CALCULATING BIODIVERSITY IN THE REAL WORLD

Students go outdoors to compare biodiversity in two wooded areas using Simpson's index

Melissa Schen and Leslie Berger

One of the standards for life science addressed in the *Next Generation Science Standards* (NGSS Lead States 2013) is "Ecosystems: Interactions, Energy, and Dynamics" (HS-LS2). A critical concept included in this core idea is biodiversity. To show competency, students are expected to design investigations, collect data, and perform mathematical analyses to explain how biodiversity changes or remains the same in ecosystems with different factors (Figure 1, p. 26).

To this end, we designed an ecosystem biodiversity investigation, posing the research question: How do two wooded areas of different ages compare in species biodiversity? Students collected plants in two areas of a local wooded ecosystem and calculated Simpson's biodiversity index (Figure 2, p. 26) to compare the relative abundance of organisms in each area. Simpson's biodiversity index gauges the number of species present and the relative abundance of each species. We chose the Simpson index, named for Edward H. Simpson, a British statistician, because it can compare data sets of unequal size and is based in probability, and research has demonstrated that students using it better comprehend relative abundance of species, or *species evenness* (Dor-Haim, Amir, and Dodick 2011).

FIGURE 1

Connections to the Standards

Dimension	Code	Description
Performance expectation	HS-LS2-2	Use mathematical representations to support and revise explanations based on evidence about factors affecting biodiversity and populations in ecosystems of different scales.
Disciplinary core ideas	LS2.A	Interdependent Relationships in Ecosystems
	LS2.C	Ecosystem Dynamics, Functioning, and Resilience
Crosscutting concept		Scale, Proportion, and Quantity
Science and engineering practice		Using Mathematics and Computational Thinking
Common Core State Standards, Mathematics	MP.2	Reason abstractly and quantitatively
	MP.4	Model with mathematics
	HSN.Q.A.2	Define appropriate quantities for the purpose of descriptive modeling

FIGURE 2

Sample calculation of Simpson's index of biodiversity.

Simpson's index of biodiversity (D) can be used to give a quantitative value of evenness in a sample. Values range from 0 → 1, with the greater the value, the greater the equality of distribution of species, and the greater the biodiversity.

$$D = 1 - \frac{\Sigma n(n-1)}{N(N-1)}$$

n = the total number of organisms of a particular species
N = the total number of organisms of all species

Example:

Species	Number of individuals (n)	n – 1	n * (n – 1)	Total
Species A	20	19	380	
Species B	30	29	870	1340
Species C	10	9	90	
Total (N)	60	59	3540	3540

Simpson's D = 1 – [1340 / 3540]
Simpson's D = 0.62

Collecting biodiversity data

Site choice

The first step in the investigation is to choose a suitable area for sampling—ideally a short walking distance from school with a variety of readily identifiable species. The sampling area could be a land lab, public park, unmaintained wild-growth area, playing fields, or grassy area on campus. Even lawn areas may have sufficient biodiversity with several types of grasses and wildflowers (e.g., clover, ground ivy). For our investigation, we used a wooded area near campus.

Preparation

After selecting your site, choose a sampling method. A common one is *quadrat sampling*. You use a quadrat—a small square plot—to sample for organisms by counting individuals or calculating percent of area covered. In our wooded site, we chose a 10 m × 10 m quadrat, which was large enough to include less densely spaced shrubs and trees. To lay out this plot, we used a tape measure and compass to make it square (walking 10 m in one cardinal direction, turning 90° to the next cardinal direction and walking another 10 m, and so on) as well as string for marking the boundaries. To save time, we marked two plots so each class section would have a fresh, nontrampled plot to explore and for the sake of comparison. The plots were in two wooded areas of known age: one that had never been cleared (climax

FIGURE 3

Assigned tasks for sampling wooded area.

Trees and bushes	Herbaceous plants
Tree circumference (4 students) Ash tree count Box elder tree count Dogwood tree count Sugar maple tree count	Ground ivy count Honeysuckle count (3 students) Raspberry count Virginia creeper count Garlic mustard count (2 students) Wild ginger count (3 students) Mayapple count False Soloman's seal count Wild leek count Waterleaf count Jewelweed count

Note: The above distribution is for a class of 24. Unless otherwise noted, each task was given to one student. To accommodate smaller classes, simple tasks can be doubled up if necessary. For larger classes, students can work in pairs on more complex tasks.

community) and one cleared about 60 years ago.

We focused on local plants, because plants compose a single trophic level, are easier to quantify, and can also readily demonstrate the diversity of an area. Before the investigation, teachers gathered samples of local flora to help prepare students. Provided with samples, students could learn to identify the plants while still in the classroom via direct instruction and field guides. This introduction also helped ease student apprehension about what they may encounter, especially those that had rarely explored the outdoors beyond their own backyards.

Collection

On data collection day, students randomly drew index cards identifying their collection task (Figure 3). Since our sampling site was large, groups of two to four students were responsible for counting a single, preassigned species. (Some students collected data on tree circumference for a later discussion on succession.) As the class headed to the site, students were permitted to trade task assignments as desired. The initial data collection was completed in one class period. A second period was needed to collect data in the second site used for comparison. However, teachers could have another class section collect data at the second site so field work could be completed in a single day. The data can then be compiled into one spreadsheet to share with all students.

Safety

Safety precautions depend on the sampling site. Grassy areas on campus may not require unusual safety precautions, but wooded or wild-growth areas require long pants, long-sleeve shirts, closed-toe shoes, and eye protection to reduce the chance of injury from sticks, branches, insect bites, and plants such as poison ivy.

Analyzing data

Mathematical analyses and indices provide several methods to descriptively and quantitatively compare biodiversity in multiple areas. Data analysis can focus on the number of species (richness) and the relative abundance (evenness) at the sampling site. To illustrate a sample's richness, one simply counts the number of different species present, with the higher value indicating higher biodiversity. In comparing multiple sites, richness alone doesn't provide an accurate analysis, as it doesn't describe how many individuals are in each species. For example, as noted in Figure 4 (p. 28), in two areas with equal richness (nine species), one area is dominated by a single species (wild ginger) with relatively fewer individuals in the additional species, while the other area has a more even distribution of individuals across the different species. The more even the distribution, the greater the biodiversity.

Evenness can be mathematically calculated several ways. First, students can calculate the percent abundance of each species and compare the distributions using a pie or bar chart. However, this gives only an approximate comparison and may be distorted when sample sizes differ, which is often the case. For a more direct comparison, ecologists use an index, such as Simpson's biodiversity index (Figure 2, p. 26). To do so, students use the data they collected on the counts for each species (n), as well as a sum of these values to get the total number of individuals across all the species (N). After multiplying each value by itself minus one, the sum of n*(n-1) will be strongly influenced by any species with very large numbers [e.g., $2 \times (2-1) = 2$ versus $20 \times (20-1) = 398$]. An area dominated by one species will have a proportionally larger value when this sum is divided by the total number of individuals (N) multiplied by N-1. To make more logical sense, subtracting this value from 1 will give a small number, indicating less biodiversity. A statistical way to interpret the value is that it gives the probability that two random individuals will be from different species—the higher the value, the greater the evenness. The key is that it takes into account all the relative proportions of individuals in each species. This allows students to readily compare the evenness of the sample areas.

Results of the student investigation (Figure 4) show that Area 2, which was cleared approximately 60 years ago, has a greater ecosystem biodiversity than Area 1. Even though it has fewer total individuals and the same number of species as Area 1, the relative abundance of those species

FIGURE 4

Ecosystem biodiversity calculation comparison using student data.

Species	Area 1	Area 2
Sugar maple	11	28
Ash		3
Box elder		2
Dogwood		2
Ground ivy	1	24
Honeysuckle	7	95
Garlic mustard	43	
Mayapple	3	
False Soloman's seal	9	
Wild leek	19	
Waterleaf	6	
Wild ginger	521	
Jewelweed		39
Raspberry		4
Virginia creeper		10
Total number of species (richness)	9	9
Total number of individuals	620	207
Simpson's biodiversity index	0.288	0.723

Note: Area 2, which was cleared approximately 60 years ago, has a greater ecosystem biodiversity. Even though it has fewer total individuals and the same number of species as Area 1, the relative abundance of those species is more evenly distributed as noted by the higher Simpson's biodiversity index. Area 1, an area never cleared, has the same richness and more individuals but is dominated by wild ginger.

is more evenly distributed as demonstrated by the higher Simpson's biodiversity index. Area 1, an area never cleared, has the same richness and more individuals but is dominated by wild ginger.

Assessment and expansion

Students' ability to quantify and compare biodiversity was assessed through a problem similar to the data analysis described here. Students were given a sample data set of two hypothetical areas' species composition that included fewer species than what they collected in this activity. The students were asked to "determine which area has the greater species biodiversity and support your answer." Ideally, students calculated the richness and evenness to compare the two areas represented in the given data set. The key was explaining how their mathematical evidence supported their conclusion, with references to richness and evenness of each area.

We found our classes did demonstrate significant gains in conceptual understanding (pretest average = 38%, posttest average = 60%, $p > 0.01$). The greatest gains were achieved by those students who could use their calculations as evidence to support their conclusion, the critical skill envisioned by the *NGSS* (in the scientific practices of Using Mathematics and Computational Thinking and Constructing Explanations and Designing Solutions).

Although we only investigated one research question comparing biodiversity, this methodology and Simpson's biodiversity index can be applied to other questions as well. Instead of comparing wooded sites of different ages, students could compare maintained versus wild-growth grassy areas, variation among local neighborhoods, the effect of seasons, or how invasive species affect biodiversity. For a field setup including smaller and denser plants such as grasses or wildflowers, the typical size for plots is 1 m^2. A hula-hoop or 1 m × 1 m square of connected PVC pipe lengths or metersticks can be used to delineate sample sites. A transect-line string along a field with knots at equal increments or a random throw of a tennis ball can be used to choose locations for the plots. For data collection, students in pairs can identify and count the species present in their square or hoop, completing multiple sites in one visit instead of one large site completed by the whole class. However, students need not be restricted to plants; they could identify and count birds at feeders or collect insects with nets, so long as all samples are from a single classification group for direct comparison.

Challenges

The logistics of taking students outdoors were the main challenge in collecting the biodiversity data. Preparation was key; preparing sites, precollecting organism samples, and dividing tasks allowed the data collection to go smoothly. Working with the plant samples in the classroom helped to ease anxieties, and preassignment of individual sampling tasks helped to keep students focused.

Weather can also be an issue, especially in winter. One possible solution is to collect data early in the school year, while the weather is still mild, during nature of science lessons. The data could be analyzed later during the ecology unit. Additionally, students could use simulation data when going outdoors is impractical.

Conclusion

Using Simpson's biodiversity index gives students experience in mathematical equations and probability in life science, an integration promoted in the *NGSS* and *Common Core State Standards, Mathematics* (2010) (Figure 1, p. 26). Using an equation such as the Simpson's biodiversity index was new to our students, requiring practice with multiple example data sets and more than one class period, but they responded well, similar to students in a previous study (Dor-Haim, Amir, and Dodick 2011). Once students became familiar with the variables, their ability to complete and interpret the calculations improved, as did their comprehension of the concept of evenness.

Paper activities and online tools can simulate a mathematical comparison of biodiversity, but collecting data outdoors helps students engage with their local ecosystems and provides a more authentic experience while answering research questions. Overall, students' engagement in collecting and analyzing meaningful data translated not only into significant learning gains, but some students also became very excited about and engaged in ecology. For example, the focus on collecting plants increased students' respect for plants' role in the ecosystem and as the foundation of the ecological pyramid.

As the focus in science education shifts to integrating authentic scientific practices with content learning, so must teaching. The biodiversity lesson described here can help achieve that goal. ■

Melissa Schen (missy.schen@gmail.com) *is the director of educational assessment at the College of Wooster in Wooster, Ohio, and Leslie Berger* (bergerl@limashawnee.com) *is a science teacher at Shawnee High School in Lima, Ohio.*

References

Dor-Haim, S., R. Amir, and J. Dodick. 2011. What do Israeli high school students understand about biodiversity? An evaluation of the high school biology programme, 'Nature in a World of Change.' *Journal of Biological Education* 45 (4): 198–207.

National Governors Association Center for Best Practices and Council of Chief State School Officers (NGAC and CCSSO). 2010. *Common core state standards.* Washington, DC: NGAC and CCSSO.

NGSS Lead States. 2013. *Next Generation Science Standards: For states, by states.* Washington, DC: National Academies Press.