The Innovative Use of Mathematica to Teach Biodiversity

ELIZABETH C. DAVIS-BERG, DANIEL R. JORDAN

ABSTRACT

We investigated using Mathematica to create an interactive exercise that allows students to explore biological concepts. Exporting these active learning experiences to the Wolfram Computable Document Format (CDF) lets students enter data and modify parameters with only a browser plug-in and without needing to know how to use Mathematica. This exercise enhances student understanding by allowing them to focus on the biological concepts and visualize the data and calculations without the tedium of hand computation. In this article, we will highlight an interactive version of a biodiversity laboratory that explores Simpson's Diversity Index. Students are walked through sample calculations with a small data set, and then analyze a larger data set, which is presented as a set of interactive graphs and automatic calculations so that they can explore the data to answer questions. For our non-major audience, this laboratory leads to a more sophisticated discussion without getting stuck on the calculations. We will present the exercise and assessment results.

Key Words: *biodiversity; inquiry; technology; mathematical reasoning; ecology.*

○ Introduction

Teaching mathematics-based concepts in a biology classroom is challenging. Particularly in classes for non-majors, students may have weak mathematics skills. Those students need not only to learn the new biology concepts, but also to address the relevant deficiencies in their mathematics background. This increases their cognitive load, making already difficult material vastly more so (Sweller, 1994). This activity aims to decrease students' cognitive load by designing a learning experience that helps first to focus students' attention on the mechanics of the needed computation, and then to relieve students of the need to attend to the mathematical content, thereby facilitating their focus on the biology concepts. In addition, students learn better when presented with active learning and inter-active experiences (Prince, 2004).

One concept that students often struggle with in general biology and ecology courses is biodiversity and calculating biodiversity indices. Biodiversity is a very important topic in most biology, ecology, and environmental science classes because our current rate of biodiversity loss may be high enough relative to background extinction rates to be considered a sixth mass extinction event (Ceballos et al., 2015; Hull et al., 2015; Waters et al., 2016). There are many different types of biodiversity indices, and this exercise could be modified to support various learning objectives (Bandeira et al., 2013).

Quantitative reasoning is a core part of the AAAS report, *Vision and Change in Undergraduate Biology Education* (AAAS, 2011). Undergraduates do not have strong mathematics skills, and in fact, those skills are declining (Mulhern & Wylie, 2004). To achieve the goals of incorporating additional quantitative reasoning elements in an undergraduate non-majors biology course, bolstering student understanding of biodiversity and biodiversity indices, and supporting students for whom the additional mathematics content presents an additional barrier to learning, we created an experience to help students visualize and understand both the biology and mathematics behind Simpson's Diversity Index (Simpson, 1949). Students often struggle with understanding how indices and other mathematical equations relate to the biology. This exercise aims to make the calculation easier to understand and the effects of changing the environment on the index visual and clear.

To create this lab, we examined other published diversity labs (Boyce, 2005) and textbooks (Roach et al., 2010; Smith & Smith, 2012) to determine how others were teaching this material. We wanted to create an activity that would complement the materials in the books we were using in the course (Roach et al., 2010; Smith & Smith, 2012). The lab's interactive elements were created in Mathematica and then exported to the Wolfram Computable Document Format (CDF), which allows students to use and interact with the lab using a free browser plug-in or stand-alone player application (https://www.wolfram.com/cdf-player/). Distributing the lab as a CDF also has the benefit that students do not need to be concerned with formal Mathematica code.

Course Context

This exercise was created for undergraduate students in non-majors or introductory biology courses, but could easily be adapted to upper-level undergraduate or to high-school courses. It can be used on any computer that has the Wolfram CDF Player or browser plug-in installed, and takes students about two to three hours to complete. We have tested it in both Ecology and Environmental Science classes. These courses take place over a fifteen-week semester, with each week consisting of either a three-hour combination lecture and laboratory block, or two, two-hour meetings per week, with one meeting dedicated to a laboratory section. Typically, our laboratory classes have a cap of twenty students per section. Students' majors include art and communication fields, including film, journalism, art, business, and music.

How to Do it

Students work in pairs within the learning management system (LMS) at a computer that has the CDF browser plug-in installed. The interactive CDF elements, together with background information, definitions, instructions, and other lab elements, are embedded within an activity in the LMS. This allows students to interact with the CDF elements and answer questions, including multiple-choice, short-answer, and essay, within the LMS. Alternatively, the entire lab, including both the interactive elements, text, and questions, can be assembled into a single CDF document and used with the stand-alone CDF Player application, but that requires students to answer questions on paper or in a separate document.

The Exercise

Learning Goals and Objectives

The objectives of this exercise were to (1) understand the basics of quantifying biodiversity using species richness, (2) graph and interpret a rank abundance curve, (3) calculate Simpson's Diversity Index, and (4) interpret and answer questions about a larger data set.

Vocabulary

We defined species richness, relative abundance, a rank abundance diagram, species evenness, and diversity indices in the introduction, as well as in the glossary of the LMS. We also provided the following formula:

Simpson's Diversity Index is $\frac{1}{\sum_i p_i^2}$, or equivalently $\frac{1}{\sum_i \left(\frac{n_i}{N}\right)^2}$, where

- n_i is the number of individuals of species i,
- N is the total number of individuals of all species,
- p_i is the proportional abundance of species i and is defined by $p_i = \frac{n_i}{N}$, and
- $\sum_i p_i^2$ means that you add over all species.

Background Information for Students

"Is it pristine?" is often a question asked about natural environments. When planning parks, managers want to know if the park made an environment more natural. But these questions are really addressing the health of the community in an area. To figure out if a community is healthy, you can measure the biodiversity. Generally, polluted or heavily human-impacted areas will have lower levels of biodiversity than more pristine areas.

So, how do you measure biodiversity? The simplest way is to count up how many individuals of each species are in a specific area. If you are assessing a large area, this is usually done with lots of smaller samples. Your results will be a table of species with counts of individuals. From this table you can calculate the *relative abundance* (p_i) of each species.

Once you have the relative abundance calculations, you can graph this as a *rank abundance curve*. To do this, you rank the species by number of individuals (rank 1 has the most individuals), and then graph the rank on the x-axis and the relative abundance on the y-axis.

Then you can calculate Simpson's Diversity Index. This will give you a number that has a minimum value of 1. If the Simpson's Diversity Index equals 1, then there is only one species in the community. The maximum value is the number of species in the community (*species richness*). This exercise will have you complete a simple calculation by hand and then interpret a larger community afterward.

Overview: The first section of the exercise will walk you through creating a rank abundance curve and calculating Simpson's Diversity Index using sample data. The second section will let you explore real data from a deciduous forest in Virginia. You will be able to manipulate the data to see how changes in the forest would affect the diversity index and graphs.

The Protocol

Preliminary Protocol Notes For Instructors

We started with a sample calculation, using a hypothetical community with six species in a table. After showing students a table with only the numbers of individuals of six species, we then explained that the rank of a species refers to its position in the list when the species are ordered from most abundant to least abundant, with the most abundant species having rank 1, the next most abundant species having rank 2, etc. We showed them the ranked data (Table 1). Then, as part of the CDF, students computed the *proportional abundance*, or the percentage of the total number of individuals, of each species.

Rank Abundance Curves (for students)

1. Compute N, the total number of individuals of all species, by adding the numbers in the *Number of Individuals* column.

Table 1. The species are sorted according to rank so species A has rank 1, and species F has rank 6.

Species	Number of individuals	Rank
A	73	1
B	15	2
C	11	3
D	8	4
E	3	5
F	1	6

Mathematica will check your work, so please press the Check button after entering your result to see if you did the calculation correctly (Figure 1A).

2. Now, for each species, we will compute the proportional abundance, $p_i = \frac{n_i}{N}$, where n_i is the number of individuals of the species. Enter your answer rounded to the thousandth place (3 digits after the decimal point). For example, $p_1 = \frac{73}{111} = 0.658$.

3. Complete the table with the proportional abundance for the other species. Again, check your work by pressing the Check button. It will check each answer individually. Do not continue until all your answers are correct (Figure 1B).

4. The rank abundance curve is the graph of rank versus the proportional abundance. The rank abundance curve has been started for you. Complete it by clicking and dragging the points for ranks 2 through 6 to the appropriate positions (Figure 2).

Simpson's Diversity Index (for students)

1. Simpson's Diversity Index is a number that represents the diversity, in terms of species, of an area. It is calculated using the formula $D_S = \frac{1}{\sum_i p_i^2}$. You just computed the proportional abundance for each species in the sample data. Those results are duplicated in the next table.

2. To compute Simpson's Diversity Index, first square each of the proportional abundance values. For example, for species A, the proportional abundance was $p_1 = 0.658$, so $p_1^2 = 0.4330$ (rounded to the nearest ten-thousandth, i.e., 4 digits after the decimal point) (Figure 3A).

3. Now, add the p_i^2 values and finally compute Simpson's Diversity Index by dividing 1 by that sum, rounding your answer to the nearest thousandth (Figure 3B).
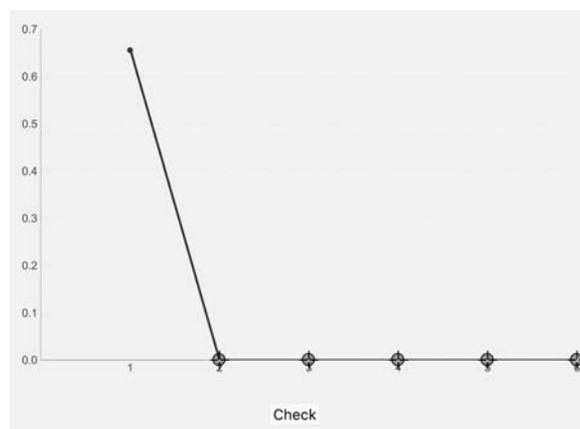
Evaluation Questions (for students)

1. How does the slope of the rank abundance curve vary with increasing species evenness? Why?

2. Species diversity measurements take into account both species richness and species evenness. Why would these measurements be preferred to species richness alone?

3. What do rank abundance curves add to one's knowledge about community structure?

4. You can make up your own sample data in the table, and the rank abundance curve and Simpson's Index will be computed for you. How does the rank abundance curve change if all the species are equally abundant? (Figure 4)

5. The table in Figure 5 shows data obtained from a mature deciduous forest stand in northern West Virginia, together with a rank abundance curve and the value of Simpson's Diversity Index (data from Smith & Smith, 2012). You can move the sliders to change the numbers of individuals of each species, and see the effect on the curve and the index

Figure 2. This image shows how the rank abundance graph appeared to students. They practiced graphing using this interactive element.

Figure 1. Screen shots of the proportional abundance calculation. **(A)** Step 1 of the methods and **(B)** step 2 of the methods as described. We have shown the example with an incorrect answer.
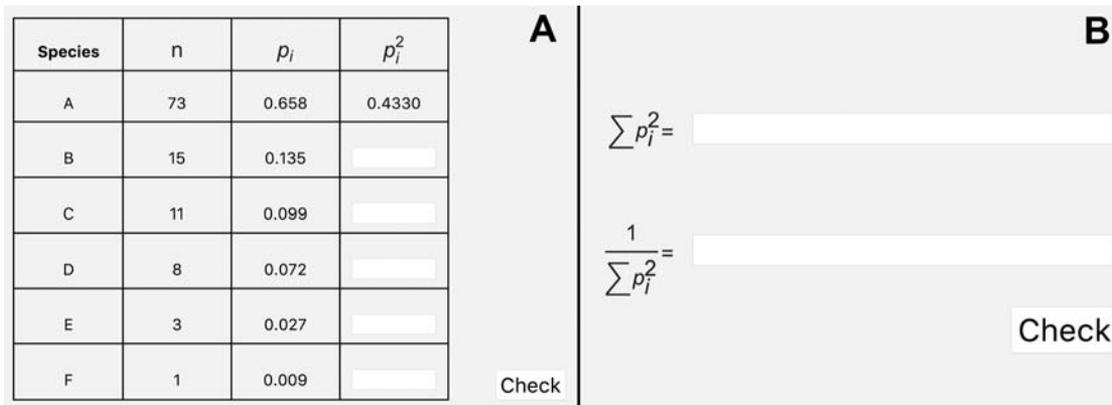
Figure 3. Simpson's Diversity Index was calculated using the table in part **(A)**. Then, once the correct numbers were obtained, we slowly led the students through the calculation of the Index itself in part **(B)**.
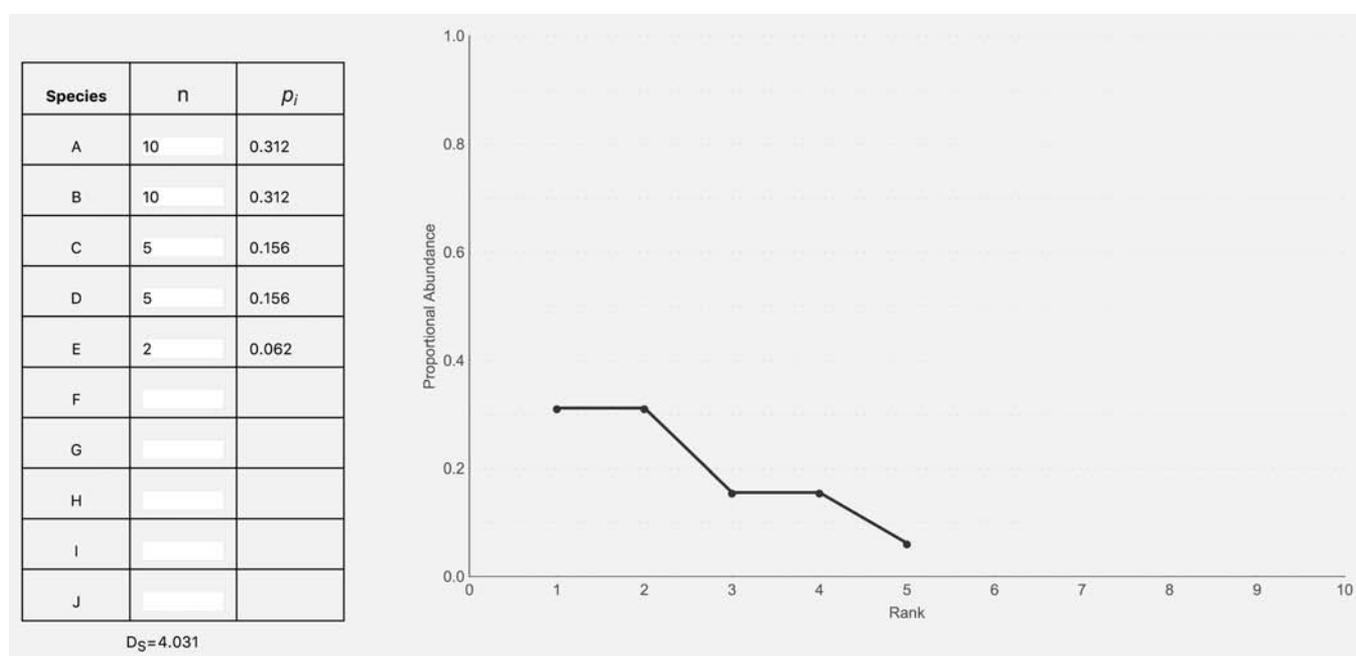

Figure 4. This interactive element, supporting Question 4, allowed students to create their own data and see both the rank abundance curve and the calculated Simpson's Diversity Index.

value (Figure 5). Use this table to answer the two questions below.

a Which species has the most impact on the community if removed? Why? Explain what you did and why you reached this conclusion. Be sure to explain your result with both the rank abundance curve and the Simpson's Diversity Index.

b Which species has the least impact on the community if removed? Why? Explain what you did and why you reached this conclusion. Be sure to explain your result with both the rank abundance curve and the Simpson's Diversity Index.

6. Why do ecologists use both diversity indices and rank abundance curves when interpreting their data?

Notes for Instructors on This Exercise

We used this exercise in conjunction with activities where students went outside, counted species, and created species accumulation curves—often plants but more recently Pokémon (Drew et al., 2017). In class, we presented community ecology data, and in small groups students determined how to assess the community (e.g., soil type, species diversity, species richness, species abundance).

The evaluation questions should be adjusted for the level of students and type of class. Some questions are designed to test critical thinking; for example, in Question 5b students may decide that the removal of a low-abundance species always has a small effect. However, if that species happened to be a predator, that could have a large long-term cascading effect on the community diversity, but only a small initial effect.

| Species	n	p_i
Yellow Poplar	76.	0.297
White Oak	36.	0.141
Black Oak	17.	0.066
Red Maple	14.	0.055
Sugar Maple	14.	0.055
American Beech	13.	0.051
Red Oak	12.	0.047
Sassafras	12.	0.047
Black cherry	11.	0.043
Mockernut hickory	11.	0.043
Slippery elm	10.	0.039
Shagbark hickory	7.	0.027
Bitternut hickory	5.	0.02
Flowering Dogwood	3.	0.012
Pignut hickory	3.	0.012
Cucumber magnolia	2.	0.008
Hornbeam	2.	0.008
White ash	2.	0.008
Tree of Heaven	1.	0.004
Sourwood	1.	0.004
Black locust	1.	0.004
Black mample	1.	0.004
Black walnut	1.	0.004
American elm	1.	0.004

$D_S = 7.566$

Reset
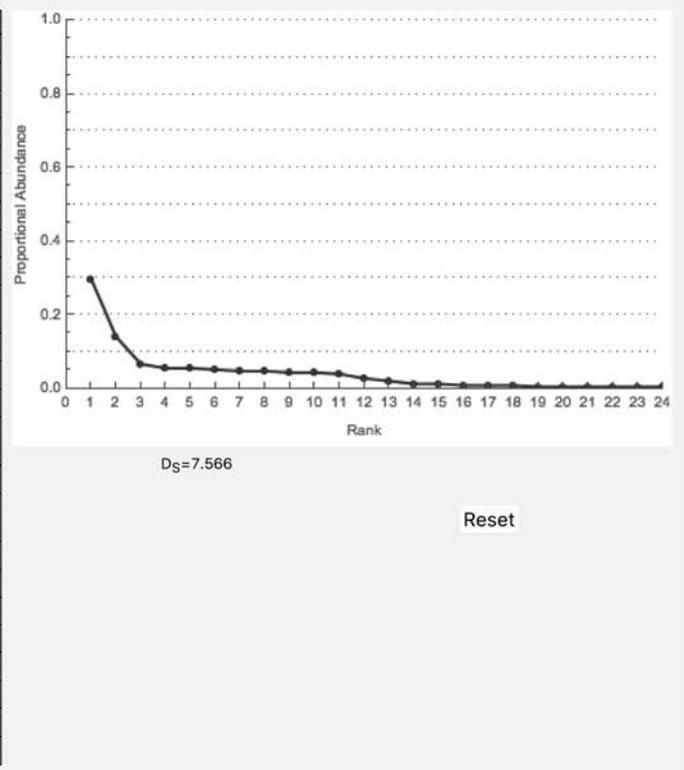
Figure 5. For Question 5, students can move the sliders to change the numbers of individuals of each species and see the effect on the rank abundance curve and Simpson's Diversity index.

○ **Assessment and Discussion**

The CDF was initially piloted in a section of Principles of Ecology in early fall 2013 and revised based on informal student feedback. Then, with IRB approval (IRB 2014-00171), during fall 2013 and spring 2014, we compared three experimental sections (CDF exercise) and two control sections (paper-and-pencil-based version) of Environmental Science and of Ecology (both non-majors courses). In all sections, we gave a pre- and post-survey to assess student preparedness and understanding; students answered using a 5-point Likert scale (Table 2). We analyzed the numerical results of the surveys using Microsoft Excel, R, and Mathematica. The post-survey included four free-response questions. We identified the topics students commented on in their responses (e.g., math, interactivity, etc.) and counted the number of students whose comments included each topic in each question.

Both groups responded more favorably in the post-survey as compared to the pre-survey. For example, for the control section, a Wilcoxon signed-rank test indicated that the students reported more confidence in their ability to compute Simpson's Diversity Index on the post-survey than on the pre-survey ($V = 148$, $p < .001$). On the same question, the experimental group also reported more confidence on the post-survey compared to the pre-survey ($V = 147.5$, $p < .001$). A Mann-Whitney test found no statistically significant difference in the post-test responses between the groups ($U = 251$, $p = .062$) (Figure 6). These numerical results suggest both that the CDF exercise is effective in promoting student learning and that the use of technology did not become a distraction or otherwise detract from students' learning.

The interactive CDF lab was just as effective in promoting learning, but also, based on instructors' observations and responses to the open-ended survey questions, increased student engagement. Students enjoyed interactive graphs and having the software check their work, with one-sixth of the students who completed the CDF lab commenting that the interactivity was their favorite aspect. In fact, interactivity was the most frequently mentioned aspect of the lab mentioned in Question 9 (Table 2). Student comments included, "My favorite part was that the program explained everything very well and you could move the graphs," and "I liked the graphs and stuff it displayed to help us understand."

Not all students seemed as comfortable with computer-based labs as we had expected, although only one student specifically mentioned the computer as a least favorite aspect, and two included it as a positive. There was some evidence in the open-ended questions to suggest that the CDF format made the lab easier to follow. While two students from each group (paper-based [control] and CDF) commented in Question 10 that the instructions were confusing to them, two other students who completed the CDF version commented in Question 9 that the clarity of instructions was one of their favorite features, compared to none in the control group.

Even with the interactive exercise, we had comments like "I disliked the large equation," although math was mentioned as the least favorite aspect of the activity by only about a quarter of the students in the experimental section compared to half for the students with the paper version of the lab. Interestingly, the most dramatic pattern in the open-ended responses was that well over two-thirds of the students in the control group commented about mathematics

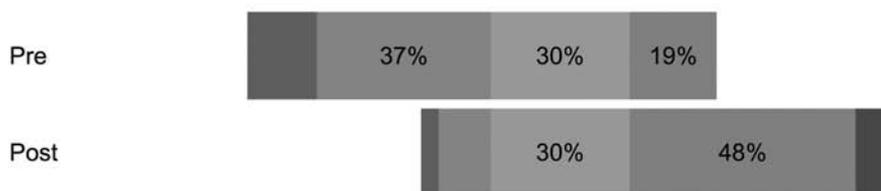**Control Sections**
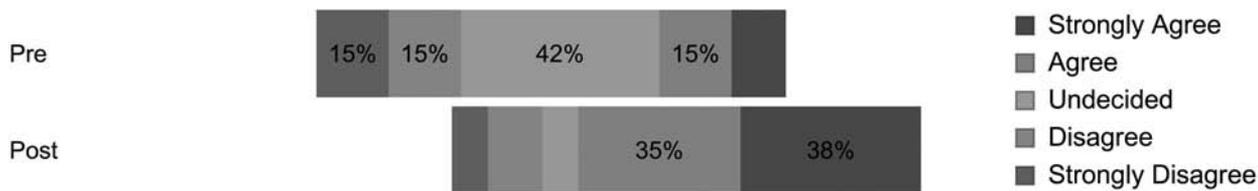
Experimental Sections

Figure 6. Results for the survey Question 4.

Table 2. Survey questions.

Pre-Test
1. I have carefully read the lab exercise.
2. I feel comfortable with the basic concepts of this exercise.
3. I feel confident in my ability to do the necessary computations in this exercise.
4. I feel confident that I can compute the value of the expression below where $p_i = \frac{n_i}{N}$ and n_i = number of individuals of *species i* and N = the total number of individuals of all species $$\frac{1}{\sum_i p_i^2}$$
5. I feel confident that I can compute the value of the expression below where $\mathbf{p_i} = \frac{\mathbf{n_i}}{\mathbf{N}}$ and n_i = number of individuals of *species i* and N = the total number of individuals of all species $$H' = -\sum_i p_i \ln p_i$$
6. I do not feel confident that I understand why rank abundance curves are important for interpreting biodiversity.
7. I feel confident that I understand why a diversity index is important when interpreting biodiversity.
Post-Test
Questions 1 through 7 above, and additionally:
8. This exercise improved my understanding of biodiversity and Simpson's Diversity Index.
Open-Answer Questions
9. What was your favorite part of this exercise?
10. What was your least favorite part of the exercise?
11. Please list any improvements you would make to this exercise.
12. Please list any comments you would like to make about this exercise.

in at least one of the three questions, while barely more than one-third of the students in the experimental group did. This suggests that the main effect of the CDF lab was to help students focus less on the mathematical computations and formulae and more on the biology.

○ Possible Modifications and other uses

We recommend creating a how-to-use-a-CDF exercise for students unfamiliar with computer-based work. You could easily make other versions to teach different math-intensive exercises (e.g., life tables, Lotka-Volterra equations, or a graphing tutorial). Also, for ease in grading or for hybrid or online classes, this exercise can be embedded within the LMS (e.g., Moodle) to include boxes for short-answer responses, multiple-choice questions, etc.

Supplementary Material

The interactive CDF elements shown in the figures can be downloaded here: Figure 1A, Figure 1B, Figure 2, Figure 3A, Figure 3B, Figure 4, and Figure 5. You will need to download the free stand-alone Mathematica CDF player (https://www.wolfram.com/cdf-player/) to open the files.

○ Acknowledgments

References

AAAS. (2011). *Vision and Change in Undergraduate Biology Education: A Call to Action*. Washington, DC: AAAS.

Bandeira, B., Jamet, J-L, Jamet, D., & Ginoux, J-M. (2013). Mathematical convergences of biodiversity indices. *Ecological Indicators*, *29*, 522–528. doi: https://doi.org/10.1016/j.ecolind.2013.01.028

Boyce, R. L. (2005). Life under your feet: Measuring soil invertebrate diversity. *Teaching Issues and Experiments in Ecology*, *3*(1). Retrieved from http://tiee.esa.org/vol/v3/experiments/soil/abstract.html

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, *1*(5), 1–5. doi:http://dx.doi.org/10.1126/sciadv.1400253

Drew, J., Sardelis, S., & Davis-Berg, E. C. (2017). Using a game-to-class pipeline to teach ecology. *Frontiers in Ecology and the Environment*, *15*(7), 357–358. doi: http://dx.doi.org/10.1002/fee.1520

Hull, P. M., Darroch, S. A. F., & Erwin, D. H. (2015). Rarity in mass extinctions and the future of ecosystems. *Nature*, *528*(7582), 345–351. doi:http://dx.doi.org/10.1038/nature16160

Mulhern, G., & Wylie, J. (2004). Changing levels of numeracy and other core mathematical skills among psychology undergraduates between 1992 and 2002. *British Journal of Psychology*, *95*(3), 355–370. doi: http://dx.doi.org/10.1348/0007126041528176

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*(3), 223–231. doi:http://dx.doi.org/10.1002/j.2168-9830.2004.tb00809.x

Roach, W. J., Bricker, M., Maruca, S., & Meir, E. (2010). Biogeography SimUText Ecology: http://www.Simbio.com.

Simpson, E. H. (1949). Measurement of diversity. *Nature*, *163*, 688. doi: http://dx.doi.org/10.1038/163688a0

Smith, T. M., & Smith, R. L. (2012). *Elements of Ecology* (8th ed.). San Francisco: Benjamin Cummings.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. doi:http://dx.doi.org/10.1016/0959-4752(94)90003-5

Waters, C. N., Zalasiewicz, J., Summerhayes, C., Barnosky, A. D., Poirier, C., Gałuszka, A., . . . Wolfe, A. P. (2016). The Anthropocene is functionally and stratigraphically distinct from the Holocene. *Science*, *351*(6269). doi:http://dx.doi.org/10.1126/science.aad2622

ELIZABETH C. DAVIS-BERG is an Associate Professor of Biology in the Science and Mathematics Department at Columbia College Chicago, 600 S. Michigan Ave, Chicago, IL 60605; email edavisberg@colum.edu. DANIEL R. JORDAN is an Associate Professor of Mathematics in the Science and Mathematics Department at Columbia College Chicago, 600 S. Michigan Ave, Chicago, IL 60605; email djordan@colum.edu.