Dees and Momsen *Evo Edu Outreach (2016) 9:3*
DOI 10.1186/s12052-016-0054-y

Evolution: Education and Outreach

RESEARCH ARTICLE

Open Access

Student construction of phylogenetic trees in an introductory biology course

Jonathan Dees and Jennifer L. Momsen*

Abstract

Background: Phylogenetic trees have become increasingly essential across biology disciplines. Consequently, 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 investigations that attempt to link construction tasks with student learning.

Results: 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 very common minor errors.

Conclusions: The method we developed to describe student-constructed phylogenetic trees uncovered several trends that warrant further 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 also 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 likely limited. We encourage researchers to use our method as a baseline for developing a more generalizable tool, which will support future investigations that attempt to link construction tasks with student learning.

Keywords: Phylogenetic trees, Cladograms, Conceptual models, Construction tasks, Evolution, Tree thinking

Background

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 biology education 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 fundamental concept for biological literacy (American Association

---

*Correspondence: jennifer.momsen@ndsu.edu
Department of Biological Sciences, North Dakota State University, PO Box 6050, Fargo, ND 58108, USA

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; Halverson et al. 2011; 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 basic interpretation skills among students necessitate continued research to address this discrepancy.

Some of the most common instructional activities concerning phylogenetic trees are construction exercises, 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 that students must develop interpretation skills before construction abilities. Thus, the effects of construction exercises 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 investigation by Eddy et al. (2013) did not describe the 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 6 weeks), form and function of plants and animals (next 5 weeks), and ecology (last 5 weeks). Students often collaborated in permanent, 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 sections (Cortright et al. 2003). All classes were observed, and instructional materials and assessments were collected to document instruction throughout the course.

Phylogenetic tree instruction

Phylogenetic trees were introduced during the evolution unit through reading assignments 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.

Following the phylogenetic tree introduction, students completed a group homework featuring a diagonal phylogenetic tree of chordates accompanied by a series of 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.

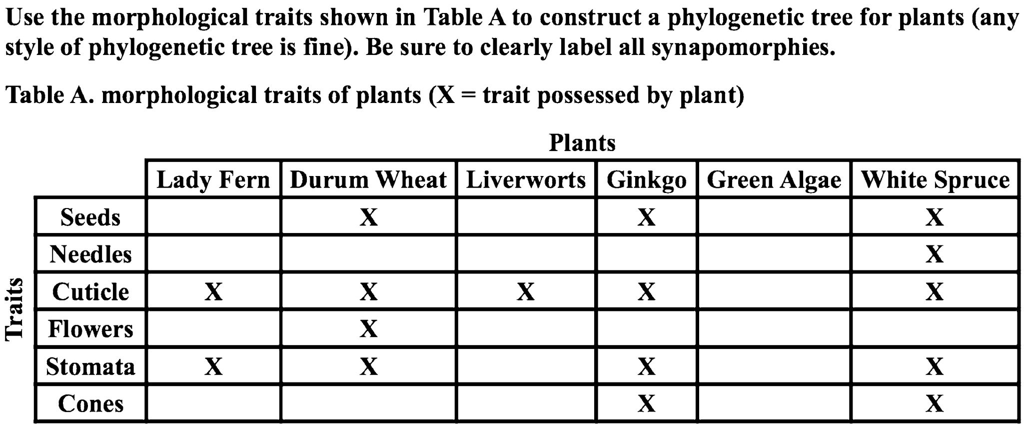
Data collection

The first phylogenetic tree construction exercise (Fig. 1) appeared on the group component 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.

The second phylogenetic tree construction exercise (Additional file 1: Figures S1–S2) 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 task version. The third phylogenetic tree construction exercise (Additional file 1: Figure S3) 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 section of the evolution unit exam (n = 23), individual component of the final exam (n = 77), and group section of the final exam (n = 22) constitute the data for this investigation.

Rubric development and coding

Rubrics were developed to code 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 (Fig. 2; "ladder" and "tree" formats described by Novick and Catley 2007), and rare cases containing both diagonal and bracket characteristics were coded as the predominant style. For example, a mainly diagonal phylogenetic tree



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

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

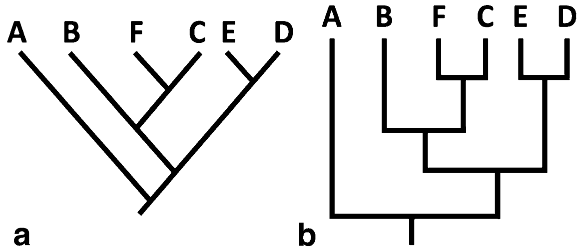


Fig. 2 Two common phylogenetic tree styles with equivalent branching patterns: **a** diagonal and **b** bracket (Dees et al. 2014; adapted from Gregory 2008)

with one divergence constructed in the bracket style was coded as diagonal.

Conventionality was used to describe characteristics 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 leftmost 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.

Phylogenetic trees constructed by students were assessed for accuracy based on major and minor errors. Major errors, such as incorrect relative 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).

Statistical analyses

Phylogenetic trees constructed during the individual component of the comprehensive final exam (only data obtained 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.

Results

Phylogenetic trees generated by introductory biology students during the group component of the evolution unit exam ($n = 23$), individual section of the final exam ($n = 77$), and group component of the final exam ($n = 22$) were evaluated in terms of style, conventionality, and accuracy.

Construction style

Students constructed diagonal phylogenetic trees more frequently than bracket phylogenetic trees during all three assessments (Fig. 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 ($p = 1.00$).

Construction conventionality

The majority of phylogenetic trees generated by students were coded as conventional for all three assessments (Fig. 4). Overall, 64 % of phylogenetic trees were constructed 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 (Fig. 5). Overall, 73 % of phylogenetic trees were generated correctly or adequately, including the group section of the final exam that contained convergent evolution (64 % correct or adequate). The most common major construction error was contemporary descent (extant taxa are descended from other extant taxa),

Table 1 Unconventional features observed in phylogenetic trees constructed by students

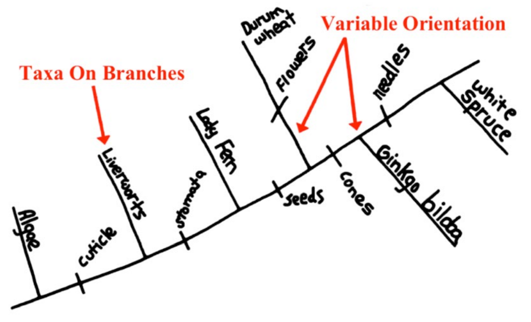
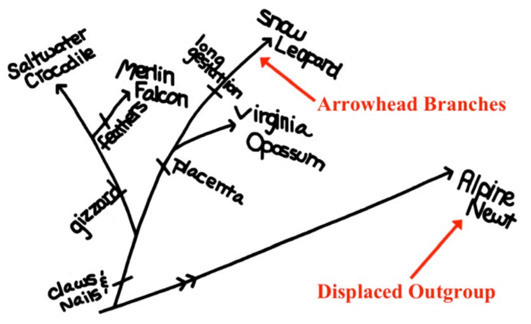
Feature	Description	Student-generated example
Variable orientation	Branches are not oriented in a consistent direction	
Taxa on branches	Taxa are on the branches rather than at the tips	
Arrowhead branches	Some or all branches are drawn with arrowheads	
Displaced outgroup	Outgroup is placed in an unconventional location	

Table 2 Major errors observed in phylogenetic trees constructed by students

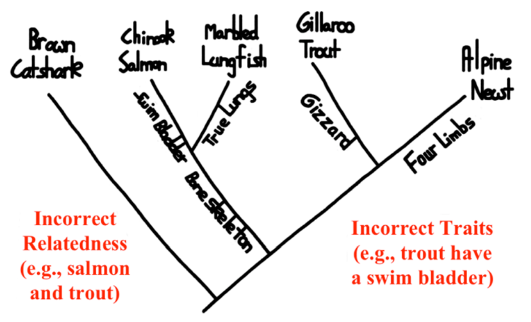
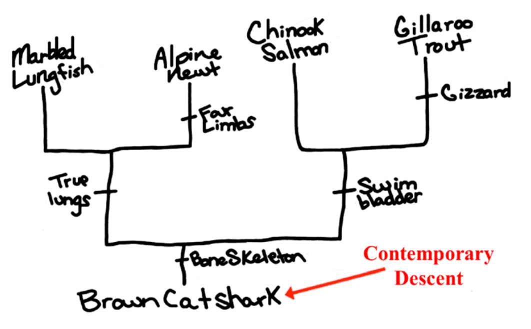
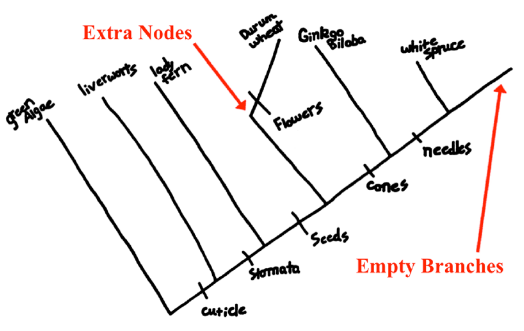
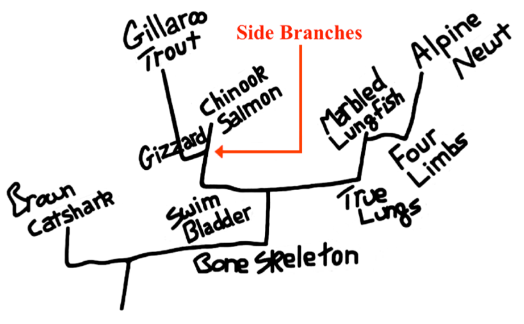
Major error	Description	Student-generated example
Incorrect relatedness	Relative placement of taxa is incorrect based on provided data	
Incorrect traits	Traits assigned to taxa are incorrect based on provided data	
Contemporary descent	Representation implies extant taxa are directly descended from one or more other extant taxa	

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)	

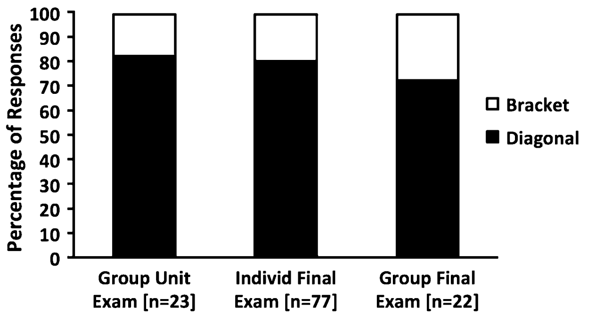


Fig. 3 Styles used by students to construct phylogenetic trees. *Columns* from *left* to *right* correspond to the activities in Fig. 1, Additional file 1: Figures S1, S2 and Additional file 1: Figures S3, respectively

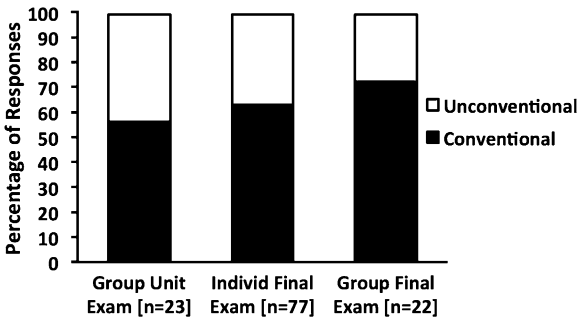


Fig. 4 Conventionality of phylogenetic trees constructed by students. *Columns* from *left* to *right* correspond to the activities in Fig. 1, Additional file 1: Figures S1, S2 and Additional file 1: Figures S3, respectively

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.

Discussion

Construction tasks are some of the most common instructional activities concerning 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

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 %)

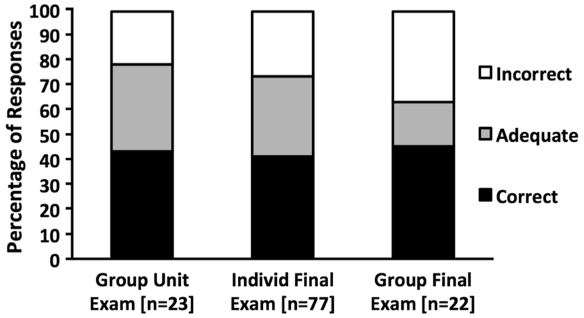


Fig. 5 Accuracy of phylogenetic trees constructed by students. *Columns* from *left* to *right* correspond to the activities in Fig. 1, Additional file 1: Figures S1, S2 and Additional file 1: Figures S3, respectively

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 strong preference for constructing diagonal phylogenetic trees across all three assessments (Fig. 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.

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			
Empty branches	6 (26 %)	31 (40 %)	5 (23 %)
Extra nodes	10 (43 %)	30 (39 %)	8 (36 %)
Side branches	0 (0 %)	7 (9 %)	1 (5 %)

Construction conventionality

The majority of students generated conventional phylogenetic trees during all three assessments (Fig. 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. However, the relationship between conventionality and accuracy is an important topic for future research.

Construction accuracy

The majority of phylogenetic trees were correct or adequate in terms of accuracy across all three assessments (Fig. 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 often than bracket phylogenetic trees, this outcome could have been impacted by the curriculum (Additional file 1: Table S1). The course textbook (Freeman 2011) contained only bracket phylogenetic trees, and instructional materials were also biased toward the bracket style. However, assessments (homework, reading quizzes, and exams) were skewed toward diagonal phylogenetic trees. Because assessment strongly impacts learning behaviors [e.g., (Cohen-Schotanus 1999; Wormald et al. 2009)], students could 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 only required to build one phylogenetic tree, in the style of their choice, during the individual section of the final exam (only data obtained 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 somewhat rare in phylogenetic trees constructed by students (Table 5). However, some of these errors could 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 exercises on student learning is uncertain based on the 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.

Additional file

Additional file 1. Phylogenetic tree construction tasks from the comprehensive final exam and a summary of phylogenetic tree styles used for instruction.

Authors' contributions

JD designed the assessment items, completed all data analyses, prepared figures, and contributed to data collection and manuscript preparation. JLM contributed to data collection and manuscript preparation. Both authors read and approved the final manuscript.

Acknowledgements

This investigation was conducted in compliance with the Institutional Review Board regulations (protocol #SM12217) and was funded by the National Science Foundation (DRL-1420321) and a STEM Education Fellowship from North Dakota State University. We are grateful to Rob Zastre, Julia Bowsher, and Lisa Montplaisir for research support and Elena Bray Speth for comments on earlier versions of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 25 November 2015 Accepted: 14 April 2016

References

American Association for the Advancement of Science. Vision and change in undergraduate biology education: a call to action. Washington, DC. 2011.
Baum DA, Offner S. Phylogenies & tree-thinking. Am Biol Teach. 2008;70(4):222–9.
Baum DA, Smith SD, Donovan SS. The tree-thinking challenge. Science. 2005;310:979–80.
Catley KM, Novick LR. Seeing the wood for the trees: an analysis of evolutionary diagrams in biology textbooks. Bioscience. 2008;58(10):976–87.
Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Measur. 1960;20(1):37–46.
Cohen-Schotanus J. Student assessment and examination rules. Med Teach. 1999;21(3):318–21.

Cortright RN, Collins HL, Rodenbaugh DW, DiCarlo SE. Student retention of course content is improved by collaborative-group testing. Adv Physiol Educ. 2003;27(3):102–8.

Dees J, Momsen JL, Niemi J, Montplaisir L. Student interpretations of phylogenetic trees in an introductory biology course. CBE-Life Sci Educ. 2014;13:666–76.

Dobzhansky T. Nothing in Biology Makes Sense Except in the Light of Evolution. Am Biol Teach. 1973;35(3):125–9.

Eddy SL, Crowe AJ, Wenderoth MP, Freeman S. How should we teach tree-thinking? An experimental test of two hypotheses. Evol: Educ Outreach. 2013. 6:13.

Fisher RA. Statistical Methods for Research Workers. 5th ed. Edinburgh: Oliver and Boyd; 1934.

Freeman S. Biological Science. 4th ed. San Francisco: Benjamin Cummings; 2011.

Freeman S, O'Connor E, Parks JW, Cunningham M, Hurley D, Haak D, Dirks C, Wenderoth MP. Prescribed Active Learning Increases Performance in Introductory Biology. CBE-Life Sci Educ. 2007;6:132–9.

Gregory TR. Understanding evolutionary trees. Evol: Educ Outreach. 2008;1:121–37.

Halverson KL. Improving tree-thinking one learnable skill at a time. Evol: Educ Outreach. 2011;4:95–106.

Halverson KL, Pires CJ, Abell SK. Exploring the complexity of tree thinking expertise in an undergraduate systematics course. Sci Educ. 2011;95:794–823.

Johnson DW, Johnson RT, Smith KA. Cooperative learning returns to college: what evidence is there that it works? Change. 1998;30(4):26–35.

McDonald JH. Handbook of Biological Statistics. 3rd ed. Baltimore: Sparky House Publishing; 2014.

Meir E, Perry J, Herron JC, Kingsolver J. College students' misconceptions about evolutionary trees. Am Biol Teach. 2007;69(7):71–6.

Morabito NP, Catley KM, Novick LR. Reasoning about evolutionary history: post-secondary students' knowledge of most recent common ancestry and homoplasy. J Biol Educ. 2010;44(4):166–74.

National Research Council. Discipline-based education research: understanding and improving learning in undergraduate science and engineering. Washington: The National Academies Press; 2012.

Novick LR, Catley KM. Understanding Phylogenies in Biology: the Influence of a Gestalt Perceptual Principle. J Exp Psychol: Appl. 2007;13(4):197–223.

Novick LR, Catley KM. Reasoning about evolution's grand patterns: college students' understanding of the tree of life. Am Educ Res J. 2013;50(1):138–77.

Novick LR, Stull AT, Catley KM. Reading phylogenetic trees: the effects of tree orientation and text processing on comprehension. Bioscience. 2012;62(8):757–64.

Omland KE, Cook LG, Crisp MD. Tree thinking for all biology: the problem with reading phylogenies as ladders of progress. BioEssays. 2008;30(9):854–67.

Smith KA. Going deeper: formal small-group learning in large classes. New Dir Teach Learn. 2000;81:25–46.

Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval. 2006;27(2):237–46.

Wormald BW, Schoeman S, Somasunderam A, Penn M. Assessment drives learning: an unavoidable truth? Anat Sci Educ. 2009;2(5):199–204.

Young AK, White BT, Skurtu T. Teaching undergraduate students to draw phylogenetic trees: performance measures and partial successes. Evol: Educ Outreach. 2013. 6:16.