Using data as a scaffold to build abstract thinking skills in undergraduate ecology students

As undergraduate science instructors, we all want our students to be sensitive, curious, humble observers of the natural world and to learn to ask insightful questions about how that world works. We also know that the right questions are necessary but not sufficient for scientific thinking. Students must also be able to translate their questions into testable hypotheses and design experiments to collect data that are relevant to those hypotheses. At this point, students must make the jump from concrete to abstract thinking as they realize that their question about how an organism responds to a change in temperature (for example) can be represented as a change in a dependent variable, such as size or growth rate. They also become aware that there are known statistical structures (such as a t-test) that can support this question, and start to see how their question is specific case of this general structure. They need to become accustomed to the idea that a column of numbers represents a certain type of information about their question and that rows represent individual observations in their experiment. In my experience, this transitional moment between a concrete question such as “Do plants grow faster at warmer temperatures?” and the translation of that question into more abstract numbers that represent plant growth is one of the most challenging pedagogical situations. The challenge is heightened when we start to ask questions about more complex systems with more interacting variables, longer time frames, and larger spatial perspectives, as is demanded at the frontiers of today’s ecological science. Yet, students learn science best by doing it, so the only option to help students see the connections between the concrete question and the abstract numbers is to get them working with all kinds of different data in all kinds of different contexts.

I use data in many situations in my classroom. I often have students design small-scale greenhouse experiments from start to finish over the course of a semester, and have them collect and analyze their own data. These experiments are student-conceived, open-ended and are usually in the form of a two-way analysis of variance examining the effects of two independent variables on some aspect of plant performance. I also have students collect data on soil pH, soil moisture, tree diameter, or plant density in contrasting field sites during one lab and then analyze those data themselves in a future lab period. In these situations, my goal is to get students practicing the steps of the scientific process and seeing data, and the abstract concepts it represents, in the context of their personal project so that they can understand more deeply what these abstractions represent.

These small scale experiments are excellent learning tools, but I am also interested in exploring ways to move students from thinking in the context of a single site or pot experiment to comparing data across sites, to improving their understanding of the workings of climate models, and help them recognize connections between remotely sensed patterns in the environment and ecological processes. These exercises require additional layers of abstract thinking and an understanding that the impacts of different ecological processes vary across spatial scales. Recently, I have tried three teaching activities related to these larger scales. First, I am the President of the Board of Directors for the Ecological Research as Education Network (EREN), a group that received NSF support from 2010 to 2016. EREN facilitates development of collaborative ecological research projects that engage undergraduate students and faculty at primarily undergraduate institutions in coordinated data collection across multiple sites. We have just started to develop formal teaching modules from EREN data sets that encourage students to compare data from multiple sites. One of these uses tree diameter data from a series of about 60 20 x 20 m permanent plots distributed in forests across the United States. My colleagues and I selected a subset of these data for students to explore using R order to learn more about environmental (and phylogenetic) variation that may affect carbon accumulation in forests at regional scales. In this exercise, students are introduced to multiple regression, which allows them to identify which of several possible independent predictor variables (such as temperature or precipitation) is having the strongest effect on a dependent variable (in this case, carbon accumulation by trees). This exercise was one of three presented in a workshop focused on the EREN Permanent Forest Plot Project at the American Biological Laboratory Educators (ABLE) Meeting in Columbus, OH in 2018 and will be published in the proceedings of this meeting later this year. I hope to get more ideas for developing quantitative modules based on EREN during this workshop.

To help students explore climate models, I like the following data exercises provided by Teaching Issues and Experiments in Ecology (TIEE) from the Ecological Society of America:

- **Investigating the footprint of climate change on phenology and ecological interactions in north-central North America** by Kellen M. Calinger
- **Global Temperature Change in the 21st Century** by Daniel R. Taub and Gillian S. Graham

Both exercises have students making graphs of temperature change over time in Excel and fitting regression lines to the graphs. The students learn that the slope represents a way to measure the extent of change over time, something that even advanced students seem to need a reminder about.

Finally, I recently joined forces with a colleague in remote sensing as part of a team teaching endeavor. Our goal to help my ecology students understand how remote sensing to could be applied as a tool to investigate ecological patterns and to help the remote sensing students have a stronger context for their technical skills. We trained both sets of students on the fundamentals of using reflected energy as a tool to analyze land cover and then had the remote sensing students analyze two images to quantify the land area converted from tropical forest to oil palm plantations in Costa Rica and Brazil. The ecology students, using published literature values, then calculated the carbon losses that occurred as a result of this conversion, taking into account carbon storage as standing biomass, carbon uptake rates as estimated by NPP, and carbon losses as via soil respiration. While this did not involve working with a large data set directly, students had to grasp the process of converting land cover into reflectance information in pixels, and then go through a series of calculation steps to convert this pixel information into carbon values. The students were very challenged, but seemed to gain a lot from the exercise.

In summary, I have tried a lot of teaching techniques that involve data. Some have worked well, but I believe there are still many challenges to using data in teaching. These include the fact that I often under-estimate the time it takes to teach quantitative skills, leaving me wondering if I am compromising biology content. Students vary widely in their quantitative skills and some can become so frustrated during an exercise that the learning value is lost. Finally, the intimidation factor for students and instructors is significant, such that some students are dismayed before they even begin and instructors will avoid such content for that reason. I think the development of well-structured modules is a good approach to supporting both students and instructors in integrating more quantitative content into teaching. If we are to teach abstract thinking as applied to the ecological world, working directly with data is the way forward.