Inference is Bliss: Using Evolutionary Relationship to Guide Inferences about Biological Categories

Laura R. Novick (Laura.Novick@vanderbilt.edu)
Dept. of Psychology & Human Dev., Vanderbilt University, 552 GPC, 230 Appleton Place
Nashville, TN 37203 USA

Kefyn M. Catley (kcatley@wcu.edu)
Department of Biology, Western Carolina University
Cullowhee, NC 28723 USA

Abstract

College students with stronger and weaker backgrounds in biology were asked to infer whether one taxon (e.g., a certain reptile) uses the same enzyme to help regulate cell function as comparison taxon A (e.g., a certain amphibian) or comparison taxon B (e.g., a certain mammal). Color pictures of the named taxa were provided. Inferences were made in one of three conditions: Two conditions gave subjects the evolutionary relationships among the taxa (amphibians, reptiles, and mammals in the example), in one of two diagrammatic formats; the third, baseline, condition did not provide such information. Subjects uniformly made the incorrect inference in the baseline condition. Providing the evolutionary diagrams had different effects depending on the diagrammatic format and the biology background of the subjects.

Keywords: Categorical inferences, evolutionary diagrams, cladograms, expertise, macroevolution.

Introduction

Category-Based Induction

One of the most important functions of categories is to enable people to "know" things, and to take corresponding action, by inference rather than by perception. As López, Atran, Coley, Medin, and Smith (1997, p. 252) noted, "humans evolved in a world where knowing whether an animal was an antelope or a lion was essential for their survival: they could eat the antelope, and they could be eaten by the lion. Accordingly, the human mind seems to have evolved to organize its knowledge of the natural world into sets of related categories (e.g., cheetahs and lions are kinds of felines, aardwolves and spotted hyenas are kinds of hyenas, and felines and hyenas are kinds of carnivores) and to use these relations to make inductions about those categories." By definition, such inferences are best guesses, not fact. Thus, for example, one might infer that parrots are more likely to share a biological property possessed by squirrels than one possessed by cobras when in fact the opposite inference is warranted because, evolutionarily, birds are more closely related to snakes than to mammals (http://tolweb.org/Amniota/14990). In the experiment reported here, we assessed whether college students are able to change the inferences they would normally make if they are given information about evolutionary relationship that is inconsistent with their folkbiological categories.

There is an extensive literature on the inferences that people make from one biological category to another (primarily though not exclusively involving animal categories). This research has examined a variety of factors that influence categorical inferences concerning biological categories, such as: (a) the similarity between the premise and conclusion categories (e.g., mouse and rat vs. fox and rat) and the typicality of the premise category as a member of the conclusion category (e.g., rabbit/mammal vs. porcupine/mammal; López et al., 1997; Rips, 1975), (b) the level of the premise category within the class inclusion hierarchy (e.g., bald eagle, eagle, bird, animal; Coley, Medin, & Atran, 1997), and (c) domain expertise or cultural background (e.g., fishermen or Mayan Amerindians compared with U.S. college students; Coley et al., 1997; López et al., 1997; Shafto & Coley, 2003).

This research on category-based inferences has been conducted within the frameworks of, and motivated by the perspectives of, cognitive psychology and cultural anthropology. In our research, we approach this topic from a different perspective, namely evolutionary biology.

Cladograms and Tree-Thinking

Although U.S. college students know a fair amount about the biological world, particularly types of animals that live in their local area or that are commonly found in zoos, their knowledge of and experience with biological categories is relatively impoverished compared with that of certain more homogeneous groups (e.g., Mayan Amerindians, the Menominee tribe in Wisconsin; Medin & Atran, 2004). It should not be too surprising, therefore, to discover major misconceptions and errors in the categorical knowledge of even very bright college students. For example, Morabito, Catley, and Novick (2009) asked students at Vanderbilt University whether lizards are evolutionarily more closely related to frogs or to mammals. The correct answer is *mammals* (http://tolweb.org/Terrestrial_Vertebrates/14952), but 91% of students with weaker backgrounds in biology and 74% of students with stronger backgrounds in biology incorrectly grouped lizards with *frogs*.

If students' knowledge of biological categories is incorrect, their inferences will likely reflect this flawed knowledge. For example, if students are asked whether the gecko shares character X with the horned frog or character

Y with the prairie dog, we would expect them to select character X because of their incorrect belief that lizards are more closely related to frogs than to mammals. Although in psychological experiments one may base inferences on whatever criteria one wishes, and one goal of cognitive psychology is to uncover those criteria, evolutionary biology provides an external yardstick for judging inference quality. Although one perhaps cannot deem one inference to be correct and another to be wrong, evolutionary biology does allow one to identify some inferences as better or more useful than others. For example, inferring which antivenin would be best for treating the bite of a venomous king brown snake based on its close evolutionary relationship to the red-bellied black snake is more likely to lead to a desirable outcome (survival!) than basing the choice of antivenin on the king brown snake's similar coloration to the western brown snake (Freeman, 2005).

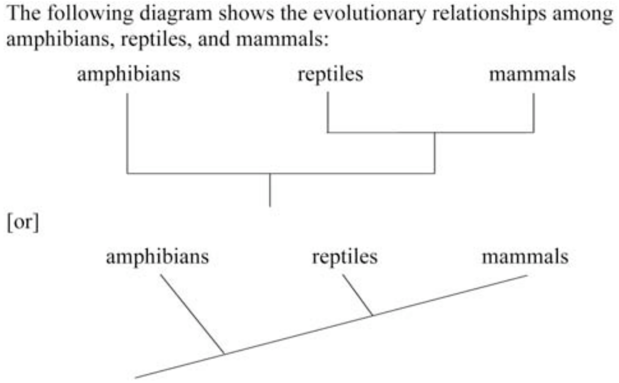


Figure 1: How the evolutionary relationship information was provided to subjects: The top and bottom cladograms are in the tree and ladder formats, respectively.

Within the field of evolutionary biology, evolutionary relationships are depicted as nested levels of most recent common ancestry in a hierarchical diagram known as a cladogram. For example, the cladograms in Figure 1 show that reptiles and mammals share a more recent common ancestor with each other than either of those two taxa share with amphibians. It is more parsimonious, therefore, that the gecko shares a character with the prairie dog than with the horned frog because geckos and prairie dogs share a most recent common ancestor that is not shared by the horned frog. Because cladograms (a) are the most important tool used by evolutionary biologists, (b) enable evidence-based inference and prediction, and (c) provide a conceptual framework for basic and applied biology (e.g., see Cracraft & Donoghue, 2004), a growing number of biologists and science educators have called for the inclusion of cladogram-based *tree thinking* in biology curricula from middle school through the undergraduate years (e.g., Baum, Smith, & Donovan, 2005; Catley, 2006; Goldsmith, 2003; Meir, Perry, Herron, & Kingsolver, 2007; Nickels & Nelson, 2005; Novick & Catley, 2007, 2009; O'Hara, 1988; Sandvik, 2008).

Cladogram Formats

Let us return to the inference question about whether the gecko shares a property possessed by the prairie dog or one possessed by the horned frog. As noted earlier, because of students' misconception concerning the pattern of evolutionary relationships among these taxa (Morabito et al., 2009), we would expect them to choose the property of the horned frog. The question we ask in the present study is whether these same students are able to use information about the correct evolutionary relationships to overcome their incorrect categorical knowledge and make the more appropriate, evidence-based inference that the gecko shares the property possessed by the prairie dog.

As shown in Figure 1, cladograms are typically drawn in one of two formats, which we have previously labeled *tree* and *ladder*. Although these formats are informationally equivalent (Larkin & Simon, 1987), our research on students' understanding of and ability to reason from cladograms shows that they are not computationally equivalent. It is much more difficult to extract the correct hierarchical relationships from the ladder format than the tree format (Novick & Catley, 2007, 2009). Although students with stronger backgrounds in biology perform better than those with weaker backgrounds on a variety of cladogram comprehension and reasoning (i.e., tree-thinking) tasks when the ladder format is used, even they show decrements in performance with the ladder format compared with the tree format (Novick & Catley, 2007, 2009).

Overview of the Study

Our earlier studies used cladograms with 5-10 taxa. The present study examined whether students can use the simplest possible tree and ladder cladograms, involving only three taxa, to make appropriate, evidence-based inferences. Although we expect the tree format to be more helpful, consistent with our earlier results, it is possible that students can understand the ladder format when it is pared down to its minimal version even though they cannot scale up this understanding to more complex structures. The results of this study on students' baseline competence at tree thinking will inform instructional interventions.

Subjects participated in one of three conditions: The no diagram condition gave us baseline data on subjects' inferences. In the two diagram conditions, subjects received cladograms depicting the correct evolutionary relationships among the relevant taxa, in either the tree format or the ladder format. As shown in Figure 1, the taxon labels on the cladograms were superordinate to those used in the inference question. Thus, subjects had to integrate the diagrammatic information showing that reptiles share a more recent common ancestor with mammals than with amphibians with their knowledge in long-term memory concerning the biological categories to which geckos, prairie dogs, and horned frogs belong. This integration is critical to preserve the inferential nature of the task.

Following López et al. (1997) and Gelman (1986), the inference questions were presented as sets of three taxa (triads), with subjects being asked whether the focal taxon shared a property with reference taxon A or reference taxon B. The property we used was reliance on a certain enzyme to help regulate cell function. The psychological literature on biological inferences has used a variety of different properties, including obscure body parts, behavioral characteristics, genes, enzymes, and diseases. We selected enzymes as the property of interest because they are likely to be shared due to recent common ancestry. In contrast, shared behavioral characteristics are often convergently evolved, and diseases may be passed on by ecological mechanisms such as predation. Obscure body parts would have served equally well, except for the difficulty of identifying such characters for all the different taxa we used.

There were two sets of experimental inference triads. One set involved amphibians, reptiles, and mammals. As already noted, our previous data (Morabito et al., 2009) strongly suggest that subjects will make the inference that contradicts evolutionary relationships in the control condition, thereby leaving room for improved performance when a cladogram is provided. The other set involved plants, fungi, and animals. Fungi are more closely related to animals than to plants, but the second author's experience teaching zoology to biology majors led us to believe that our subjects would make the inappropriate inference in the control condition for this set as well.

Method

Subjects

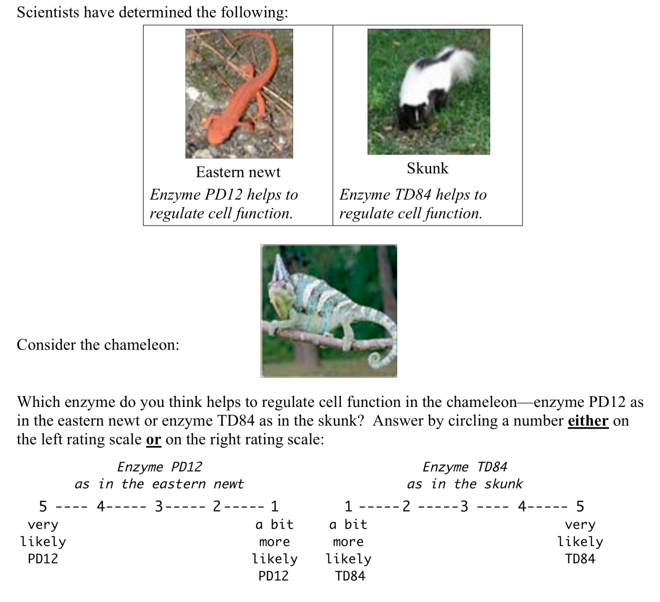
The subjects were 108 Vanderbilt University students who were recruited during two consecutive fall semesters. They participated in partial fulfillment of course requirements for introductory psychology (10 females, 5 males) or evolutionary biology (25 females, 27 males), or for extra credit in the psychology (21 females, 2 males) or education (15 females, 3 males) class from which they were recruited.

The subjects were divided into two groups based on their background in biology. Stronger background subjects had taken at least the two-semester introductory biology sequence required for biology majors and pre-med students. Weaker background subjects had not completed these two classes and typically had little exposure to biology. There is a very large difference in biology background between the two groups, as can be seen in the mean number of courses taken (of those listed on our questionnaire). The 54 stronger background students (26 females, 28 males) had taken an average of 3.31 semesters of biology (or relevant geology) classes, whereas the 54 weaker background students (45 females, 9 males) had taken an average of only 0.46 semesters of such coursework, a more than 7:1 difference.

Inference Triads

Each inference question involved two reference taxa and a focal taxon about which an inference was to be made. There were two types of taxon triads. *Reptile* triads used an amphibian and a small mammal as reference taxa and a reptile as the focal taxon. The amphibian was either a salamander/newt or a frog. The reptile either lived on land (lizard, snake, or turtle) or in the water (crocodilian, snake, or turtle). These two factors were crossed, and two triads were created for each of the four combinations. Figure 2 shows an example reptile triad in the no diagram control condition. As shown in this figure, all taxa were named and represented by a color picture. *Fungus* triads used a small flowering plant and an animal as the reference taxa and a mushroom as the focal taxon. Five types of animals were used (fish, invertebrate, large carnivore, primate, and reptile), with two triads created for each type of animal.

For each triad, subjects were told that a different enzyme helps to regulate cell function in each of the two reference taxa. The enzyme names consisted of a randomly generated consonant pair followed by a randomly generated number pair (see Figure 2). Subjects were then told to consider the focal taxon and to decide which of the two enzymes helps to regulate cell function in that taxon. If they preferred the enzyme of the reference taxon pictured on the left, they circled a number on the left rating scale; if they preferred the enzyme of the reference taxon pictured on the right, they circled a number on the right rating scale. The left and right positions were correct (i.e., supported by evidence concerning evolutionary relationships) equally often.



Scientists have determined the following:

Eastern newt *Enzyme PD12 helps to regulate cell function.*	Skunk *Enzyme TD84 helps to regulate cell function.*

Consider the chameleon:

Which enzyme do you think helps to regulate cell function in the chameleon—enzyme PD12 as in the eastern newt or enzyme TD84 as in the skunk? Answer by circling a number **either** on the left rating scale **or** on the right rating scale:

```
      Enzyme PD12                          Enzyme TD84
  as in the eastern newt                 as in the skunk
 5 ---- 4----- 3----- 2----- 1      1 -----2 -----3 ---- 4----- 5
very                      a bit    a bit                      very
likely                    more     more                      likely
PD12                      likely   likely                     TD84
                          PD12     TD84
```

Figure 2: A reptile triad in the no diagram condition.

The 8 reptile triads and 10 fungus triads were mixed with 32 filler triads that involved a variety of different kinds of animals and plants. The filler triads served to disguise the critical trials. The data from those triads were not analyzed.

Cladogram Format Conditions

As noted earlier, the most appropriate inference depends on the evolutionary relationships among the taxa. Subjects should infer that the focal taxon uses the same enzyme as the reference taxon with which it shares a more recent common ancestor. For the reptile triads, that is the mammal. For the fungus triads, that is the animal. In the no diagram condition, as illustrated in Figure 2, the inference question was presented without providing the relevant evolutionary information. This condition, then, provides data on subjects' entering biases concerning taxon relationships.

In two other conditions, we gave subjects a cladogram that depicted the relevant evolutionary relationships. This diagram was located immediately above the information depicted in Figure 2. The cladogram was drawn in one of two formats, either tree or ladder. The two cladograms used with the triad shown in Figure 2 are depicted in Figure 1. Subjects received only one of these two diagrams.

Subjects were randomly assigned to cladogram format conditions. There were 17, 18, and 19 stronger background students and 17, 21, and 16 weaker background students in the no diagram, ladder, and tree conditions, respectively.

Procedure

The 50 inference problems were randomly ordered subject to certain constraints to make sure that critical triads of a given type (reptile or fungus) did not occur on consecutive trials and that the critical triads of each type were distributed throughout the booklet. Two different random orders were used. The problems were printed two to a page. Subjects worked through the booklet at their own pace.

Results

The data from the two types of triads were analyzed separately because different variables were manipulated in their construction. For both types of triads, subjects' ratings for the evolutionarily-inappropriate taxon choice were coded on a scale from -5 (very likely) to -1 (a bit more likely); their ratings for the evolutionarily-appropriate taxon choice were coded on a scale from 1 (a bit more likely) to 5 (very likely). Thus mean negative ratings indicate a preference for the more-distantly-related taxon (amphibian choices for the reptile triads and plant choices for the fungus triads), mean ratings near 0 indicate no preference for one reference taxon over the other, and mean positive ratings indicate a preference for the more-closely-related taxon (mammal choices for the reptile triads and animal choices for the fungus triads).

Effect sizes are reported as partial η^2. A value of .01 was the minimum taken to indicate a small effect, .09 was the minimum taken to indicate a medium effect, and .25 was the minimum taken to indicate a large effect.

Reptile Triads

We conducted a 2 (reptile type; within) X 2 (amphibian type; within) X 2 (biology background; between) X 3 (diagram condition; between) ANOVA on the mean ratings for the reptile triads. There was a main effect of type of reptile, $F(1, 102) = 8.11$, $p < .01$, $MSE = 0.93$, partial $\eta^2 = 0.07$. Subjects showed a stronger bias to pick the enzyme of the amphibian when the focal taxon was a land mammal ($M = -1.48$) than when it was a water mammal ($M = -1.22$). This might have been because our pictures of the amphibians showed them on land. There was also a main effect of type of amphibian, $F(1, 102) = 23.08$, $p < .001$, $MSE = 1.13$, partial $\eta^2 = 0.19$. Subjects showed a stronger bias to pick the enzyme of the amphibian when it was a salamander ($M = -1.60$) than when it was a frog ($M = -1.10$). We speculate that this is due to the greater similarity of locomotion between salamanders and reptiles than between frogs and reptiles. Neither the type of reptile nor the type of amphibian interacted with any other factor. Thus, we will not discuss these factors further.

There was a main effect of diagram condition, $F(2, 102) = 40.99$, $p < .001$, $MSE = 13.59$, partial $\eta^2 = 0.45$. As predicted, the ratings were strongly negative in the no diagram condition ($M = -3.10$), less negative when the correct evolutionary relationships were shown in the ladder format ($M = -1.85$), and positive when the evolutionary relationships were depicted in the tree format ($M = 0.91$). The main effect of biology background was marginally significant, $F(1, 102) = 2.96$, $p < .09$, partial $\eta^2 = 0.03$.

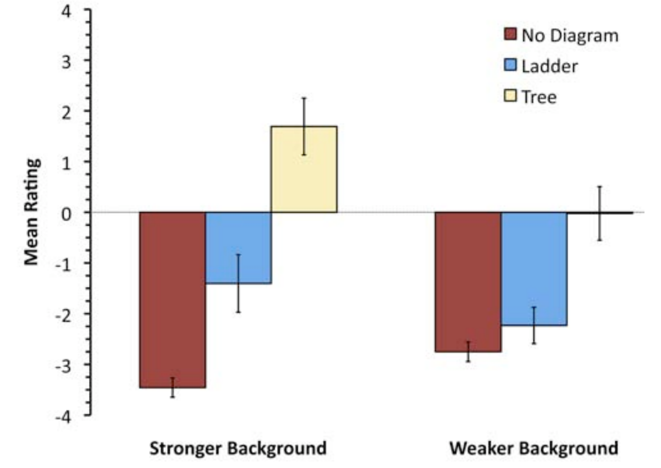


Figure 3: Mean inference ratings for the reptile triads as a function of biology background and diagram condition. Negative numbers indicate (incorrect) amphibian choices; positive numbers indicate (correct) mammal choices. Error bars are standard errors for the relevant mean.

These main effects were qualified by a diagram condition by biology background interaction, $F(2, 102) = 3.79$, $p < .03$, partial $\eta^2 = 0.07$, which is depicted in Figure 3. As shown in the figure, the overall pattern discussed for the main effect of diagram condition is largely driven by the results for the stronger background subjects. They had a strong negative rating in the no diagram condition, a less negative rating in the ladder condition, and a positive rating in the tree condition. The weaker background subjects, in contrast, had very similar ratings in the no diagram and ladder conditions, and in the tree condition their mean rating was essentially 0, indicating no preference for placing the reptile with either the amphibian or the mammal.

Fungus Triads

We conducted a 5 (animal taxon; within) X 2 (biology background; between) X 3 (diagram condition; between) ANOVA on the mean ratings for the fungus triads. The main effect of type of animal taxon was not statistically significant, *F*(4, 408) = 1.64, *p* > .16, *MSE* = 0.79, partial η^2 = 0.02. Type of animal taxon did not interact with any other factor. Therefore, it will not be discussed further.

The results for the between-subjects factors were very similar to those found for the reptile triads. There was a main effect of diagram condition, *F*(2, 102) = 37.66, *p* < .001, *MSE* = 15.35, partial η^2 = 0.43. As predicted, the ratings were most negative in the no diagram condition (*M* = -2.04), less negative when the correct evolutionary relationships were presented in the ladder format (*M* = -1.16), and positive when the evolutionary relationships were depicted in the tree format (*M* = 1.54). The main effect of biology background was marginally significant, *F*(1, 102) = 3.70, *p* < .06, partial η^2 = 0.03. The interaction between diagram condition and biology background also was marginally significant, *F*(2, 102) = 2.51, *p* < .09, partial η^2 = 0.05. Figure 4 shows that the form of this interaction is similar to that found for the reptile triads, with all the means shifted in a positive direction.

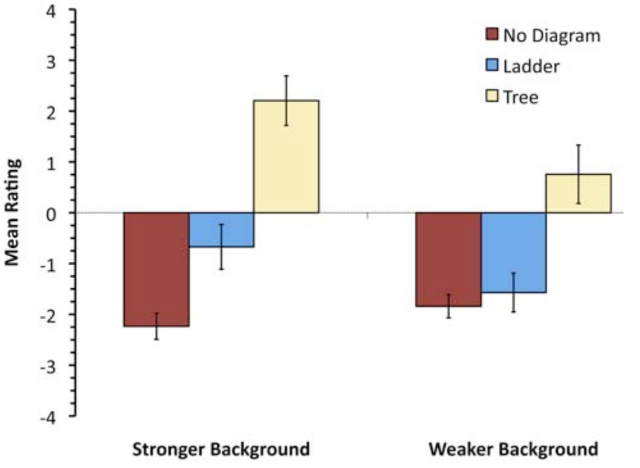


Figure 4: Mean inference ratings for the fungus triads as a function of biology background and diagram condition. Negative numbers indicate (incorrect) plant choices; positive numbers indicate (correct) animal choices. Error bars are standard errors for the relevant mean.

Discussion

As expected, both weaker and stronger background subjects made the poorer inference in the no diagram control condition, strongly believing that the reptile would share a cell-function-regulating enzyme with the amphibian and that the mushroom would share such an enzyme with the plant. If anything, this tendency to make an incorrect inference was stronger in the students who were more knowledgeable about biology.

Providing subjects with the correct evolutionary relationships among the taxa had different effects depending on the format in which the cladogram was presented and the biology background of the subjects. When the cladogram was depicted in the ladder format, the responses of weaker background students were unchanged. Three explanations for this finding are possible: (a) they did not understand the diagram and so ignored it, (b) they misinterpreted the ladder to indicate that reptiles are more closely related to amphibians and fungi are more closely related to plants, and (c) they misinterpreted the ladder to indicate that the three taxa are equally closely related and so responded on some other basis. Although, we cannot discriminate among these alternatives based on the present data, our previous data (Novick & Catley, 2007, 2009) suggest that these students may interpret three-taxon ladders as indicating that all three taxa are equally closely related. This would mean that the ladders used in the present study provided them with no relevant information for choosing one reference taxon over the other, leading to the results we found. Presenting the cladograms in the tree format did help the weaker background subjects, at least in a relative sense: For both the reptile and fungus triads, the bias to respond incorrectly disappeared. However, for neither type of triad did weaker background subjects have a mean rating that was significantly above zero, indicating that these subjects never consistently made the evolutionarily-appropriate inference, even when the relevant evolutionary evidence was provided.

As might be expected, the stronger background subjects fared better with the cladograms. When given the evolutionary relationship information in the tree format, they switched from making the evolutionarily-inappropriate inference to making the evolutionarily-appropriate inference for both the reptile and fungus triads. For the fungus triads, these subjects were as sure that the animal enzyme was the correct inference in the tree condition (*M* = 2.21) as they had been in the no diagram condition that the plant enzyme was the correct inference (*M* = –2.23). The incorrect inference for the reptile triads was more difficult to overcome, as the mean rating in the tree condition (*M* = 1.69) was less in absolute value than the mean rating in the no diagram condition (*M* = –3.46). These data strongly suggest that the incorrect reptile/amphibian link is deeply entrenched. Providing the evolutionary relationship information in the ladder format reduced stronger background subjects' tendency to make an incorrect inference for both types of triads, but their mean ratings were still negative in both cases. This result is consistent with our earlier research showing that even students with stronger backgrounds in biology have difficulty understanding the structure of the ladder format.

These results have important implications for instruction. Cladograms are powerful tools for inference because they depict the structural information upon which biological inferences should be based. If students understand this function of cladograms, they should make better inferences when they receive a cladogram than when they do not. Our results suggest that college students understand this function of cladograms to some extent. Clearly, though, there is room for improvement, as subjects reliably chose the evolutionarily-appropriate inference in only one of the four relevant cells of our design even though the cladograms we used are the simplest ones possible.

Although we have argued that non-biology majors do not need to understand the ladder format (Novick & Catley, 2007), biology majors should because it appears in the research literature (although less frequently than the tree format). In collaboration with Dan Funk in the biology department at Vanderbilt University, we are examining the effects of instruction in phylogenetics (i.e., tree thinking) on the ability of students enrolled in an evolution class to use the evolutionary relationship information provided in the tree and ladder formats to make appropriate inferences.

College students with weaker backgrounds in biology, and even some with stronger backgrounds, need help understanding and using cladograms in the tree format. We are currently evaluating the effectiveness of an instructional booklet for improving tree-thinking skills in students with varying backgrounds in biology.

In a third extension of the present findings, we are investigating the effects of providing evolutionary relationship information in the tree format on the inferences made by tenth graders enrolled in a high school biology class. The results of all three of these studies in progress will help guide us in the development of a tree-thinking curriculum at both the college and high school levels.

Acknowledgments

We thank Dan Funk at Vanderbilt University for allowing us to collect data from the students in his evolution classes across two semesters. We thank John Coley at Northeastern University for giving us triads to use for the filler trials. We thank Julie Denham and Caryn Wang for their help in creating the stimulus pages; Amy Garza, Courtney Shade, Marissa Mencio, and Emily Schreiber for their help with collecting the data; and Amy Garza, Marissa Mencio, and Aaron Novick for creating and checking the data file. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A080621 to Vanderbilt University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education

References

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

Catley, K. M. (2006). Darwin's missing link: a new paradigm for evolution education. *Science Education, 90,* 767-783.

Coley, J. D., Medin, D. L., & Atran, S. (1997). Does rank have its privilege? Inductive inferences within folkbiological taxonomies. *Cognition, 64,* 73-112.

Cracraft, J. & Donoghue, M. J. (Eds.) (2004). *Assembling the tree of life*. New York, NY: Oxford University Press.

Freeman, S. (2005). *Biological science* (2nd ed.). Upper Saddle River (NJ): Prentice Hall.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23,* 183-209.

Goldsmith, D.W. (2003). Presenting cladistic thinking to biology majors & general science students. *The American Biology Teacher, 65,* 679- 682.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65-99.

López, A., Atran, S., Coley, J. D., Medin, D. L., & Smith, E. E. (1997). The tree of life: Universal and cultural features of folkbiological taxonomies and inductions. *Cognitive Psychology*, *32*, 251-295.

Medin, D. L., & Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. *Psychological Review, 111,* 960-983.

Meir, E., Perry, J., Herron, J. C., & Kingsolver, J. (2007, September). College students' misconceptions about evolutionary trees. *The American Biology Teacher Online, 69 (7).*

Morabito, N., Catley, K. M., & Novick, L. R. (2009). *Seeing both sides of the coin: The effects of biology background on university students' knowledge of most recent common ancestry and homoplasy*. Manuscript under revision.

Nickels, M. K., & Nelson, C. E. (2005). Beware of nuts & bolts: Putting evolution into the teaching of biological classification. *The American Biology Teacher, 67,* 283-289.

Novick, L. R., & Catley, K. M. (2007). Understanding phylogenies in biology: The influence of a Gestalt perceptual principle. *Journal of Experimental Psychology: Applied, 13,* 197-223.

Novick, L. R., & Catley, K. M. (2009). *Understanding the tree of life: The effects of biology background, cladogram format, and cognitive/perceptual factors on tree thinking.* Manuscript under revision for an invited resubmission.

O'Hara, R. J. (1988). Homage to Clio, or, toward an historical philosophy for evolutionary biology. *Systematic Zoology, 37,* 142-155.

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14,* 665-681.

Sandvik, H. (2008). Tree thinking cannot taken for granted: challenges for teaching phylogenetics. *Theory in Biosciences, 127,* 45–51. Retrieved from: http://www.springerlink.com/content/eu62420p381402xr/ March 19, 2008.

Shafto, P., & Coley, J. D. (2003). Development of categorization and reasoning in the natural world: Novices to experts, naive similarity to ecological knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 641-649.