurs: A Concise Natural History* by Fastovsky and Weishampel from Cambridge University Press at http://www.cambridge.org/gb/knowledge/isbn/item5708889/?site_locale=en_GB.

tions.

Finally, we want to also discuss a recommendation other authors have given. Although we did not test for it directly in this study, we feel that we would be remiss not to include a discussion of the presentation of cladograms themselves. The use of 'ladder' cladograms with straight diagonal lines (e.g., Fig. 1) may partially lead to the misconceptions observed in our study. In particular, ladder cladograms make it more difficult to visualize the clades as being able to rotate about the nodes (leading to the proximity hypothesis), they have long, straight lines (leading to the primitive hypothesis), and they have branches off of a main line (also leading to the primitive hypothesis) (Morabito, 2010; Bair, 2011). Therefore, a 'tree' (or 'goalpost') diagram (e.g., Fig. 2, Question 3A–F) is recommended (Morabito, 2010; Bair, 2011). This diagram reduces the misconceptions by appearing more like a mobile that can be rotated, and the goalposts break up long lines without having a single, main line.

CONCLUSIONS

Because we believe that phylogenetic systematics ought to be used when teaching organismal relationships, and because, in this study, we have identified the persistence of misunderstandings, focused effort can now be devoted to helping students to understand cladograms without the misconceptions identified in this study. Our observations indicate that it is important to target the proximity hypothesis and highlight why this reading of cladograms is incorrect, especially by discussing the fact that clades can be rotated about the nodes without changing the relationships presented. Because students do not apply the 'proximity' hypothesis consistently, they likely can be brought around more easily to the application of the scientific hypothesis. Misconceptions illustrated by the primitive hypothesis appear to be more deeply ingrained. However, both hypotheses likely can be partially addressed by switching the presentation of cladograms from ladder cladograms to tree or goalpost cladograms.

In order for students to read cladograms properly, instructors teaching this subject must be keenly aware of these common, but fundamental misconceptions. Our view is that phylogenetic systematics cannot be effectively taught, particularly in large, general-education settings, unless these areas of misunderstanding and preconception are confronted and explicitly remediated. Our results indicate that although cladograms are excellent tools to convey knowledge to experts, there are common misunderstandings that students have about cladograms that require a significant amount of focused teaching and explanation to overcome.

ACKNOWLEDGMENTS

We thank the editors of this volume and two reviewers, W. I. Ausich and an anonymous reviewer, for their thoughtful reviews of the manuscript. We thank A. R. Bair for her enlightening discussions. The work has been partially supported by the Department of Geosciences, University of Rhode Island, by the Community College of Rhode Island, and by the Institute for Integrative Bird Behavior Studies, College of William and Mary. IRB approval for this study (Investigating students' understanding of four-legged-animal relationships, 12-08-2009) was attained.

REFERENCES

BAIR, A. R. 2011. Teaching and learning about cladograms in geology courses: Insights from work in biology, a pilot study in a paleobiology course and cognitive science, 221-6. Geological Society of America Abstracts with Programs, 43(4):533.

BAUM, D. A., S. DEWITT SMITH, AND S. S. S. DONOVAN. 2005. The tree-thinking challenge. Science, 310:979–980.

BREWER, C. A. 2004. Near real-time assessment of student learning and understanding in biology courses. BioScience, 54:1034–1039.

BURKS, R. L., AND L. C. BOLES. 2007. Evolution of the chocolate bar: A creative approach to teaching phylogenetic relationships within evolutionary biology. American Biology Teacher, 69:229–237.

BUSCH, M. 2011. Late Quaternary Normal Faulting and Hanging Wall Basin Evolution of the Southwestern Rift Margin From Gravity and Geology, B.C.S., MX and Exploring the Influence of Text-Figure Format on Introductory Geology Learning. Unpublished Ph.D. dissertation, Arizona State University, 209 p.

CHAN, C., J. BURTIS, AND C. BEREITER. 1997. Knowledge building as a mediator of conflict in conceptual change. Cognition and Instruction, 15:1–40.

COOPER, S., D. HANMER, AND B. CERBIN. 2006. Problem solving modules in large introductory biology lectures enhance student understanding. American Biology Teacher, 68:524–529.

CRACRAFT, J. 1979. Phylogenetic analysis, evolutionary models, and paleontology, p. 7–39. *In* J. Cracraft and N. Eldredge (eds.). Columbia University Press, New York, New York.

CRISP, M. D., AND L. G. COOK. 2005. Do early branching lineages signify ancestral traits? Trends in Ecology and Evolution, 20: 122–128.

CROUCH, C. H., AND E. MAZUR. 2001. Peer instruction: Ten years of experience and results. American Journal of Physics, 69:970–977.

DE QUEIROZ, K., AND J. GAUTHIER. 1992. Phylogenetic taxonomy. Annual Review of Ecology and Systematics, 23:449–480.

DE QUEIROZ, K., AND J. GAUTHIER. 1994. Toward a phylogenetic system of biological nomenclature. Trends in Ecology and Evolution, 9:27–31.

DINGUS, L., AND T. ROWE. 1998. The Mistaken Extinction. W. H. Freeman, New York, New York.

ELDREDGE, N., AND J. CRACRAFT. 1980. Phylogenetic Patterns and the Evolutionary Process. Columbia University Press, New York, New York.

FASTOVSKY, D. E. AND D. B. WEISHAMPEL. 2009. Dinosaurs: A Concise Natural History. Cambridge University Press, New York, New York.

FLORY, S. L., E. L INGRAM, B. J. HEIDINGER, AND T. TINTJER. 2005. Hands-on in the non-laboratory classroom. American Biology Teacher, 67:542–547.

FOOTE, M., AND A. MILLER. 2007. Principles of Paleontology. 3rd ed., W. H. Freeman, New York.

GOLDSMITH, D. W. 2003. The great clade race. American Biology Teacher, 65:679–682.

GOULD, S. J. 1989. Wonderful Life: the Burgess Shale and the Nature of History. W. W. Norton and Company, New York.

GOWER, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika, 53:325–338.

GREGORY, T. R. 2008. Understanding evolutionary trees. Evolution Education Outreach, 1:121–137.

GUZZETTI, B. J. 2000. Learning counter-intuitive science concepts: What have we learned from over a decade of research? Reading and Writing Quarterly, 16:89–98.

HAMMER, Ø., D. A. T. HARPER, AND P. D. RYAN. 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica, 4:9.

HENNIG, W. 1979. Phylogenetic Systematics. Translated by D. D. Davis and R. Zangerl. University of Illinois Press, Urbana, Illinois.

HEWSON, P. W., AND M. G. A. B. HEWSON. 1988. An appropriate conception of teaching science: A view from studies of science learning. Science Education, 72:597–614.

JULIUS, M. L., AND H. L. SCHOENFUSS. 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, July/August:40–45.

KALINOWSKI, S. T., M. L. TAPER, AND A. M. METZ. 2006. How are humans related to other primates? A guided inquiry laboratory for undergraduate students. Genetics, 172:1379–1383.

KORTZ, K. M., AND J. J. SMAY. 2010. Lecture Tutorials for Introductory Geoscience. W. H. Freeman, New York, New York, 88 p.

KORTZ, K. M., J. J. SMAY, AND D. P. MURRAY. 2008. Increasing student learning in introductory geoscience courses using lecture tutorials. Journal of Geoscience Education, 56:280–290.

MAZUR, E. 1997. Peer Instruction: A User's Manual, Prentice Hall, Upper Saddle River, 253 p.

MCCONNELL, D. A., D. N. STEER, AND K. D. OWENS. 2003. Assessment and active learning strategies for introductory geology courses. Journal of Geoscience Education, 51:205–216.

MCCONNELL, D. A., D. N. STEER, K. D. OWENS, AND C. C. KNIGHT. 2005. How students think: Implications for learning in introductory geoscience courses. Journal of Geoscience Education, 53:462–470.

MCCONNELL, D. A., D. N. STEER, K. D. OWENS, J. R. KNOTT, S. VAN HORN, W. BOROWSKI, J. DICK, A. FOOS, M. MALONE, H. MCGREW, L. GREER, AND P. J. HEANEY. 2006. Using conceptests to assess and improve student conceptual understanding in introductory geoscience courses. Journal of Geoscience Education, 54:61–68.

MEIR, E., J. PERRY, J. C. HERRON, J. KINGSOLVER. 2007. College students' misconceptions about evolutionary trees. American Biology Teacher, 69:71–76.

MEISEL, R. P. 2010. Teaching tree-thinking to undergraduate biology students. Evolution: Education and Outreach, 3:621–628.

MORABITO, N., K. M. CATLEY, AND L. R. NOVICK. Reasoning about evolutionary history: The effects of biology background on post-secondary students' knowledge of most recent common ancestry and homoplasy. Journal of Biological Education, 44:166–174.

NEE, S. 2005. The great chain of being. Nature, 435:429.

NELSON, G., AND N. PLATNICK. 1981. Systematics and Biogeography. Columbia University Press, New York, New York.

NOVICK, L. R., AND K. M. CATLEY. 2007. Understanding phylogenies in biology: The influence of a Gestalt perceptual principle. Journal of Experimental Psychology: Applied, 13:197–223.

NOVICK, L. R., K. M. CATLEY, AND D. J. FUNK. 2011. Inference is Bliss: Using evolutionary relationship to guide categorical inferences. Cognitive Science, 35:712–743.

O'HARA, R. J. 1988. Homage to Clio, or, toward an historical philosophy for evolutionary biology. Systematic Zoology, 37:142–155.

O'HARA, R. J. 1994. Evolutionary history and the species problem. American Zoology, 34:12–22.

O'HARA, R. J. 1998. Population thinking and tree thinking in systematics. Zoologica Scripta. 26:323–329.

POSNER, G. J., K. A. STRIKE, P. W. HEWSON, AND W. A. GERTZOG. 1982. Accommodation of a scientific conception: Toward a theory of conceptual change. Science Education, 66:211–227.

ROMER, A. S. 1966. Vertebrate Paleontology (third edition). University of Chicago Press, Chicago, Illinois, 468 p.

RUDOLPH, J. L. AND J. H. STEWART. 1998. Evolution and the nature of science: On the historical discord and its implications for education. Journal of Research in Science Teaching, 35:1069–1089.

WILEY, E. O., D. SIEGEL-CAUSEY, D. BROOKS, AND V. A. FUNK. 1991. The Compleat Cladist: A Primer of Phylogenetic Procedures. University of Kansas Museum of Natural History Special Publication 19. University of Kansas, Lawrence, Kansas.