JOURNAL OF GEOSCIENCE EDUCATION **62**, 578–586 (2014)

# Using Rich Context and Data Exploration to Improve Engagement with Climate Data and Data Literacy: Bringing a Field Station into the College Classroom

Amy L. Ellwein,[1,a] Laurel M. Hartley,[2] Sam Donovan,[3] and Ian Billick[1]


## ABSTRACT

Authentic scientific data, when richly contextualized, can provide the basis for compelling learning experiences. Many undergraduate students either do not have access to primary data, or if they do, the data are so abstract that student engagement is limited. Here, we describe contextual information and data-rich, student-centered activities we developed to give life to data sets from an intensely studied place, the Rocky Mountain Biological Laboratory (RMBL). The project Web site, Digital RMBL, highlights charismatic organisms, scientists, and long-term data sets as a tool for engaging students who are unable to physically visit a field station. The *Biology of Climate Change* module, the focus of this paper, has been tested in college-level classrooms by 10 collaborating faculty and 243 undergraduate students at a variety of colleges and universities across the nation. Authentic long-term data sets, primary literature, data visualizations, and a flexible format suitable for laboratory sections have led to very high usability ratings by collaborating faculty. In student surveys, a surprising number of undergraduate science majors (20%–30%) report that they have never worked with authentic scientific data—even at major research universities. The percentage of students who have not worked with data is much higher at collaborating 2 y institutions (60%–80%). The majority of students report that they appreciate the opportunity to explore long-term climate science data sets despite the frustrations they experience with the "messiness" of authentic scientific data. The impact of this climate change activity is achieved through *engagement* with people, place, and research subjects, followed by student-centered data *exploration* that builds personal interest and scientific discovery skills. This paper outlines one model in which scientists can meet funding agency requirements to share data publicly while providing excellent opportunities for improving climate and data literacy at the college level. © *2014 National Association of Geoscience Teachers*. [DOI: 10.5408/13-034]


*Key words*: climate literacy, data literacy, place-based learning, 5E instructional model, phenology, data visualization, engagement

## INTRODUCTION

Introducing students to inquiry-based science within an active and collaborative environment is critical to training the next generation of scientists and citizens (Manduca and Mogk, 2003). Long-term data sets collected at specific sites provide valuable records of change over time that are instrumental to understanding ecological and Earth system phenomena, and those data increasingly play an important role in undergraduate science education (e.g., Trautmann and McLinn, 2012). Access to authentic data is improving exponentially (e.g., DataONE, 2014; Dryad Digital Repository, 2014; etc.), but without rich context and compelling narrative, these data streams are inaccessible to those without advanced training. Context for data, especially interesting information about people and place, has been shown to improve student willingness to engage with science, especially for nonscientists (Feinstein et al., 2013) and populations traditionally underrepresented in science, technology, engineering, and mathematics (STEM) careers (e.g., Wilson et al., 2010; Brownell et al., 2012; NRC, 2012).

The field station experience is often transformative for undergraduates (Jack, 1939; Brussard, 1982; Eisner, 1982; Wilson, 1982; Smith, 1993), in part because immersion in science and a community of scientists is particularly *engaging*. Field stations have long been important centers for biological training. Since 1991, the Rocky Mountain Biological Laboratory (RMBL) has been providing opportunities for undergraduate research through the National Science Foundation (NSF) Research Experiences for Undergraduates (REU) program and related coursework and research opportunities. Approximately 30% of current RMBL scientists spent time at RMBL as undergraduates or shortly after graduating. Almost 50% (90 of 187) of the RMBL REU students since 1991 have gone on to graduate school, 20 have received advanced degrees, and seven have tenure-track positions. Digital RMBL (www.digitalrmbl.org) is an attempt to broaden aspects of the impact of the field stations to reach students who are unable to participate directly in the field station experience. The overarching goal of the Digital RMBL project was to bring long-term, place-based studies of natural systems into college-level classrooms and to assess the extent to which Web-based resources can provide some of the educational benefits (e.g., engagement with science and science practices) of field stations.

Through the Digital RMBL project, we have leveraged the scientific community and their long-term data sets collected at the RMBL in Gothic, Colorado, to produce a

*Received 13 May 2014; revised 28 May 2014 and 9 September 2014; accepted 10 September 2014; published online 19 November 2014.*

[1]The Rocky Mountain Biological Laboratory, PO Box 519, Crested Butte, Colorado 81224, USA
[2]Department of Integrative Biology, University of Colorado, Denver, Colorado 80217, USA
[3]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA
[a]Author to whom correspondence should be addressed. Electronic mail: amy@rmbl.org.

 

Web-accessible collection of data, contextual narrative, and research resources for college-level biology instruction. Our objectives with the Digital RMBL project were to (1) improve the scientific discovery skills of students, including the ability to frame scientific questions, work with data, and access primary literature, (2) increase the willingness and ability of students to engage with science, and (3) train students, faculty collaborators, and public audiences to use RMBL's online digital repositories. We developed four data-rich modules that incorporate richly contextualized data and opportunities to mine those data in pursuit of students' own research questions. Here, we focus on the *Biology of Climate Change* module, the most widely tested of the four curricular modules in Digital RMBL. This module challenges learners to explore the effects of changing climatic conditions on high-altitude ecologic systems in the Rocky Mountains.

In this study, we primarily test the ability of the *Biology of Climate Change* module to impact students' willingness to engage with science. Additionally, we were interested in learning how many undergraduate students have had opportunities to engage with authentic data through their own questions and analysis. Because the module was tested in a wide range of institutions, we were also in a position to analyze key factors that lead to faculty adoption of data-rich activities in their undergraduate courses. Open-ended, primarily attitudinal survey responses from this study support an extensive research base showing that data-rich, student-centered activities build interest and engagement with science and scientific research (e.g., Bransford et al., 2000; Manduca and Mogk, 2003; Brownell et al., 2012; Feinstein et al., 2013), a critical step in building science literacy as well as scientific skills and habits of mind (e.g., Rutherford, 1990; Ebert-May and Hodder, 2008; NRC, 2012).

## TARGET POPULATIONS AND SETTINGS

Eight RMBL-affiliated scientists were recruited as a steering committee at the conception of the project. After commencement of the project, four steering committee members were no longer able to participate, and 10 new collaborators were added. Six of the new collaborators used the *Biology of Climate Change* module with their students, for a total of 10 faculty collaborators testing this particular module. Of those 10 collaborators, six had previous personal experience with the RMBL, and four had not. Our testing collaborator mix allowed us to achieve diversity in both institutional nature and regional setting: five collaborators taught at research universities, three taught at primarily undergraduate 4 y universities, and two taught at 2 y institutions. Two collaborators taught at Hispanic-serving institutions (one research institution and one 2 y college). Three faculty that tested the modules were located in the southern U.S., four were on the west coast, two were in the Rocky Mountains, and one was in the upper Midwest.

The institutional diversity of faculty ensured diversity in the targeted student population. Student engagement was our primary goal, but we did not collect demographic data at the individual student level. Digital RMBL modules were tested in introductory biology and environmental science courses, biology and environmental science courses for nonmajors, and junior- to senior-level majors-only biology courses. Class sizes ranged from large introductory courses of 200+ students to laboratory sections or small courses with 10 or fewer students. The intended audience for all Digital RMBL modules was an introductory college-level audience, and advanced skills or detailed content knowledge were not expected.

## METHODS

### Understanding Faculty Needs

Here, we describe the methods used for assessing faculty curricular needs and translating those needs into curriculum. The module development process was preceded by a qualitative faculty needs assessment of the initial steering committee. The needs assessment was conducted through telephone interviews; questions were delivered to faculty a few days before each interview. The steering committee members were asked to describe (1) their interest in the project, (2) how and why they currently use data in their courses, (3) barriers to using data with their students, (4) their opinion on the relative importance of using networked computing and communications tools to support teaching and learning (i.e., cyberlearning), and (5) the most useful formats of supporting materials for their courses (2 h laboratories, homework assignments, demonstrations, capstone projects, etc.). Faculty were also asked to provide feedback on proposed module content, module design, and teaching strategies before materials were developed to maximize the potential for widespread faculty adoption.

All eight members of the steering committee had personal connections to RMBL (e.g., had conducted research at RMBL and/or taught summer courses at the field station) and were currently teaching in research universities (3), small public liberal arts colleges (4), and community colleges (1). All expressed interest in using authentic long-term data sets in their courses, and all were interested in using high-quality instructional materials, especially data-centered and inquiry-based stand-alone modules. One collaborator taught an upper-level environmental systems modeling course and was very interested in his students using authentic data sets to construct dynamic simulation models. Another stressed the importance of student interaction with data as a necessary component of learning scientific concepts and skills.

All steering committee members reported that they provided opportunities for their students to interact with data in their courses. Five of eight often had students collect and analyze their own data in order to foster personal connection to the data and process of science, including communication of scientific results. One collaborator remarked that using authentic data sets in instruction (1) increases the sense of relevance, generating student interest, and (2) can illustrate how uncertainty and "noise" in authentic data sets contrast with idealized or "sanitized" data sets used in textbook examples. When asked which module formats were most appealing, by far the most requested format for data-rich activities (100%) was the 2 h laboratory assignment, followed by or supplemented with homework assignments, and lecture-ready demonstrations, especially data visualizations.

Steering committee members were asked to describe limitations or barriers they face in using data with their students. Of the faculty that had students collect and analyze their own data, limitations included the small size of student-collected data sets and the short duration of data

**TABLE I: Components of the *Biology of Climate Change* module.**

| Major Components | Specific Components |
|---|---|
| Data set(s) | Daily and summary weather observations (billy barr) |
| | Yearly phenology records (first sightings; billy barr) |
| Primary scientific literature | Climate change is affecting altitudinal migrants and hibernating species Inouye et al. (2000) |
| Contextual information | Colorado climate summary |
| | Summary of major drivers of ecological change in high-altitude Rocky Mountain systems |
| | Audio interview with David Inouye and billy barr (Colorado Public Radio) |
| | Snowpack video from the Snodgrass RMBL Meteorological Station |
| | Natural history pages for several organisms on the first sightings list |
| Assignment | Use the Data Visualizer to "explore" the data |
| | Reproduce plots from Inouye et al. (2000) |
| | Make new plots with additional data from billy barr |
| | Analyze new data |
| | Evaluate predictions in Inouye et al. (2000) |
| | Class discussion of findings |
| For instructors | Summary of data and publications |
| | Learning goals |
| | Teaching strategies |
| | Additional tools (readings, videos, etc.) |
| | Examples of how the module was used by others |

collection. Other barriers included the time-consuming nature of data manipulation, especially if students needed to learn new software, and the difficulty of concurrently assisting large numbers of students. Two of the faculty remarked on the difficulty of securing computer classrooms for data manipulation exercises. Most either provided data for their students to analyze and/or had students collect their own data to analyze. Only two of the initial collaborators asked students to find data online; both of them required that the data originate from the primary literature.

## Developing Contextualized Case Studies from Primary Literature and Authentic Scientific Data

Digital RMBL modules combine data sets, published research articles, and contextual background in student-centered activities designed to promote student engagement while building important content knowledge and science practices. Module development began with authentic long-term data sets contributed by RMBL-affiliated scientists. In collaboration with scientists, we identified publications from those data sets deemed "accessible" to nonexpert readers. Working from published peer-reviewed journal articles ensured that scientists retained control of their data until publication. Contributing scientists were consulted for suggestions on using their data and publications in teaching the relevant scientific concepts or practices identified by the steering committee.

To encourage student engagement with data and science practices, we developed rich context around the data (Table I). Popular science articles (including audio interviews, video, etc.) provided relevance and a more easily understood framework for the data, as well as highlighting the personalities of the RMBL scientists. Other contextual information included background regarding RMBL researchers and their projects, natural history information about local organisms (including charismatic species, such as hummingbirds and marmots), students' descriptions of field research, and descriptions of major RMBL research strengths. Working from published research was important for context as it simplified the process of communicating a research "story." An added benefit is that students had the opportunity to learn to critically read primary scientific literature.

Modules were developed using the constructivist 5E instructional model (engage, explore, explain, elaborate, and evaluate) pioneered by the Biological Sciences Curriculum Study (Wilson et al., 2010). The 5E approach is widely used in developing K–16 inquiry-based science curricula (e.g., Bybee et al., 2006; Ebert-May and Hodder, 2008; Wilson et al., 2010). All materials were reviewed by appropriate RMBL-affiliated researchers before being uploaded to the Digital RMBL Web site for testing in college classrooms.

Each of the four modules is accompanied by information specifically for instructors, in which we outline the data set(s), peer-reviewed primary literature, and learning goals for the student investigations. Additionally, we suggest teaching strategies, lecture topics, and other useful tools. Most modules also contain a page or two outlining how collaborating faculty have used the module in their specific course. During the testing phase, each instructor section also included links to the online instructor and student surveys.

### The *Biology of Climate Change* Case Study

The *Biology of Climate Change* module was based on daily weather observations (temperature, precipitation, snowpack thickness, snowmelt date, etc.) and seasonal "first sightings" of hibernating mammals, migrating birds, and early spring flowers recorded by billy barr, RMBL's business manager and accountant (billy does not capitalize his name). billy's data were combined with other data sets to show the effects of climate change on migrating and hibernating species at this high-elevation field station and to outline predictions for future changes (Inouye et al., 2000). One major finding was that the start of the growing season in Gothic (snowmelt date) had not changed significantly since the early 1970s. In contrast, many scientists working at lower elevations had substantial evidence for lengthening growing seasons, earlier spring migrations, and earlier reproduction in a variety of different organisms. Inouye concluded that this disconnect was and will continue to be a problem for both (1) migrating species that visit high-elevation Rocky Mountain ecosystems during the summer and (2) hibernating resident species, such as the yellow-bellied marmot.

Since the 2000 publication, billy barr has continued to collect data. Students are asked to compare the results and predictions from the original publication (Inouye et al., 2000) with an additional decade of data to evaluate the conclusions of the authors. They are also directed to use the data and their newfound understanding of ecosystems around RMBL to ask and answer their own questions. For example, has the growing season at RMBL changed over the past decade? Has the timing of first sightings of migrating birds or hibernating mammals changed since Inouye's research was published?

The competency goals for the activity are to build student skills in (1) interpreting graphical and statistical data, (2) understanding statistical significance and regression, (3) using software (e.g., Excel) to plot and describe data using statistics, (4) giving effective oral group presentations, and (5) constructing an argument from data. The content goals for the activity are to help students understand (1) the biological significance of climate variables on a single species and (2) how climate change affects interactions among species in complex ways.

Contextual material is presented to *engage* learners with the setting, the people, the organisms, and the data. The self-contained module is presented through the Digital RMBL Web site, starting with background information on the climate of central Colorado, climatic drivers for ecosystems around RMBL, and an audio interview with David Inouye, the lead author, and billy barr; billy is a fascinating character, in part because his life is so different from the average person. He has lived in Gothic, a mining town turned ghost town turned field station, since the early 1970s. Along the backbone of the Rocky Mountains at almost 2900 m elevation (9500 ft), this location receives an average of 10.9 m of snowfall and is accessible only by skis or snowshoes for 6 mo or longer each year. The population in Gothic oscillates between 160+ residents (primarily scientists) during the 3 mo summer field season, to only a handful of winter caretakers for the rest of the year. billy isn't a scientist by training, but his data are exceeding useful to many RMBL scientists. His story is a particularly interesting "hook" for many students.

The first step towards using and understanding the data is to *explore* the data using the interactive Google motion chart (Fig. 1). The motion chart tool, a simplified version of the data visualization developed for Hans Rosling's Gapminder Web site (http://www.gapminder.org/), allows students to manipulate almost any component of the chart with minimal training. For example, variables can be changed instantly using dropdown menus, the color ramp used for the data points can represent any of the variables within the data set, and the graph is constructed data point by data point when students press the "play" button and are then walked through the time series. Intuitive operation allows students to focus, at least initially, on *exploring* the data and the patterns within the data without concurrently learning to manipulate specific software. The interactive data visualization can also be used as a demonstration or discussion tool in lecture settings with a projector.

After *exploring* the data using the interactive Google motion chart, students are asked to reproduce the plots from the paper (Inouye et al., 2000), using the updated data set with 10 y of additional observations. This step is intended to get students familiar with graphing software (Excel), seeing patterns in data, and using metadata. Instructors are asked to consider placing students in small groups (two to three students) for this step because many introductory students are not proficient with Excel and may benefit from working with a partner. Also, working in small groups can help boost confidence in analyzing data and presenting their findings to the class.

After plots have been constructed, students are asked to *explain*: "In your opinion, do the original plots support the original interpretations? Describe your reasoning. Do the updated plots suggest new interpretations? How have abiotic conditions changed over the extended period of record? How do first sightings change with the extended record? What are your predictions for future change?" This step forces students to analyze data and link data to the natural world.

At this point, students are familiarized with the data and are asked to generate their own testable questions, or *elaborate*. Before students get started on more open-ended inquiry, we recommend that the instructor lead the students in thinking through some answerable questions. For example, do all migrating birds arrive at the RMBL at the same time? Will all hibernating mammals show the same relationship with snowmelt date? The interactive data visualization can be used to assist students as they generate their own questions.

## Tools for Testing Module Effectiveness

From January 2011 through May 2013, the *Biology of Climate Change* module was tested in college classrooms in a wide variety of settings (as outlined above). Faculty collaborators were asked to complete an instructor survey after each test and have their students complete a module-specific student survey after each test. Tests focused on usability and engagement, not student learning gains. Both instructors and students completed paper or online surveys developed for this project to provide specific feedback on project deliverables and measure progress towards project objectives.

We developed one instructor survey for this project, which was used to evaluate any of the four Digital RMBL modules (see supplementary material 1). Instructors were asked several questions regarding (1) their observations of the relative engagement of their students, (2) alignment of the activity with their teaching goals, (3) primary cognitive skills used by their students (e.g., recognize, predict, explain, analyze, interpret), (4) the overall quality of the contextual information provided within the module (e.g., background

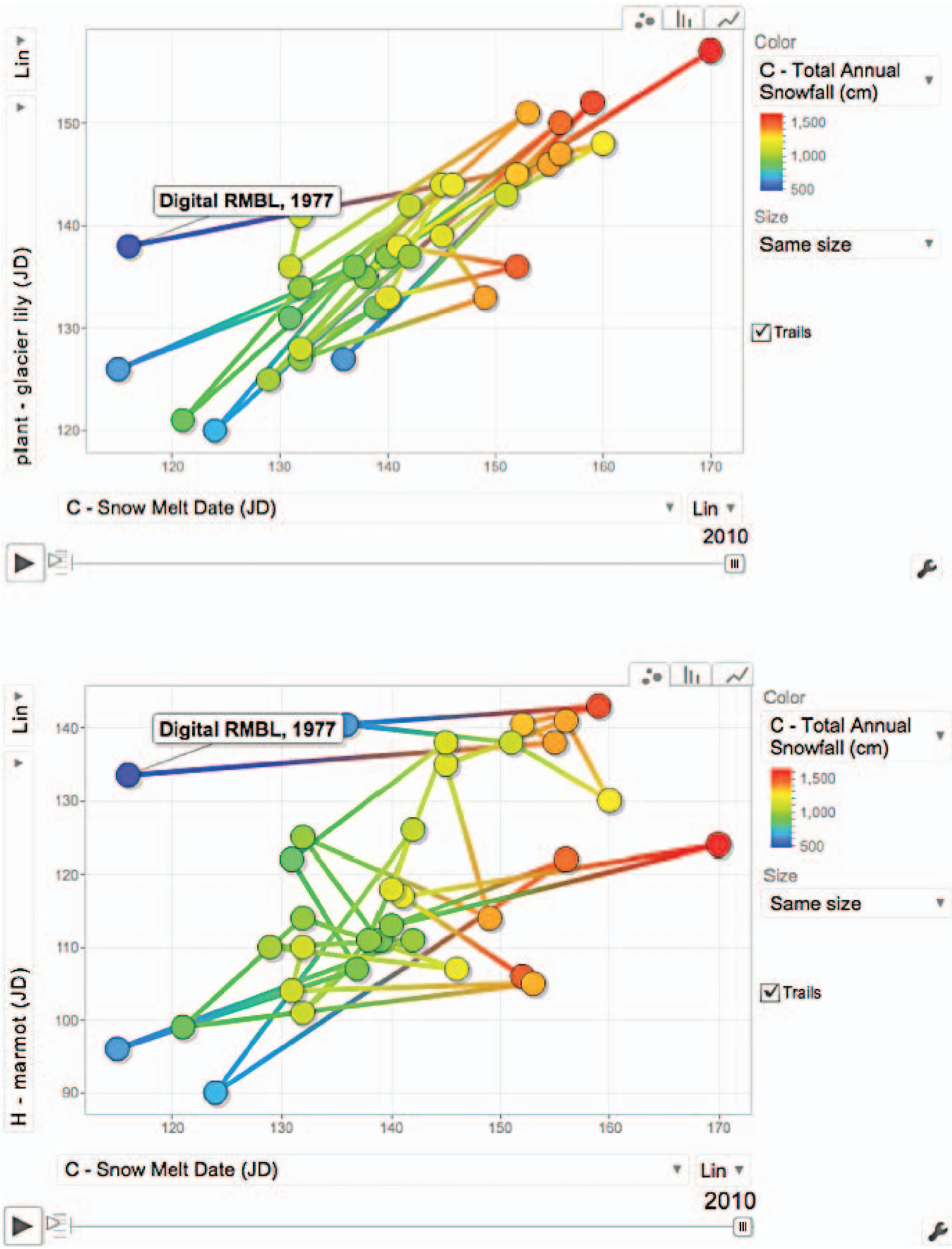


**FIGURE 1: The data set used in the *Biology of Climate Change* module contains climatic data (e.g., snowmelt date, total and average snowpack) and dates of "first sightings" for early spring flowers (e.g., glacier lilies; upper chart), as well as migrating or hibernating animals (e.g., marmots; lower chart) collected by billy barr in Gothic, Colorado, since 1973. Students can easily manipulate variables plotted, and the chart adds data points when students press the "play" button. Unfortunately, this simple (and free) Google tool generates lines between successive data points and cannot be used to produce regression lines or perform statistical analyses. (http://www.digitalrmbl.org/p85.php). (Color for this figure can be found in the online version of this paper.)**

information, interviews, primary literature articles, popular science articles, videos, etc.), (5) the straightforwardness or ease-of-use of each module, (6) the cost effectiveness (in terms of student and faculty time and effort) and/or novelty of the module (i.e., filling a need not filled by other known curricula), and (7) whether they would use the module again in the future.

Each module in Digital RMBL has a related module-specific student survey; the *Biology of Climate Change* student survey tool is included as see supplementary material 2. For each test of the module, collaborating instructors asked their students to provide information about their major, number of college years completed, and whether they had used a Web-based activity in a college course or whether they had

**TABLE II: Instructor perception of cognitive abilities used by students in the *Biology of Climate Change* module.**

| When using The Biology of Climate Change module, did your students ... | Affirmative Response | |
|---|---|---|
| RECOGNIZE the truth of a scientific principle or fact whether stated in its general form or implied based on a specific instance of a generalization | 6 | 60% |
| Use a scientific principle to PREDICT a phenomenon (e.g., what will happen to this object when this condition is changed) | 8 | 80% |
| EXPLAIN a phenomenon in terms of a scientific principle (e.g., why does this event occur given these conditions?) | 8 | 80% |
| ANALYZE a situation in terms of its component parts (e.g., what must be true for this object to behave this way?) | 10 | 100% |
| INTERPRET a representation of a scientific principle or phenomenon | 8 | 80% |

ever worked with authentic scientific data sets. Students were asked to report their comfort level with various module-specific tasks (e.g., interpreting primary literature, making graphs, using Excel, developing hypotheses, etc.) before and after their use of the module, and the relative value (using a Likert scale) of the contextual components specific to the module. Several open-ended questions were included to gauge the most and least interesting module components, one measure of engagement, and major learning outcomes from the students' perspectives.

## RESULTS OF CLASSROOM TESTS OF THE *BIOLOGY OF CLIMATE CHANGE* MODULE

### Instructor Responses

All surveyed instructors reported that their students used key cognitive skills (Table II) and that student engagement with the module activities was high. Regardless of institution type, major, or years until completion, students could comprehend the context and key ideas, were interested and motivated, and could make connections to other aspects of the course (Table III). Open-ended responses reported a high degree of student engagement, as well as close alignment of the module with instructor goals. Examples of instructor responses to open-ended questions are summarized in see supplementary material 3. The average reported duration of in-class use was 2.7 h, ranging from 40 min to 4 h. Most instructors reported that they had modified the module to some extent in order to fit their particular course or learning objectives. All collaborators were either very likely (9) or somewhat likely (1) to use this module again in the future and were very likely (7) or somewhat likely (3) to develop derivative lessons from it.

**TABLE III: Instructor evaluation of student engagement and various module components.**

| Student Engagement | Agree | Disagree |
|---|---|---|
| Students were able to comprehend the context and key ideas of the lesson. | 10 | 0 |
| Students were interested/motivated by the context and key ideas of the lesson. | 10 | 0 |
| Students were able to make connections between the lesson and other parts of your course. | 10 | 0 |
| Quality of Contextual Information | Agree | Disagree |
| The background information provides helpful elaboration of the key ideas, the data sets, and the context in which data were gathered. | 10 | 0 |
| The background information provided is sufficient to carry out the activity. | 10 | 0 |
| The background information is scientifically accurate. | 10 | 0 |
| The background material is comprehensible. | 10 | 0 |
| The background information integrates the RMBL Web site in a useful manner. | 10 | 0 |
| The background information is organized in a logical and useable way. | 9 | 1 |
| Quality of Procedure/Activity | Agree | Disagree |
| The materials needed were readily apparent and accessible on the Web site. | 7 | 3 |
| The procedure is comprehensible. | 9 | 1 |
| The formats of the graphs and tables were easy for students to comprehend and use. | 7 | 3 |
| The procedure integrates the RMBL Web site in a useful manner. | 10 | 0 |
| Cost Effectiveness and Usability | Agree | Disagree |
| The time and effort required for students to complete this activity are appropriate given the number/difficulty of concepts covered. | 9 | 1 |
| The time and effort required for students to complete this activity are reasonable compared with other activities you know that cover similar material. | 10 | 0 |

**TABLE IV: Selected student responses from student surveys.**

| | Two-Year Institutions | | Four-Year Institutions | | Research Universities | |
|---|---|---|---|---|---|---|
| | Science Major | Nonmajor | Science Major | Nonmajor | Science Major | Nonmajor |
| What year are you in college? | | | | | | |
| Freshman | 13% (2) | 19% (8) | 14% (13) | 100% (3) | 1% (1) | |
| Sophomore | 69% (11) | 55% (23) | 18% (17) | | 14% (12) | 100% (1) |
| Junior | 19% (3) | 26% (11) | 33% (32) | | 51% (43) | |
| Senior | | | 33% (32) | | 32% (27) | |
| Post-BA or BS | | | 2% (2) | | 3% (2) | |
| **Total** | **16** | **42** | **96** | **3** | **85** | **1** |
| Have you ever used a Web-based activity in a class? (Yes) | 63% (10) | 62% (26) | 83% (80) | 100% (3) | 89% (76) | 100% (1) |
| Have you ever worked with authentic data sets before this activity? (Yes) | 19% (3) | 36% (15) | 78% (75) | 67% (2) | 68% (58) | 100% (1) |

## Student Responses

Student surveys were collected from 243 students from participating 2 y institutions (58 respondents), 4 y institutions (99 respondents), and research universities (86 respondents). Most respondents were juniors or seniors majoring in a science or science-related field. Of those respondents, 20%–30% reported never having used authentic data in a college-level course, compared with a majority (60%–80%) of student respondents at 2 y institutions. Student responses to selected questions related to experience with college-level work and experience analyzing authentic data sets are summarized in Table IV.

In general, students were either not very detailed in their answers to the open-ended questions posed in the student surveys or didn't answer them at all. Of the students that responded, most students remarked that while they enjoyed using authentic data to answer their own questions, they did not enjoy using Excel or other software. Other dominant themes emerged from questions regarding student engagement with the module, here represented by self-reports of what students liked and didn't like. Sample responses, primarily from science majors at 4 y institutions and universities, are provided for each theme in see supplementary material 4.

### *Question: What Did You Like Most About the RMBL Web Activity That You Completed?*

Theme #1: Students appreciated the self-directed aspect of the activity (e.g., "doing" science), as illustrated by this quote: "I most enjoyed... creating [my] own question, hypothesis, and graphs to try and [answer] the question." The module includes an open-ended exploration in which students can generate and test additional hypotheses. What they chose to explore was dictated only by what they found to be interesting.

Theme #2: Students were engaged by using authentic data and/or exploring a research case study; "I liked being able to analyze new data for an older research project because it allowed for me to understand the process of analyzing new data." The RMBL climate change and phenology data sets are more extensive than data sets students can generate on their own with respect to the number of variables and the number of data collection years. The variables included in the data set are easily understandable and can be used to look for a large number of potential correlative patterns ranging from simple two-variable correlations to multivariate associations.

Theme #3: Students enjoyed the interactive data visualization tool, and the term "play" was used in many student responses. This tool allowed students to easily visualize associations among variables. For example, "my favorite part from this activity was the automatic self-running visualizing data plotter... it helped me to picture and imagine" [the relationships]. The tool is an exceptionally useful exploration tool because it doesn't require knowledge of data analysis programs.

### *Question: What Did You Like Least About the RMBL Web Activity That You Completed?*

Theme #1: Students did not like "messiness" or problems with analyzing authentic data; however, faculty thought "messiness" was an important trait of the data sets. Scientists rarely deal with perfect data, and students need to learn how to deal with issues such as missing data points, outliers, insufficient numbers of collection points, and unwieldy spreadsheets with thousands of cells. Interestingly, many students wrestled with the question regarding how many years of data one needs in order to have confidence that the patterns in the data are meaningful; "the paper itself was rather interesting to read, but lacked data beyond 40 years." Long-term data sets are particularly important for climate change research, but they are difficult to collect.

Theme #2: Some students were reluctant to "dig deeper" or take control of their learning. The module includes some prescribed activities to help students become familiar with the data set and some accessory tutorials on using Excel. However, the module also allows ample room for students to encounter concepts and tasks that they must decipher without the aid of step-by-step directions. For example, one student remarked that "it was hard to understand some of the concepts by myself without actually having to look them up and read them a couple of times." We assume this particular student learned more as a result.

## DISCUSSION

The importance of engaging students with primary data should not be underestimated. Our surveys revealed that many students, even science majors, do not have opportunities to work directly with primary data. The likelihood that students have experience analyzing authentic scientific data clearly increases with increasing college experience based on survey data from this (admittedly limited) study. Not surprisingly, freshmen and sophomores in introductory science courses at participating 2 y institutions report having little experience with authentic data (60%–80% report that they haven't used authentic data). However, the number of junior- and senior-level science majors that report never having used authentic data (20%–30%) in their college courses is surprisingly high (Table IV).

Based upon interviews, instructors are very willing to incorporate Digital RMBL modules into their courses because they are designed to overcome the issues inherent in teaching with real data. The primary barrier for collaborating faculty to incorporating authentic data into science courses is time—time to find large data sets or have students build their own data sets, build meaningful activities around those data, teach students data analysis skills including practice with software, and assist students with open-ended data-rich projects, while still managing to teach the breadth of content expected in their courses. Digital RMBL was designed to address such problems for instructors. Instructors serve as the gatekeepers for instructional materials. If instructors are not happy with the material, they will not incorporate it into their curriculum. Collaborators rated the case studies highly; they particularly liked the use of authentic data and primary literature in scaffolded, yet open-ended student investigations. Additionally, all collaborating instructors plan to continue using the module in the future—an important measure of success.

The *Biology of Climate Change* module heightens student engagement with data and science practices. Student engagement is attributable in part to the highly contextualized learning experience. Student surveys show that students felt connected to the data, in part because they could relate personally to billy barr, as both an interesting "character" and a citizen scientist. Additionally, students were engaged because they could explore the data through interactive visualizations before trying to quantify and analyze relationships in Excel or other software. Simple interactive data visualization tools, combined with context about people and place, made the data more compelling, easy to understand, and fun to explore. This sense of "play" seemed to help students feel more comfortable and familiar with the data, improving their ability and willingness to ask scientific questions about the data. Strong themes in student responses—irrespective of major, college type, or whether their instructor was directly affiliated with the RMBL—were appreciation for the opportunity to (1) analyze authentic scientific data, (2) ask and answer their own questions, and (3) "play" with data in interactive data visualizations.

However, simple graphical visualization tools are not the same as more sophisticated statistical analyses of data. Frustrations with "messy data" or specific software programs were also a common theme in student responses. We believe that the initial period of exploration and connection with social context (e.g., Feinstein et al., 2013) led to increased determination that helped temper students' software frustrations.

While the purpose of this study was to understand student engagement, there was anecdotal evidence that engagement translated into learning gains. Instructors claimed large gains in their students' content knowledge and data literacy skills after completing the module. One collaborator remarked that the *Biology of Climate Change* module was "indeed very useful, [her students'] exams were strong in that topic." One of the mechanisms for these learning gains was likely the increased engagement and motivation on the part of students.

## VALIDITY AND IMPLICATIONS

To reiterate, our primary goal with the *Biology of Climate Change* module is to increase undergraduate students' willingness to engage with data and to ask and answer their own scientific questions. We believe the major strengths of this module are the richly contextualized data used to understand a politicized issue through an open-ended, student-led inquiry exercise. Because this module was developed as a 2–4 h laboratory exercise, but was delivered in a wide range of settings and courses, to students with a wide range of background experiences, we focused on faculty observations and student self-reporting of student engagement to measure increased willingness to engage with data and science practices. We did not attempt to objectively investigate student learning, although that is a future goal. We piloted survey instruments with three instructors (three different courses) in diverse institutional settings before we implemented wider data collection. Faculty gave feedback about the survey instruments, and we made minor modifications to decrease the length of the faculty survey and to provide more space for open-ended feedback. Based on faculty and student feedback, we added questions to the student survey. Feedback about the module was generally consistent across institutions and class types.

Beyond the lessons learned in terms of the pedagogical impact of working with primary data, and mechanisms for increasing student and instructor engagement, our work also provides interesting insights into climate change education. In comparison with other modules developed for Digital RMBL, the *Biology of Climate Change* was particularly popular with instructors. Climate change and the effects of climate change are active research areas for many college-level faculty, and undergraduate students are interested in the topic, making the module content both relevant and engaging (Table III). Furthermore, we believe that direct and positive student engagement with climate-related data and contextualized information promote a sense of personal experience with a complex and often politicized topic. Average climatic conditions and climate change are largely statistical phenomena and therefore not directly observable; a growing body of research shows that personal experience with climate and climate change affects learning about climate change, perception of risk, and resulting behaviors (e.g., Weber, 2006, 2010; Blennow et al., 2012). The development of personalized experience with data plays a strong role in climate and data literacy because "personal experience is far more likely to capture a person's attention, and its impact dominates the often far more reliable and diagnostic statistical information" (Weber, 2010).

The Digital RMBL project is one model of how organizations (e.g., field stations, marine laboratories, online data repositories, etc.) and individual scientists can share their data with the public through curricula designed to improve scientific and data literacy. An unexpected but very positive outcome from the Digital RMBL project has been the degree of excitement by and potential future involvement of RMBL scientists. As the project has progressed, increasing numbers of RMBL scientists have become involved as collaborators, reviewers, and contributors. For example, one faculty collaborator assigned a natural history assignment to his students that resulted in most of the Digital RMBL natural history pages; another faculty collaborator's students built a dynamic interactive model that may be incorporated into the site in the future. In spring 2013, five RMBL-affiliated scientists proposed developing new modules for Digital RMBL, featuring their data, as a component of the broader impacts and data management sections of their scientific research proposals to the NSF. The Digital RMBL project connects the public with scientific data through effective online science curriculum development, increasing the value of science to society, as well as the public understanding of that value (e.g., Billick et al., 2013).

### Acknowledgments

The authors wish to thank billy barr and David Inouye for contributing data; Dan Jones, Ceci Ervin, Leah Roberts, and NEXT Interactives for their work on the Digital RMBL Web site; the RMBL community; and of course, many thanks go to our numerous project collaborators and their students. We extend special thanks to two anonymous reviewers for their constructive comments on the manuscript. This project was funded by the National Science Foundation (DUE-091875); any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. For more information about the Digital RMBL project, visit www.digitalrmbl.org.

## REFERENCES

Billick, I., Babb, I., Kloeppel, B., Leong, J.C., Hodder, J., Sanders, J., and Swain, H. 2013. Field stations and marine laboratories of the future: A strategic vision. National Association of Marine Laboratories and Organization of Biological Field Stations. Available at http://www.obfs.org/fsml-future (accessed 1 May 2013).

Blennow, K., Persson, J., Tomé, M., and Hanewinkel, M. 2012. Climate change: Believing and seeing implies adapting. *PLoS ONE*, 7:e50182.

Bransford, J.D., Brown, A.L., and Cocking, R.R., eds. 2000. How people learn: Brain, mind, experience, and school, expanded edition. Washington, DC: National Academies Press. Available at http://www.nap.edu/catalog.php?record_id=9853#toc (accessed 1 May 2013).

Brownell, S.E., Kloser, M.J., Fukami, T., and Shavelson, R. 2012. Undergraduate biology lab courses: Comparing the impact of traditionally based "cookbook" and authentic research-based courses on student lab experiences. *Journal of College Science Teaching*, 41:36–45.

Brussard, P.F. 1982. The role of field stations in the preservation of biological diversity. *BioScience*, 32:327–330.

Bybee, R.W., Taylor, J.A., Gardner, A., Van Scotter, P., Powell, J.C., Westbrook, A., and Landes, N. 2006. The BSCS 5E instructional model: Origins and effectiveness. A report prepared for the Office of Science Education National Institutes of Health. Available at http://www.bscs.org/bscs-5e-instructional-model (accessed 1 May 2013).

Data Observation Network for Earth (DataONE). 2014. DataONE. Available at https://www.dataone.org/ (accessed 9 September 2014).

Dryad Digital Data Repository. 2014. Dryad Digital Data Repository. Available at http://datadryad.org/ (accessed 9 September 2014).

Ebert-May, D., and Hodder, J., eds. 2008. Pathways to scientific teaching. Sunderland, MA: Sinauer Associates, Inc., Publishers. p. 212.

Eisner, T. 1982. For love of nature: Exploration and discovery at biological field stations. *Bioscience*, 32:321–326.

Feinstein, N.W., Allen, S., and Jenkins, E. 2013. Outside the pipeline: Reimagining science education for nonscientists. *Science*, 340:314–317.

Inouye, D.W., Barr, B., Armitage, K.B., and Inouye, B.D. 2000. Climate change is affecting altitudinal migrants and hibernating species. *Proceedings of the National Academy of Sciences of the United States of America*, 97:1630–1633.

Jack, H.A. 1939. Close to nature: Biological field stations. *American Biology Teacher*, 6:141–145.

Manduca, C.A., and Mogk, D.W. 2003. Using data in undergraduate classrooms: Final report on an interdisciplinary workshop. Available at http://serc.carleton.edu/files/usingdata/UsingData.pdf (accessed 1 May 2013).

National Research Council (NRC). 2012. A framework for K–12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: The National Academies Press. Available at http://www.nap.edu/catalog.php?record_id=13165 (accessed 1 May 2013).

Rutherford, F.J., ed. 1990. Science for all Americans: Project 2061. American Association for the Advancement of Science. Oxford, UK: Oxford University Press. p. 272

Smith, R.J. 1993. Biological field stations: Opportunities for undergraduate and faculty research. *Council on Undergraduate Research Quarterly*, 12:4.

Trautmann, N.M., and McLinn, C.M. 2012. Using online data for student investigations in biology and ecology. *In* Olofsson, A., and Lindberg, J., eds., Informed design of educational technologies in higher education: Enhanced learning and teaching. Hershey, PA: IGI Global, p. 80–100.

Weber, E. 2006. Experience-based and description-based perceptions of long-term risk: Why global warming does not scare us (yet). *Climate Change*, 77:103–120.

Weber, E. 2010. What shapes perceptions of climate change? *Wiley Interdisciplinary Reviews: Climate Change*, 1:332–342.

Wilson, C.D., Taylor, J.A., Kowalski, S.M., and Carlson, J. 2010. The relative effects and equity of inquiry-based and commonplace science teaching on students' knowledge, reasoning, and argumentation. *Journal of Research in Science Teaching*, 47:276–301.

Wilson, E.O. 1982. The importance of biological field stations. *Bioscience*, 32:320.

## SUPPLEMENTARY MATERIALS

Supplementary Materials referred to herein are available online at http:dx.doi.org/10.5408/13-034s1.