College Students' Misconceptions About Evolutionary Trees

Eli Meir Judy Perry Jon C. Herron Joel Kingsolver

Evolution is at the center of the biological sciences and is therefore a required topic for virtually every college biology student. Yet several core aspects of evolution are non-intuitive. Evolutionary biology is broadly divided into two sub-disciplines: microevolution, which looks at how the distribution of traits in a population changes over relatively short time periods; and macroevolution, which looks at how new taxa arise over long time periods. Many studies have shown that students harbor misconceptions about key ideas in microevolution such as natural selection (e.g., Greene, 1990; Ferrari & Chi, 1998; Lawson & Thompson, 1998; Anderson, Fisher & Norman, 2002), and have outlined the classes of misconceptions that are most common. Yet macroevolution has received little attention (see Baum et al, 2005), despite being the area of evolution that receives the most media attention through newsworthy topics such as fossil discoveries, speciation, and the relationships among species.

Over the past year we have been building a new simulation software package called EvoBeaker (Meir et al., 2005) to teach college-level evolutionary biology through simulated experiments. We have built both micro and macroevolutionary laboratories into the program. For labs dealing with microevolution, we were able to examine the literature to identify the misconceptions that were most prevalent among college students, and therefore most important for us to address. We could find no similar literature on misconceptions about macroevolution. We thus began our software design phase by seeking to identify misconceptions among college students about the subject of evolutionary trees, the diagrams used to display branching evolutionary relationships between populations or species (which include cladograms and other diagrams showing reconstructions of evolutionary history), and the ideas embedded in these diagrams. Here we report the most common misconceptions among college students in their understanding of evolutionary trees, and their demonstrated ability to perform typical "tree-thinking" skills.

Methods

Instrument Development

We used a seven-step process to gauge students' tree-thinking misconceptions.

Eli Meir is President of SimBiotic Software and Research Affiliate at MIT, Cambridge, MA 02139; e-mail:meir@simbio.com. Judy Perry is Research Manager, Teacher Education Program, MIT, Cambridge, MA 02139; e-mail: jperry@mit.edu. Jon C. Herron is Lecturer in the Department of Biology, The University of Washington, Seattle, WA 98195; e-mail: herronjc@comcast.net. Joel Kingsolver is Professor, Department of Biology, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599; e-mail: jgking@bio.unc.edu.

1. Two authors of this study (Herron & Kingsolver) reflected on their experiences teaching evolutionary trees and created an initial list of misconceptions they anticipated would be common among college students.
2. Using this list, we developed written questions we thought would elicit those misconceptions. To fully elicit students' thought processes and provide as little prompting as possible, all questions were open-ended, free response questions. Each required students to draw diagrams, write short essays, or provide written explanations of their answers.
3. We submitted a version of the written test to a group of subject matter experts–evolutionary biologists at the University of North Carolina, Chapel Hill–for their feedback.
4. To refine and expand the list of student misconceptions, we pilot-tested the initial open-ended questions with 34 Seattle-area college students enrolled in an evolutionary biology course.
5. The initial student responses were used to modify the pilot questions and list of misconceptions.
6. An additional 10 Boston-area college students were given varying versions of the pilot instrument. Immediately after taking the written test, these students were each interviewed and asked to rephrase the questions in their own words. After each student, we refined our list of misconceptions and modified, added or deleted questions from our written test as new, unanticipated ideas came from the students and as we observed which questions worked to elicit misconceptions.
7. Finally, based on students' written answers and verbal explanations, we selected common misconceptions and commonly-used rationales and wording to construct multiple choice answers reflecting the variety of different misconceptions students expressed in answering each question. The primary aim of constructing these multiple choice responses was to be able to correlate specific answers with specific misconceptions. However, some open-ended questions (e.g., asking students to draw diagrams) were kept to provide students with the chance to give more open-ended responses. Rubrics detailing how these open-ended questions were coded are available on request. The complete 21 question instrument is available by request from the lead author.

Instrument Validation & Data Collection Procedures

Once we felt that the instrument elicited misconceptions effectively, we recruited 124 students from Boston area colleges (herein referred to as "local" students) to take the revised survey, which took approximately 20 minutes to complete. Prior to beginning the pre-test, participating students were given several pages describing evolutionary trees from a leading biology textbook (Freeman & Herron, 2004, pp. 47-49) on evolutionary trees to read and review. After reading these materials, local students took the test individually, under proctored testing conditions. Local students participated voluntarily in our study outside of class and were compensated for their time. We also recruited instructors from nine colleges around the country to give the pre-test to students in their classes, from which we received 286 student pre-tests (herein referred to as "remote" students). Like local students, remote students took the pre-test in a proctored setting. Tests were not scored by students' professors and scores did not count toward formal grades. All students participating in this study were required to have had at least one college level lecture on evolutionary biology, and most had been introduced to evolutionary trees in their classes.

To assign test answers to particular tree-thinking misconceptions or skills, one author (Meir), who conducted the majority of the initial student interviews, used the written answers of 10 students plus his notes from pilot interviews to pair specific test responses to specific misconceptions. This was done conservatively – answers that did not clearly belong to any misconception were not assigned to any particular skill or misconception. To validate these pairings of responses and skills/misconceptions, another author with expertise in studying and teaching evolutionary biology (Herron) independently paired answers with skills/misconceptions. He had participated in constructing the tests but not the student interviews. These two authors then discussed the differences to arrive at a consensus.

Results

Four Major Misconceptions About Evolutionary Trees

Through written pre-tests in which we asked students to write explanations of their answers, and interviews following the pre-test, we identified four common misconceptions about students' interpretations of evolutionary trees. These are listed below, together with examples of written or oral responses from participants which typify these misconceptions.

Incorrect Mapping of Time

Many students were confused about spatial representation of the flow of time on an evolutionary tree. While reading vertically-oriented trees, in which time flows from bottom to top, students often thought that the horizontal order in which extant species were drawn across the top was significant, and assumed that the older species were on the left and the younger species on the right. Students also frequently thought that the tree was anchored at the upper left species, and assumed that this species was the common ancestor of the others on the tree. We saw these misconceptions in both written answers and in arrows which students were asked to draw indicating the flow of time (Figure 1b shows correct drawing). For instance, when asked the order in which the node species (P, C, W, N) in Figure 1a lived, one student said "The older animals are at the top." Another student said that "Mammals, the most recent, branched from (N), which came last [after P, C, and W]". When asked to draw an arrow representing time on an evolutionary tree (Question 8), many students drew a horizontal arrow from

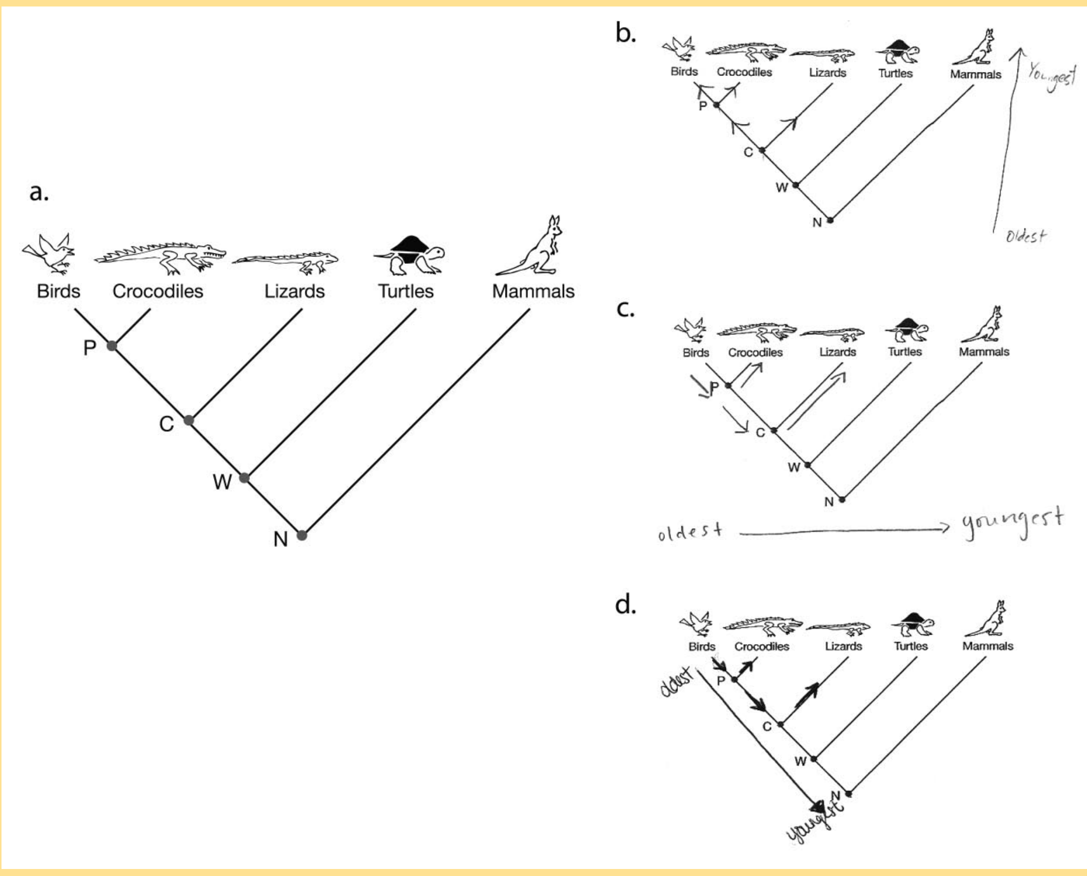

Figure 1. Students conception of time on an evolutionary tree. (a) An evolutionary tree that we presented to students to use for several of the test questions. Students were asked to draw arrows showing the direction of time on the tree from oldest to youngest, and the direction time moves between four particular nodes. (b) One student's correct answers. (c) This student incorrectly shows time going left to right rather than bottom to top. (d) This student incorrectly shows time going top to bottom. Both students in (c) and (d) show time going from the left-most tip of the tree (birds) down to the first common ancestor (P) and then back up to the next tip over (crocodiles), as if birds evolved into crocodiles.

left to right (Figure 1c), others drew an arrow straight down, and still others drew an arrow that paralleled one of the outside branches (Figure 1d). Some students even drew two arrows, one left to right and the other top to bottom. Similarly, when asked to draw arrows showing the direction of time between tips and nodes (Question 7), many students drew arrows from the top left tip down, and then back up to tips farther to the right (Figures 1c and d). Among the local students, some of this confusion may have been enhanced by the textbook chapter they read, which stated that trees are sometimes drawn vertically and other times horizontally, but we saw the same range of arrows among students from remote sites who presumably used other texts.

Tip Proximity Indicates Relationship

Students often thought that species drawn closer together at the tips of the tree were more closely related to each other than those drawn farther apart. For instance, to a question about whether, according to the phylogeny in Figure 1a, turtles are more closely related to mammals than to birds, one student answered "Yes," because "on the chart, turtles are closer to mammals." When the tree was rotated around node (W) so that the left-to-right order along the top was turtles, lizards, crocodiles, birds, and then mammals, that same student answered a similar question with a "No," because "[turtles are] closer to birds than they are mammals."

Node Counting

Some students thought that the number of nodes crossed in tracing a path between two species on a tree indicated how closely related they were. For instance, looking at the tree in Figure 1a, these students would say that turtles are more closely related to mammals than to crocodiles because there were only two nodes (W and N) between turtles and mammals, while there were three (W, C, and P) between turtles and crocodiles. While watching students answer these questions, we observed them counting the nodes. Their explanations would also involve node counting, with one student explaining, "The ancestor of turtles (W) is connected to the ancestor of mammals whereas the ancestor of birds is separated by one ancestor (C)."

Straight Line Equals No Change

Often evolutionary trees have one lineage that is drawn as a straight line extending from a common ancestor deep in the tree to one of the tips, with other species branching off that line. This is an accident of the drawing, but many students interpret it as meaning that the species at the tip of the straight line is the ancestor of the others that branch off of the line. Looking at the target tree for test questions (10-14) (see Appendix), a common interpretation of the relationship between crocodiles and birds is that "Crocodiles evolved from birds" or "Crocodiles came after birds." If you switch birds and crocodiles in the Figure, as in Question 10, but leave all evolutionary relationships the same, these students will say that it shows a different tree because birds now evolved from crocodiles rather than crocodiles evolving from birds. For instance, one student who said "crocodiles evolved from birds," when shown a tree with crocodiles and birds switched (Question 10), explained, "Here, although it [the tree] looks the same, it's not saying the bird was first. It's saying the crocodile was [first] and the birds came from that." Students will also interpret the root of the tree as being the same species as the tip species at the end of the straight line, especially if no trait changes or branches are shown coming off that lineage. For instance, often in our interviews students looking at the tree in Figure 1a would say that (N), the common ancestor, is a bird or a mammal. Question 20 on our test (see Appendix) addressed this explicitly by asking students if one extant species is the common ancestor of all the other extant species.

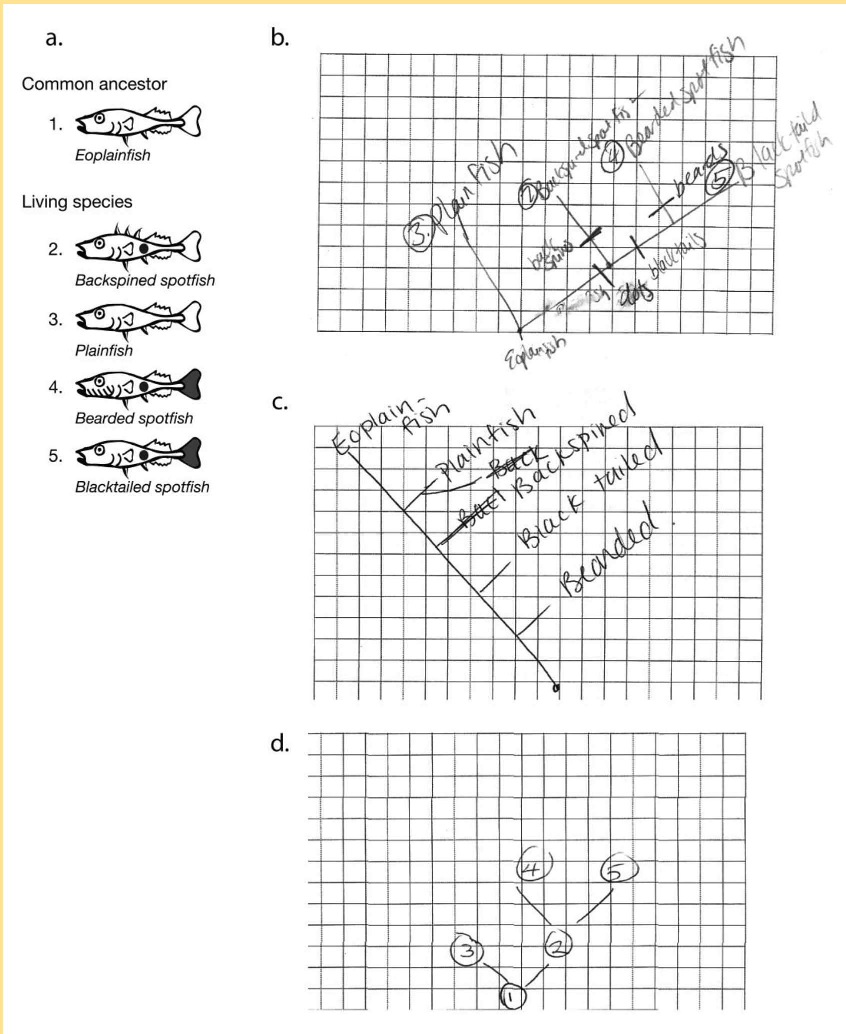

Figure 2. Students ability to reconstruct a tree from phenotypic traits. (a) We presented this collection of species to the students with four discrete traits and asked them to reconstruct the evolutionary tree for this family. (b) One student's correct answer. (c) This student drew the tree as a progression of more complicated species, starting from the ancestor (he/she also drew time going from top to bottom). (d) This student drew one of the extant species as the common ancestor of two other extant species.

Three Skills for Correctly Interpreting Evolutionary Trees

In addition to the four misconceptions identified above, we also looked for three skills that evolutionary biologists in our group (Herron & Kingsolver) identified as being important in thinking correctly about evolutionary trees. Each skill is described briefly below.

Skill 1. Reading Traits from Tree

Given a tree with trait transitions marked, the student is able to correctly deduce the traits of a particular species on the tree.

Skill 2. Deducing Ancestral Traits

Given the terminal taxa (species) on an evolutionary tree with sets of traits given for each extant species, the student is able to deduce the most likely traits shared by the common ancestor of those species.

Skill 3. Reconstructing Trees

Given a small set of extant species with no convergence or loss of novel traits, and the common ancestor of these species, the student is able to draw an evolutionary tree showing the most likely phylogeny. This was the most challenging test question we asked students, and also the most challenging to score. For the purposes of this paper, the following rubrics were used to score four aspects of the reconstructed tree: a) Were all the extant species put on the same horizontal line at the top of the tree, to show that they all exist currently? b) Was the common ancestor put at the base of the tree? c) Was the branching pattern correct? d) Were traits marked correctly? Figure 2 shows a correct answer (Figure 2b) and two examples of incorrect answers (Figures 2c and d).

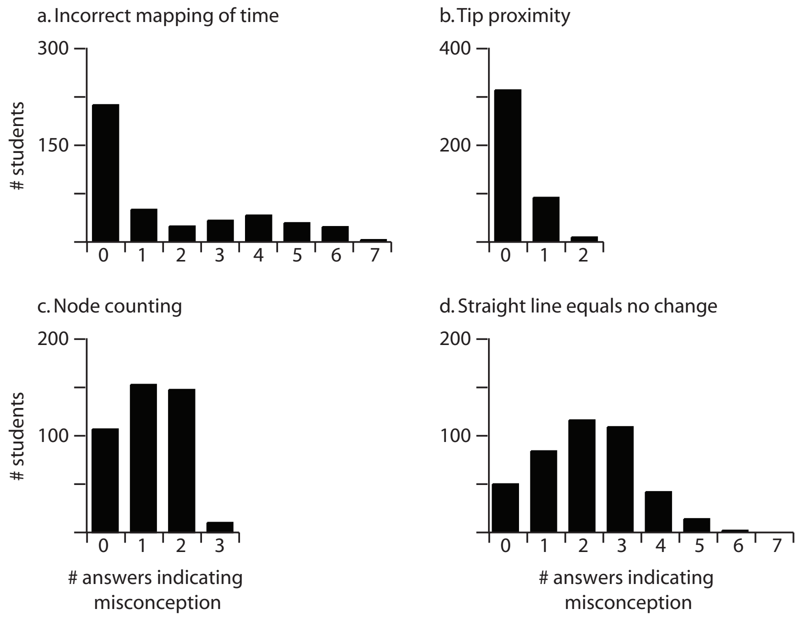

Figure 3. Number of responses each student selected indicating each misconception. For both Incorrect Mapping of Time and Tip Proximity, most students did not select any answers indicating those misconceptions, and among the remaining students there was a wide variation in how many answers they selected indicating the misconception. For Node Counting and Straight Line Equals No Change, most students selected some but not all possible answers indicating that misconception.

Prevalence of Misconceptions & Lack of Skills Among Undergraduate Students

Each misconception was captured by two to nine answers to questions on our written tests and each skill by one to three questions (Table A1 in the Appendix shows our mapping of answers to misconceptions/skills). As shown in Figures 3 and 4, large percentages of students showed each of the misconceptions in at least some of their answers on our tests, and the majority of students lacked each of the skills. We had different numbers of answers keyed to each misconception and skill, so to be able to compare the prevalence of each with the others, we arbitrarily declared that a student was showing a problem if he/she picked more than 1/3 of the answers that indicated the misconception or lost more than 1/3 of the points for a skill (roughly the cut-off for a failing grade on a typical test that is not curved). With that metric, we find the 24% - 40% of students harbor each of the misconceptions, and 65% - 84% lack each skill (Table 1). The most common misconceptions were Straight Line Equals No Change and Node Counting, with about 40% of students harboring each. Interestingly, the correlation among skills is rather low.

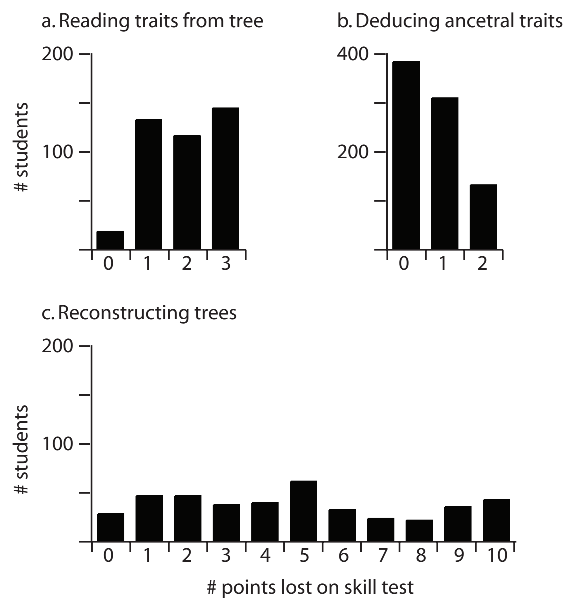

Figure 4. Number of answer points students lost on showing lack of ability for each of the skills. Students in the 0 bar of each graph received perfect scores for that skill. Very few students got all possible points for either Reading Traits from Tree or Reconstructing Trees, and all skills had a substantial number of students who did not receive any of the possible points (rightmost bar in each graph), showing a complete lack of understanding.

For none of the skills were students' abilities more than moderately correlated with any other skill (highest correlation r=0.36).

We cannot properly do this same analysis on all combinations of misconceptions because the answers for misconception questions were not entirely independent of each other. Picking an answer to a question indicating one misconception necessarily means the student did not pick an answer to that question indicating a different misconception. Nevertheless, even in misconceptions scored using questions that did not overlap (Tip Proximity and Straight Line Equals No Change) or overlapped only slightly (Incorrect Mapping of Time and Straight Line Equals No Change), we only found low-to-moderate correlations between misconceptions (highest correlation r=0.39). Thus at first blush it seems that different students harbor different sets of misconceptions and/or skills.

Local & Remote Students Are Similar

Our student sample included both students we recruited in the Boston area and proctored ourselves (n = 124), as well as students from schools outside of Boston who took our test as part of their class's participation in testing EvoBeaker laboratories (n = 286). To make sure that combining these results as we did above was legitimate, and also for future analyses (see discussion), we were interested to see whether the prevalence of misconceptions was similar between the local and remote populations of students. In all but one of our misconceptions/skills, we found no significant difference in the level of local vs. remote students (two-tailed *t*-test, none significant at the $p < .05$ level). The one exception was the Straight Line Equals No Change misconception where local students on average picked 0.5 more answers indicating the misconception than remote students ($p < 0.001$). Thus in general, it appears that the local and remote populations of students in our study are comparable in their understanding of evolutionary trees, at least as measured on our test.

Table 1. Percentage of students who exhibited a misconception or lacked a skill. Students were considered to exhibit a misconception if they picked more than 1/3 of the test answers keyed to that misconception, and were considered to lack a skill if they received less than 2/3 of the points assigned to that skill. Lower level students were students in 100 or 200 level classes (generally non-majors or intro biology); upper level were students in 300 level classes or above. * indicates that the upper level scores were significantly different from the lower level scores (one-tailed *t*-test, $p < 0.05$). Total number of students was 410. Total lower level students was 172 and upper level students was 114 (we only have class level for remote students).

% of students *exhibiting* a given misconception	Total	Lower Level	Upper Level
Incorrect Mapping of Time	31%	39%	21% *
Tip Proximity	24%	24%	25%
Node Counting	38%	35%	38%
Straight Line Equals No Change	40%	42%	24% *
% of students *lacking* a given skill			
Reading Traits from Tree	65%	67%	55% *
Deducing Ancestral Traits	84%	91%	71%
Reconstructing Trees	70%	74%	68%

Table 2. Misconceptions and skills in upper level versus lower level biology courses.
* represents statistically significant at the .05 level when Bonferoni corrected for multiple comparisons.

SKILLS	Level 1 Average Score (SD)	Level 2 Average Score (SD)	P-value
Reading Traits from a Tree (maximum score = 3)	1.8324 (0.9589)	2.1327 (0.8915)	0.0041*
Deducing Ancestral Traits (maximum score = 2)	0.6647 (0.6405)	0.8230 (0.8581)	0.0379
Reconstructing Trees (maximum score = 10)	4.4855 (3.0263)	5.0973 (3.2595)	0.0531
MISCONCEPTIONS			
Incorrect Mapping of Time (maximum score = 7)	1.9075 (2.1165)	1.1593 (1.8399)	0.0012*
Tip Proximity (maximum score = 2)	0.2486 (0.4466)	0.2832 (0.5256)	0.2754
Node Counting (maximum score = 3)	1.1214 (0.7941)	1.0885 (0.8405)	0.3901
Straight Line Equals No Change (maximum score = 7)	2.1734 (1.2455)	1.6106 (1.1835)	0.0001*

Upper Level Students Do Better Than Intro Level, but Gender Seems Unimportant

In three of the seven areas we tested, students in upper level classes (defined as 300 level or above) did better than those in lower level classes (Table 2). These included Skill 1: Reading Traits from Tree, where higher scores demonstrate stronger skills, and two of the four misconceptions: Incorrect Mapping of Time and Straight Line Equals No Change. The other four areas showed no significant difference in score between students in lower vs. upper level classes. Univariate Analyses of Variance revealed no statistically significant differences in the average overall test scores between groups compared by gender and level (upper level compared to lower level) (n=250 for this analysis as some students did not provide complete data).

Conclusions

Students in this study had all received college level instruction in biology, and presumably almost all had taken biology classes in secondary school as well. Evolutionary trees are shown in typical biology courses, and most students in this study had been formally exposed to tree-thinking either through written materials, lectures, or both. Thus the four tree-thinking misconceptions we identified here are likely to be held rather deeply and persist through at least cursory instruction on evolutionary trees. Similarly, the lack of skills needed to read and reconstruct trees persists through secondary and post-secondary introductory biology instruction. While students in upper level courses demonstrated fewer misconceptions and stronger skills related to reading trees, large percentages of students (25% or more) in the upper level courses still show signs of having trouble with those misconceptions and skills. Nevertheless, some of the students in this study may show evidence of misconceptions simply because they have never received good instruction in tree-thinking. We will bring further data to bear on this question in a subsequent article (Perry et al., in review).

These misconceptions are fundamental barriers to understanding how evolution operates and phylogenetic reasoning is central to much contemporary research on evolution. Students lacking a solid understanding of the concepts on our tests will likely have a hard time relating the disparate areas of biology through evolutionary thinking. Thus tools that go beyond standard techniques for teaching tree-thinking are sorely needed. A few such tools already exist, including the Flowers and Trees lab in our teaching simulation software EvoBeaker (Meir et al., 2005) and a card game called The Great Clade Race (Goldsmith, 2003). In a follow-up article (Perry et al., in review) we will discuss how well these two curricula aid students' understanding of tree-thinking.

Acknowledgments

Thanks to Eric Klopfer, Ilona Holland, Mark Hartman, Emily Franklin, and three thoughtful anonymous reviewers for advice and assistance, and to the group of biologists at the University of North Carolina, Chapel Hill who reviewed our tests. Eric Klopfer generously donated lab space to conduct student interviews. Thanks also to professors and students who volunteered to assist in our study (names withheld to preserve the students' anonymity). This work was partially funded by Grant #0341202 from the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Anderson, D., Fisher, K. & Norman, G. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching, 39*, 952-978.

Baum, D.A., Smith, S.D. & Donovan, S.S. (2005). Evolution – the tree-thinking challenge. *Science, 310,* 979-980.

Ferrari, M. & Chi, M. T. H. (1998). The nature of naïve explanations of natural selection.

Freeman, S. & Herron, J.C. (2004). *Evolutionary Analysis, 3rd Edition*. Upper Saddle River, NJ: Prentice Hall.

Goldsmith, D.W. (2003). The great clade race: presenting cladistic thinking to biology majors and general science students. *The American Biology Teacher, 65*, 679-682.

Greene, E. D. (1990). The logic of university students' misunderstanding of natural selection. *Journal of Research Science Teaching, 27*, 875-885.

International Journal of Science Education, 20, 1231-1256.

Lawson, A. & Thompson, L. (1998). Formal reasoning ability and misconceptions concerning genetics and natural selection. *Journal of Research in Science Teaching, 25*, 733-746.

Meir, E., Herron, J.C., Maruca, S., Stal, D. & Kingsolver, J. (2005). *EvoBeaker 1.0*. Available online at : http://www.simbio.com. Ithaca, NY: SimBiotic Software.

Appendix

We constructed a 20-question written test to elicit student misconceptions around evolutionary trees. We then mapped answers to these questions to particular misconceptions or skills. The test and answer key is available by writing to the corresponding author at SimBiotic Software. Table A1 shows our mapping of questions to misconceptions.

Table A1. Test answers used to score each misconception and skill (*skills italicized*). Most questions asked for multiple choice responses with four or five choices. Questions 7 and 8 asked students to draw arrows on a tree showing the direction of time. Questions 10–14 ask students to compare two trees and mark "X" if they are different from each other.

Misconception/Skill	Answers used in scoring
Incorrect Mapping of Time	1a, 4b, 5a, 5b, 6b, 7 wrong, 8 horizontal or downward arrow, 14 X
Tip Proximity	1b, 2a
Node Counting	1c, 2b, 4a
Straight Line Equals No Change	3b, 3c, 4d, 6a, 6c, 10X, 12X, 13X, 20 Yes
Reading Traits from Tree	17, 18, 19
Deducing Ancestral Traits	15, 16
Reconstructing Trees	21