EARTHTIME: Teaching Geochronology to High School Students in the USA

Britta Bookhagen, Noah McLean, Robert Buchwaldt, Matthew Rioux, Francis Dudás, and Samuel Bowring

1 Introduction

1.1 Why Teach Deep Time in School?

There is widespread recognition among scientists and education policymakers that student engagement in science must be improved. In order to maintain our technological standard, we need to ensure a scientifically literate society and continued contributions by competent scientists (American Geophysical Union 1994; National Science Foundation 1996; National Research Council 1997; National Science Board 2002, 2003). Geologic time ("Deep Time") is an important concept in geology, as already established by Hutton 1788 and Lyell 1830, giving a logical timescale to many Earth processes and events. Understanding the timing and rates of geologic processes is critical for understanding such diverse topics as the age and assembly of the Earth, plate tectonics, the timing and causes of mass extinctions, and the recurrence rates of volcanic activity and other natural hazards. Because geologic time and geochronology integrate elements of chemistry, physics, biology, and mathematics, teaching "Deep Time" provides an opportunity to introduce students to a key scientific concept that integrates knowledge from a diverse range of disciplines.

Research has demonstrated that several common preconceptions in science should be individually targeted. Project 2061, an initiative of the American Association for the Advancement of Science (AAAS, 2009), was founded in 1985 to advance literacy in science and has included more than a decade of research and development on preconceptions in science. Philips (1991) confirmed and summarized some earth science preconceptions that address time and time measurement

B. Bookhagen (✉) • N. McLean • R. Buchwaldt • M. Rioux • F. Dudás • S. Bowring
Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA
e-mail: britta.bookhagen@gmail.com

V.C.H. Tong (ed.), *Geoscience Research and Outreach: Schools and Public Engagement*, Innovations in Science Education and Technology 21, DOI 10.1007/978-94-007-6943-4_11,

(e.g., mountains are created rapidly, all radioactivity is man-made, dinosaurs and cavemen lived at the same time). Research has also demonstrated that university students have difficulty grasping extended timescales, on the order of millions of years (DeLaughter et al. 1998; Libarkin et al. 2007). It is generally accepted that students create conceptions about the world from their experiences (e.g., Lederman 1992), and "Deep Time" is not something that one can experience.

The Global Science Literacy (GSL) states that the understanding of "Deep Time" is one of the central constructs of the geosciences (McPhee 1981; Mayer 1991a; Rudwick 1992) and forms one of the seven basic science understandings that construct science literacy (Hurd 1958; Clary et al. 2009). Geologic time is described in the GSL as a "key concept that spans natural science" and a "core element or critical barrier" (Trend 2002). The Earth Science Literacy Initiative (ESLI 2009), funded by the US National Science Foundation, developed a framework of underlying principles in earth science and identified the age of the Earth as the Big Idea #2 in earth science literacy. Many others agree that understanding "Deep Time" can change people's view of nature and humanity's role in it (e.g., Mayer 1991b; Lederman and O'Malley 1990) and that "Deep Time" is a key concept that affects other subjects (Dodick and Orion 2003). In addition, several books (e.g., Gould 1987; Haber 1959; Gorst 2001, and references summarized by Dean 1981) and papers address the topic of "Deep Time" (e.g., Patterson et al. 1955; Ault 1982; Trend 2001; Dodick and Orion 2003; Clary et al. 2009) and state its importance. Thus, teaching high school students about geochronology provides an opportunity to introduce students to a key concept of geology, address pre- or misconceptions, and improve science literacy. Toward this goal, we developed a teaching module to introduce high school students to how geologic time is measured and used for understanding Earth processes.

Following Piaget's constructivist learning theory (Piaget 1967), inquiry-based teaching and hands-on activities have been shown to facilitate learning complex topics (e.g., Tobin 1990; Mintzes et al. 1998). In addition, a recent study showed that improved collaboration between research scientists and K-12 (primary and secondary) educators enriches student experiences using otherwise unavailable resources (Gosselin et al. 2003). For teachers in the USA, there are abundant opportunities for hands-on activities and field trips in science subjects such as biology and chemistry, but we observe a dearth of similar opportunities in the geosciences, which are usually taught as elective sciences class in schools; the US National Science Teachers Association (NSTA) has been trying to elevate earth science in parity with physical and biological sciences. Field trips can be a great way to engage students with different topics, but due to time and money restrictions, they are not always feasible. Our feedback from teachers suggests that ready-to-use lesson plans in earth sciences would be greatly appreciated and could be another way to strengthen the application of earth sciences in school.

To create a productive education module on geologic time, we developed, adapted, and tested strategies for teaching U-Pb geochronology in different settings in the Greater Boston area, USA, over a 2-year period. This project was part of the EARTHTIME outreach initiative and was primarily developed by professors, postdocs, researchers, and graduate students in the MIT isotope lab. The goal of the project was to build theoretical and hands-on exercises to introduce students ages 14 and older to the concepts of "Deep Time" and geochronology. The modules we developed focus on introducing students to the techniques used for measuring geologic time and how these techniques can be used to study Earth history. To test the effectiveness of the teaching modules, we utilized multiple qualitative and quantitative datasets collected through questionnaires for students and science teachers. In this chapter, we will focus on data collected from 144 participants from one school, where pre- and post-tests were conducted to measure learning outcomes.

The purpose of this chapter is to describe the different parts of the outreach project and to use this example to qualitatively assess the effectiveness of our hands-on approaches in the classroom.

1.2 Description of the EARTHTIME Initiative

EARTHTIME is a scientific initiative aimed at sequencing Earth's history through international cooperation and collaboration (http://www.earth-time.org/). The initiative is supported in the USA by the National Science Foundation (NSF) and in the EU under the Framework Program 7. The need for this organizational structure to facilitate a better communication between geochronologists, stratigraphers, and paleontologists was recognized in 2003, and the aim of the EARTHTIME decade (2006–2016) is to create better tools to constrain Earth history through high-precision geochronology. A key component of the EARTHTIME initiative is educational outreach, with the goals of (1) developing educational tools with exercises and teaching material that provide students with multiple opportunities to explore difficult concepts in earth science and integrate new learning into their knowledge framework and (2) providing opportunities for students to interact with scientists and gain an understanding of how scientific research is done.

The MIT isotope laboratory has played a leading role in the EARTHTIME outreach initiative. The isotope lab specializes in the application of geochronology for understanding the rates and timing of geologic processes. Researchers in the lab are involved in a large number of ongoing projects on a wide range of topics, providing the opportunity to incorporate diverse research areas into outreach material.

2 Methods

2.1 Teaching Module

The goal of the teaching module we developed is to introduce students to the theory and techniques used in geochronology (Fig. 1). The module is based around U-Pb dating and includes theoretical exercises on radioactive decay and isotopic measurements and practical exercises on mineral separation. The combined exercises expose the

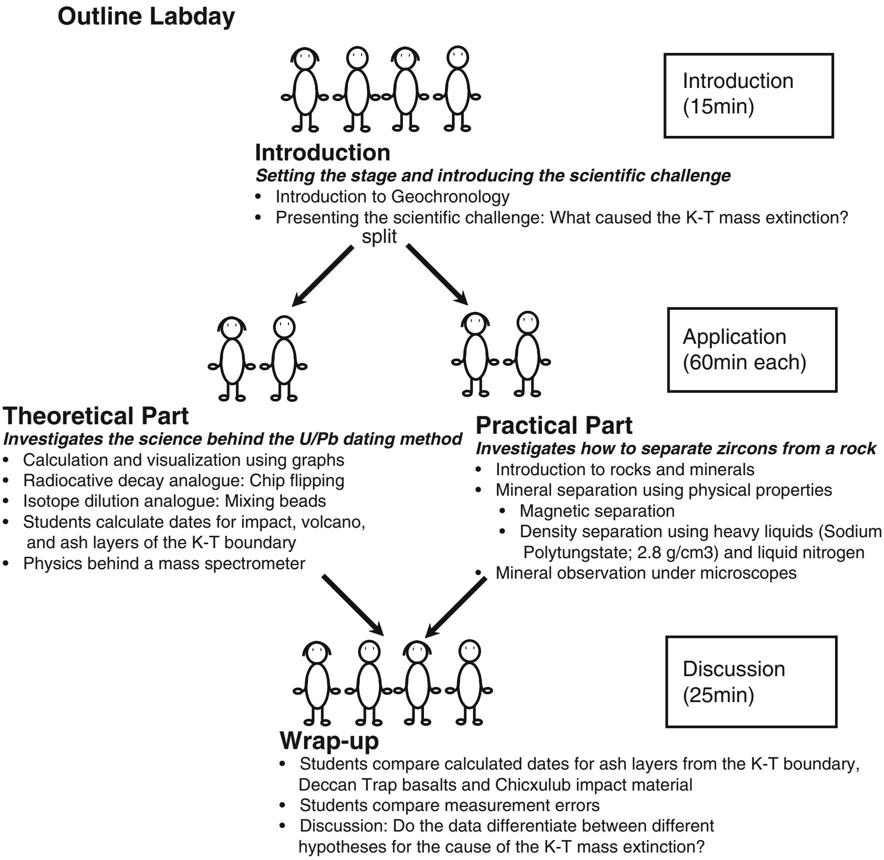

Fig. 1 Overview and outline of the Lab Day to demonstrate the different teaching actions (This figure was produced using the program Adobe Illustrator)

students to all of the steps in U-Pb dating, from sample collection to data interpretation. To provide a context for the exercises, we structured the module around the overarching theme of determining "What killed the dinosaurs?" However, the module is flexible and can be adjusted to a range of different topics (e.g., early Earth, climate change).

The experience starts with a short lecture that introduces to the idea of geologic time, the importance of geochronology for understanding Earth processes, and some background on the Cretaceous-Tertiary (K-T) boundary. The students are then split into two groups, to work on theoretical and practical exercises.

(A) *Practical Exercises*
In the practical section, students are introduced to the difference between rocks and minerals and examples of both. This exercise models the steps from collecting a rock sample to separating zircon, a key mineral in U-Pb geochronology.

Fig. 2 Students working with sodium polytungstate (yellow "heavy liquid") to separate the minerals by density. Liquid nitrogen (foreground) is used to freeze the heavy minerals in the bottom of the vial while the light minerals are poured out

Students go through a series of experiments to show how different minerals can be separated based on their physical properties. The techniques introduced in this section are regularly used in geochronology laboratories around the world. The mineral separation experiments include:

- Magnetic separation: Students are given a crushed rock sample and use strong magnets to separate magnetic and nonmagnetic minerals.
- Density separations: Students take the nonmagnetic minerals from the previous experiment and separate them based on density. The students place the nonmagnetic minerals in vials of sodium polytungstate (Fig. 2), a heavy liquid with a density of ~2.8 g/cm^3. The vials are put in a centrifuge and the dense minerals sink to the bottom, while the lighter minerals float on top. The students then freeze the bottom of the vials in liquid nitrogen and dump the less dense minerals into filter paper. After the dense minerals defrost, they are dumped into a separate filter paper.
- Mineral observation: Finally, students observe pre-prepared heavy mineral separates under an optical microscope with a large working distance (Fig. 3). The students discuss how different minerals look under the microscope and are given the opportunity to manipulate individual grains using fine-tipped tweezers.

Fig. 3 Students working with microscopes to pick zircons in the practical part

(B) *Theoretical Exercises*

In the theoretical section, students are introduced to isotopes, radioactive decay, and how isotope ratios and half-lives can be used to determine the age of a mineral. This content is described in detail in the online lesson plan for teachers. The students complete three exercises:

- Exercise 1: Students use worksheets to track the relative abundance of parent and daughter isotopes during radioactive decay. They plot changes in parent, daughter, and parent/daughter ratio through time. This exercise introduces radioactive decay, half-lives, and the use of parent/daughter ratios for determining the age of a mineral.
- Exercise 2: Students model radioactive decay by flipping chips with periodic table abbreviations for uranium on one side and lead on the other (Fig. 4). The students then determine the ratio of chips representing uranium to those that represent lead and use the ratio to calculate an age. This reinforces the concept that the parent/daughter ratio can be used to calculate an age, regardless of the initial number of parent isotopes.
- Exercise 3: Students use plastic beads to model isotopic ratio measurement made by isotope dilution. In the isotope dilution method, a known amount of a tracer isotope is added to a sample and used to determine the amount of

Fig. 4 Student working with the two-sided chips (part of the teacher material kit) to model radioactive decay in the theoretical part

Fig. 5 Students working with the beads (part of the teacher material kit) to determine ratios. The *colored beads* represent atoms and were mixed in a bowl beforehand for the isotope dilution exercise of the theoretical part (Color figure online)

> each isotope in the sample. The students are given a large tub full of clear beads, which represent ^{206}Pb atoms. To determine the total number of ^{206}Pb atoms, without counting every bead in the tub, a known number of ^{205}Pb atoms (red beads) are added to the tub and mixed in. The students then count the ratio of ^{206}Pb/^{205}Pb beads in small subsets of beads (Fig. 5) and then use the counted ratio and the known number of ^{205}Pb beads to calculate the total number of ^{206}Pb beads in the original tub. This provides a very good model for how isotope ratios are actually measured, and the data the students generate is used to introduce the concept of measurement uncertainties and to calculate the age of different events related to the K-T boundary.

After both units conclude, the students are reassembled to summarize and reinforce the information presented. Students pool their results from the first unit, where they determined the age of the extinction and two possible extinction mechanisms. The "dated" samples include ash beds from above and below the dinosaur extinction event within a sedimentary section and material related to both a large volume volcanic eruption in India (Deccan Traps) and a major impact structure in Mexico (Chicxulub crater), which have both been proposed as possible causes of the extinction. After a discussion about cause-and-effect relationships and geochronological data, students are then guided through interpretation of their data.

2.2 Teaching Settings

We named the educational module Lab Day because the main goal of the workshop was a laboratory, hands-on teaching style. We adapted and taught the module in three different settings over the course of 2 years: (A) "Lab Days" at the MIT museum, where students were brought to the museum as part of a daylong field trip; (B) "Lab Day on the road," where researchers brought equipment into a high school classroom; and (C) "Lab Day teacher workshop," where researchers met with teachers, demonstrated the module, and answered questions to help them incorporate it into their lesson plans.

(A) *Lab Days*

The first Lab Day event was held at the MIT museum in April 2008 as part of the internationally recognized Cambridge Science Festival, a weeklong event designed to promote interaction between the science, technology, engineering, and math communities in Cambridge, MA, and the general public. We advertised within schools in the greater Boston area for a geology-related school field trip. Three 12th grade classes (35 students) from different high schools were invited to join us for a 4-h workshop, followed by a video-streaming webcast, called "Q and A with the scientists." The program was successful and received positive feedback from students and teachers. Based on the feedback and our own observations, we adapted the lessons and expanded the program for the Cambridge Science Festival in April 2009. This event was again held at the MIT museum and included 9th and 10th graders (58 students) from two different schools. This time, as a prearrangement, students prepared short essays about their conception of geochronology. The outstanding results and the strong interest for outreach in geosciences led us to organize a third Lab Day later that year. In December 2009, we adapted the program again, with 9th grade students from one school (214 students) visiting the MIT museum over 3 days. The Lab Day program was shortened to facilitate the large number of students. During the event, the instructors were largely the same five scientists that led the other workshops. This time, to measure and compare students' understanding and to evaluate long-term results of our methods, pre-test were completed 1 week before the visit, and two post-tests were taken, 1 week and then 4 months after the visit. Detailed evaluation will be described below. To ensure a convenient campaign and participation of schools regardless their financial background, bus transportation and bagged lunches were provided for visiting students and teachers for all of the workshops.

(B) *Lab Day on the Road*
In May 2009, we modified the Lab Day exercises so that they could be transported to a classroom. In the first test of the new program, five scientists visited an all-girls school in Massachusetts to teach a class of 10th grade students. The scientists ran a 4-h workshop in the classroom using laboratory equipment brought to the school. Since the original design of the module utilized equipment available in the MIT museum, some presentations needed to be adjusted and replaced by posters and other activities. Written feedback from the students was collected 1 week after the visit.

(C) *Teacher Workshop*
After the initial Lab Days, we saw the need to include teachers in the program and to obtain more detailed feedback from educators, as well as to reach out to schools to advertise our program. In July 2009, as part of a teacher workshop at a local university, we introduced 27 science teachers to our curriculum. Written feedback from the teachers was obtained directly after the course and is not further addressed here.

2.3 Online Module

Based on our experiences in the Lab Day workshops, we produced a detailed lesson plan (31 pages) for the theoretical unit, which covers the principles of radioactive decay and isotopic dating. The lesson plan has been http://www.earth-time.org/Lesson_Plan.pdf from the EARTHTIME webpage since May 2009, and a teacher material kit is available by request in the USA. The kit and material provide a ready-to-use lesson plan, which was one of the most common requests in the teacher workshop. The online article describes teaching suggestions, teacher background knowledge, and learning goals and their correlation to national science concepts (a standard for unifying concepts and processes, given by the National Committee on Science Education). It provides worksheets for the exercises and a Microsoft Excel spreadsheet containing the necessary calculations. The material kit consists of three bags with 500 colored beads each (red, white, and blue), a bag with 100 two-sided chips (U/Pb on either side), a cup to take out samples, and a mixing bowl; a video demonstration of how to use the kit is available online. More than 50 kits have been requested and sent out to different schools in the USA so far.

3 Evaluation/Data Analysis

To evaluate the effectiveness of the module and use the participants as a resource for criticism, we solicited written feedback after each Lab Day presentation and used these comments to continually improve the exercises. To obtain more detailed statistical data for the December 2009 "Lab Day," we conducted a pre- and two different post-tests to compare short- and long-term learning outcomes. The pre-test was completed the week before the visit to the MIT museum, the first post-test was carried out 1 week after the visit, and the second post-test was done 4 months later. The middle and high school teachers gave each of the tests in their class room setting, and we had no influence on their implementation. However, we asked teachers not to help students with the questions. We also made it apparent that students would not be graded for their answers and that the survey would be solely for assessment of our teaching methods and would be treated anonymously. Students were given sufficient time to complete the questionnaires, which usually took less than 7–10 min. In the following sections, we focus on the results from the December 2009 Lab Days, because the pre- and post-tests from these workshops provide the best quantitative measure of the success of the teaching modules and Lab Day model.

The pre-test questionnaires consisted of the following four quantitative multiple-choice questions to test general content knowledge:

1. How old do you think the Earth is?
   (a) 1.23 million years
   (b) 2.34 thousand years
   (c) 3.45 trillion years
   (d) 4.56 billion years
   (e) 5.67 million years
2. What is a half-life?
   (a) The time when one half of a rock is eroded
   (b) The time when one half of the earth was formed
   (c) The time when one half of radioactive atoms decayed
   (d) The time when magma cooled enough to form a mineral structure
   (e) The time when half of the molecules have formed covalent bonds
3. What minerals are often used to date older rocks?
   (a) Quartz
   (b) Ruby (Corundum)
   (c) Zircon
   (d) Olivine
   (e) Obsidian
4. What is one dating method for determining the age of a really old rock sample?
   (a) Uranium-lead dating
   (b) Radiocarbon dating
   (c) Potassium-argon dating
   (d) Uranium-thorium dating
   (e) Tree-ring counting

We selected these questions because they provided the best quantitative measure of learning in our initial questionnaires from earlier "Lab Days." The first two questions deal with topics that might have been covered at some point in the school curriculum, while the last two questions are more specific to our module. In hindsight, we realized that question four was poorly worded, although it still likely serves as a useful monitor of the effectiveness of the modules (discussed in the study limitations). Both post-tests contained the same quantitative questions as the pre-test and two additional qualitative questions:

5. Name three things you learned at your Lab Day. (Three open answers possible)
6. What exercise was the most interesting and fun part of your Lab Day experience? (One open answer possible)

Parents/guardians were informed of the project and testing beforehand and signed written consent forms. Although students were granted anonymity, some of them did use their full names and others used initials or first names only. Names/initials were only used to match the pre- and post-tests. There was no distinction made between male and female students. Of the visiting 214 students, all three tests (pre-test and two post-tests) could be matched for 144 individuals ($n = 144$). This is due to some students not being present at one of the three testing dates in school or being unable to correlate pre- and post-tests due to missing names/initials. The long-term post-test from one participating group could not be obtained and was omitted from further consideration.

All tests results were normally distributed. To measure the difference between pre- and post-test, we ran two paired *t*-tests for the quantitative set of multiple-choice questions as a whole, one for the pre- and post-test and another one for the pre- and the long-term post-tests. We also investigated the change in each of the four quantitative questions and ran the paired t-tests for each question, examining the change from pre- to post-test, from pre- to long-term post-test, and also between the two post-tests. To quantify the effectiveness, we used the standard *t*-test *p* value (probability value): a small number indicates the module is effective while a large p value would indicate that the taught module seem to be ineffective. We defined the means of statistical significant difference as follows: p value <0.05 validates statistical relevance, marked with one *; p value <0.01 shows a strong relevance (**); p value <0.001 states high significance (***).

4 Results

4.1 Quantitative Questions

The results of the two paired *t*-tests examining the overall performance are shown in Fig. 6. Participants achieved significantly different performances in the overall pre- and post-phase and in the pre- and long-term post-phase, with a p value <0.01 in both cases. The increase in correct answers from the pre- to the first post-test demonstrated the effectiveness of our module: 67 % of the questions in the pre-test were answered correctly, while in the first post-test, 85 % were answered accurately. The long-term post-test shows a slight decrease in correct answers to 83 %, but the change is not

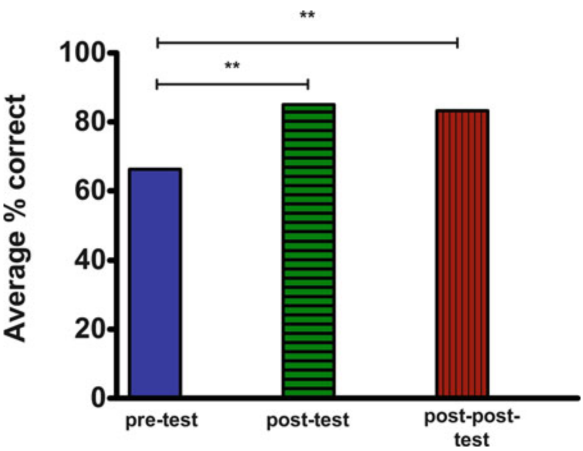

Fig. 6 Averaged results of the four quantitative questions for each test phase (percentage of correct answers). Statistical significance was measured by running two paired *t*-tests. $n = 144$, *p* value <0.01 for both *t*-tests. *Stars* indicate statistical significance as described in the Results section. Pre-test, 1 week before the visit: 67 % correct answers. Post-test, 1 week after the visit: 85 % correct answers. Post-post, 4 months after the visit: 83 % correct answers (This figure was produced using the program Graphpad)

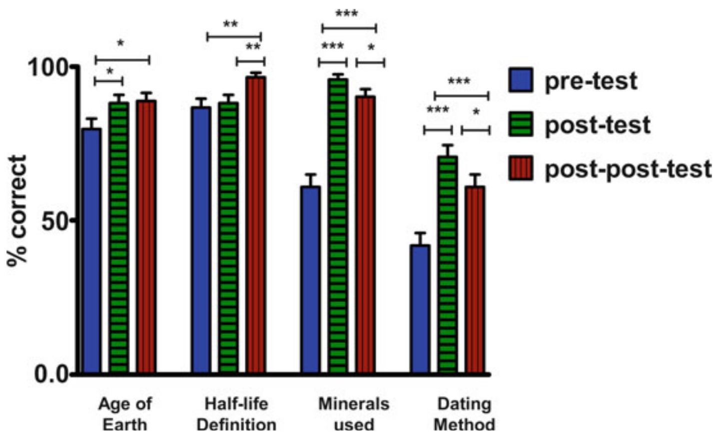

Fig. 7 Detailed chart with the four questions that were used for the paired *t*-tests for each test phase in Fig. 1. $n = 144$. *Stars* indicate statistical significance, as described earlier; *no star* indicates there is no statistical change between test phases (*p* value >0.05) (This figure was produced using the program Graphpad)

statistically significant (*p* value > 0.05). The decrease in correct responses was largely related to questions 3 and 4, visible in Fig. 7 and described below.

Figure 7 illustrates responses for the four questions used for the paired t-tests. Most students already answered question one (How old do you think the Earth is?) correctly in the pre-test (79 %). However, there was a statistically significant improvement for the post-test (87 %, *p* value $= 0.02$) and from the pre- to the long-term post-tests (88 %; p value =0.02). For the second question (What is a half-life?), the students did well in the pre-test (86 % correct answers), and there is no significant increase for the post-test (87 %; p value = 0.6). Interestingly, the half-life definition advanced further 4 months later for the long-term post-test (96 %) and thus shows a significant increase for both t-tests (p value =0.0015 from pre- to long-term post-test and $p = 0.004$ for post- to long-term post-test). This may be due to later reinforcement in students' secondary school curriculum. Overall, correct answers in the post- and long-term post-tests improved for these two questions.

For the third and fourth questions ("Which minerals are used to date older rocks?" and "What is one method used to date older rocks?"), there was very significant improvement in both post-tests relative to the pre-test. For the third question, 60 % answered the question correctly in the pre-test, whereas 95 % (p value <0.0001) answered it correctly in the post-test and 89 % (p value <0.0001) answered it correctly in the long-term post-test. For the fourth question, 40 % answered the question correctly in the pre-test, whereas 70 % (p value <0.0001) answered it correctly in the post-test and 60 % (p value =0.002) answered in correctly in the long-term post-test. The improvement can be attributed to our curriculum, since these topics were not covered in other classes.

Small score decreases for the long-term post-tests compared to the immediate post-tests are not significant for the overall performance (Fig. 6), but are significant when the questions are compared separately. The long-term retention of the knowledge was not as strong for the third and fourth questions (p value = 0.045 and 0.023, respectively). This may reflect the fact that the first two questions were likely reinforced in other classes. However, despite the slight decrease in correct answers between the two post-tests, the last two questions still exhibited large improvements from the pre-test to both the post-test and long-term post-test, demonstrating the positive impact of our program and its long-term benefit. Students may have later forgotten what they learned in the module and reverted back to familiar answers or preconceptions. For the third question, students correctly answered "zircon" a week after the visit, while 4 months later a significant percentage of students changed their answer to quartz, a better known mineral. For the fourth question, although most of the students correctly answered that the U-Pb system is used for dating very old rocks in the first post-test, some students changed their answer to "radiocarbon dating" 4 months later. Radiocarbon dating is commonly referenced in public and in the media, and it is sometimes used as a general term for measuring arbitrarily old dates. The distinction between different isotopic dating methods and which samples can be dated with which method is rarely made. This might explain why students ticked the well-known name radiocarbon instead of the lesser-known U-Pb dating method.

4.2 Qualitative Questions

Describing all the results of the qualitative questions of the pre- and post-tests is beyond the scope of this chapter. However, in this section we summarize the main outcomes and implications. All answers to the two qualitative questions were collected and then, if possible, categorized and summarized by topic.

Question #5 (*Name three things you learned at your Lab Day*)

For the first qualitative question, students could name three things they learned. Answers were categorized by similar answers or specific terms. Most students named zircons in some way. This would include the most stated answer "separate zircons from rock" as well as physical properties of zircons ("zircons are nonmagnetic" or "zircons are heavier/denser than other minerals"). Some even answered "zircons can be used to date rocks." The second most common category consisted of answers that mentioned the framework theme of our exercise, the K/T boundary and mass extinction. Both answers were still well represented in the long-term post-test 4 months later. Answers containing something about mass spectrometers and isotopic ratios were prevalent in the week after the experiment, but were marginal 4 months later. These again are topics that not usually covered in school, and even though they may have made an impression, without further reinforcement, they are not the first terms that come to students' minds when reflecting on the experience. On the other hand, categories that included radioactivity and half-lives were not the most common in the post-test, but gained more attention after 4 months – possibly because they had been covered in science classes and teachers could refer to our experiment. This suggests that out-of-school trips are an effective way to introduce a new topic, which can then be further discussed in the classroom to reinforce understanding and establish a long-term effect.

Question #6 (*What exercise was the most interesting and fun part of your Lab Day experience*?)

The second qualitative question was easier to categorize. Students were asked to name the most fun and interesting part of their experience, and we divided the responses based on whether the experience they listed was part of the practical or theoretical section. The practical and theoretical parts were then further subdivided into the different exercises in each section that were most popular. The practical part was clearly favored by students: 84 % (88 % in the long-term post-test) of the students named one of the hands-on laboratory elements as their favorite part in the two post-tests, while 9 % (11 % in the long-term post-test) named fractions of the theoretical part as more fun. The most favored practical exercise was using the liquid nitrogen to freeze heavy minerals in the sodium polytungstate, and the most favored theoretical exercise was the isotope dilution bead problem. These responses support our hands-on teaching approach.

4.3 General Written Feedback from Teachers and Students

Teachers noted that there is a need for educational modules to be taught in different settings as well as with different applications. Our approach with the geochronological relevance of radioactivity was praised by geology, chemistry, and physics teachers as a unique approach. Even though the laboratory modules cannot be carried out in school due to lack of equipment, our material kit was stated to be an excellent representation of radioactive decay. Some teachers made it clear that they were trying to avoid the topic of radioactive decay but now feel more confident in teaching it after completing the teacher workshop and being able to use the bead model.

Notable student quotes included the following: “cool doing something new and different,” “didn’t know rocks could be so cool,” “interesting that mathematics can actually be used for exciting topics,” and “nice to be at MIT and work with real scientists and see what they do.” The students valued being included in something that felt like how science is done in a lab, and they were fascinated to see the dedication to our jobs/studies and that scientists are excited about what they are doing. They were amazed to hear how many different places geologists can visit and conduct research in and what a broad range of topics are included in geology.

4.4 Study Limitations

All our results were confirmed by testing day-to-day variation in the responses (not further explained here), so we are confident in our positive learning outcomes in general. However, testing methods always have drawbacks, and we would like to illustrate possible factors: for question four, we did not clearly quantify “really old.” The uranium-thorium and radiocarbon techniques can be used to date samples that are thousands of years old and so could be considered correct answers. In addition, potassium-argon dating can be used to date geologic events over the same time range as uranium-lead dating and is also an appropriate answer. Because the Lab Day teaching module stressed the uranium-lead technique, and other techniques such as potassium-argon are likely not well known to the students, we do feel that the statistical variations in the number of “correct” answers likely reflect the amount of material the students retained from the Lab Day exercises, despite the poor wording of the question. We could not supervise the testing in class, and even though we asked students to give their true opinion and told them they would not be graded, we can never check if they copied ideas from one another. Although we asked teachers not to mention specific terms, it is possible that a teacher or a student made a well-intentioned suggestion (e.g., “remember the Lab Day where we did…”) that could have altered the data. Also, not all long-term post-tests came back, and in general, the feedback from teachers and students still seems to be limited when written. An oral feedback discussion would be ideal but was not feasible due to time constraints. Also, another third longer-term post-test (e.g., 8 or more months after teaching the module) would be useful for assessing the long-term benefit of the Lab Day model. In general, the very positive test results might not be completely representative for a typical 14-year-old high school student. One of the teachers we contacted for the project was already known to be interested in the subject and therefore might have covered parts of the topic in class beforehand. Also, teachers would not want students to look too uninformed when visiting a research university, so the topic might have been covered briefly beforehand. Finally, nonresponses may have selectively biased the later tests toward students who were confident of the correct answers.

5 Outcome of the Project and Conclusions

To quantitatively assess learning outcomes and the long-term impact of the outreach program, we used pre- and post-testing. The data show that many students learned and retained knowledge of U-Pb dating from the practical and theoretical exercises. We used the additional input to adapt the lesson plan and provide more detailed instructions for science teachers. The assessment results demonstrate that complex concepts are retained over the short and long term. However, topics that have not yet been covered in school and are not repeated after the visit are not retained as well over longer timescales. We suggest reinforcing difficult concepts in multiple settings (i.e., out of school and in school) and that a wrap-up after school trips might have a stronger impact on learning. In general, we propose that conducting outreach with scientists is a highly successful way of engaging students and should be a part of every research facility to foster curiosity and appreciation of science.

From this study we are able to conclude that teachers appreciated the hands-on activities that placed complex subjects in a wider context and ready-to-use lesson plans, especially in interdisciplinary subjects such as the geosciences that are not usually part of the school curricula. Teachers also welcomed the chance to further explore unfamiliar material during workshops taught by research scientists. Students generally enjoyed the hands-on laboratory experiments and the opportunity to encounter and scientifically evaluate an ongoing research project with professional scientists.

Our results also show that the EARTHTIME outreach initiative and its efforts to bring cutting-edge scientific research to schools and the public are effective at fostering scientific literacy at an early age. We created an educational module to provide students with hands-on exercises in geoscience, which covered material from geology, mathematics, physics, and biology. Continuous feedback from students and teachers helped adapt the module to their needs.

5.1 Perspective

The ready-to-use lesson plan for teachers with clear instructions is downloadable from the EARTHTIME web page (www.earth-time.org). Material kits accompanying the lesson plans have been sent out to more than 50 schools US-wide, and the module has been presented at three international conferences. In Vienna, Austria, the German translation of the lesson plan has been adapted for younger students and is in frequent use with the material kit, e.g., at the Children's University.

Overview

Background and Motivation

– An educational module about uranium/lead dating and geologic time was developed as part of the EARTHTIME outreach initiative. Our goal was to combine theoretical exercises and hands-on experiments that expose high school students (age 14+) to real research methods and problems. The module builds on multiple existing school subjects, including physics, chemistry, and biology. The ambition was to test whether complex scientific topics can be successfully taught in high schools, using hands-on activities to model multifaceted processes.
– We have provided a convenient lesson plan that covers topics of general interest and incorporates cutting-edge research, because science teachers often do not have the chance to be up-to-date on current research.

Innovations and Findings

– Students appreciated exploring a new topic in a different environment, and students and teachers both enjoyed interactions with "real scientists about a real scientific topic."
– Although the curriculum was demanding, we observed a statistically significant increase in short- and long-term learning outcomes for the program.
– Students also made connections between different branches of science and observed that mathematics and physics could be applied to solve problems in earth science.

Implications for Wider Practice

– Our data suggests that high school students benefit from visiting earth science research facilities. We suggest that even complex topics taught by scientists using hands-on activities can facilitate learning. If teachers have accurate teaching materials and training, they can work together with students to teach sophisticated scientific concepts. For a research institution, the development of such a module requires commitment in administration.
– We propose that topics that seem to contend with general preconceptions need to be addressed further, repeatedly, and with different teaching methods to gain a durable effect. Also, to accomplish an understanding and knowledge gain for all students with different learning skills, it might build a longer-lasting impact if challenging topics are covered in lessons in school as well as field-trip experiments with different applications.
– With the teaching module and kit, we would like to give teachers the opportunity to address different topics without a field trip, even if these can only work as inspirational samples since teachers already have to cover a substantial amount of science topics in classes.

Acknowledgments We would like to thank Jessica Creveling, Stacia Gordon, and Christy Till for helping with the Lab Day programs. We very much appreciate the help of all the participating students and teachers, especially Rita Chang from Wellesley High School, Massachusetts.

References

American Association for the Advancement of Science (AAAS). (2009). http://www.project2061.org/research/assessment.htm

American Geophysical Union. (1994). *Report of the AGU Chapman conference on scrutiny of undergraduate geoscience education*, 55 p.

Ault, C. R. (1982). Time in geological explanations as perceived by elementary school students. *Journal of Geological Education, 30*, 304–309.

Clary, R. M., Brzuszek, R. F., & Wandersee, J. H. (2009). Students' geocognition of deep time, conceptualized in an informal educational setting. *Journal of Geoscience Education, 57*(4), 275–285.

Dean, D. R. (1981). The age of the earth controversy: Beginnings to Hutton. *Annals of Science, 38*(4), 435–456.

DeLaughter, J. E., Stein, S., Stein, C., & Bain, K. (1998). Preconceptions about Earth science among students in an introductory course. *Eos, 79*, 429–432.

Dodick, J., & Orion, N. (2003). Measuring student understanding of geological time. *Science Education, 87*(5), 708–731.

Earth Science Literacy Initiative (ESLI). (2009). www.earthscienceliteracy.org/

Gorst, M. (2001). *Measuring eternity. The search for the beginning of time*. New York: Broadway. 352 pages.

Gosselin, D., Levy, R., & Bonnstetter, R. (2003). Using earth science research projects to develop collaboration between scientists at a research university and K-12 educators: Insights for future efforts. *Journal of Geoscience Education, 51*(1), 114–120.

Gould, S. J. (1987). *Time's arrow, time's cycle: Myth and metaphor in the discovery of geologic time*. Cambridge: Harvard University Press.

Haber, F. C. (1959). *The age of the world. Moses to Darwin*. Baltimore: Johns Hopkins Press.

Hurd, P. D. (1958). Science literacy: Its meaning for American schools. *Educational Leadership, 16*, 13–16.

Hutton, J. (1788). Theory of the Earth. *Transaction of the Royal Society of Edinburgh*, I(Part II), 209–304. http://www.uwmc.uwc.edu/geography/hutton/hutton.htm

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29*, 331–359.

Lederman, N., & O'Malley, M. (1990). Students' perceptions of the tentativeness in science: Development, use, and sources of change. *Science Education, 74*, 225–239.

Libarkin, J. C., Kurdziel, J. P., & Anderson, S. W. (2007). College student conceptions of geological time and the disconnect between ordering and scale. *Journal of Geoscience Education, 55*, 413–422.

Lyell, C. (1830). *Principles of geology, being an attempt to explain the former changes of the Earth's surface, by reference to causes now in operation*. London: John Murray. Volume 1–3.

Mayer, V. J. (1991a). Earth-system science: A planetary perspective. *The Science Teacher, 58*(1), 31–36.

Mayer, V. J. (1991b). Framework for Earth systems education. *Science Activities, 28*(1), 8–9.

McPhee, J. (1981). *Basin and range* (229 p.). New York: Farrar, Straus, and Giroux.

Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (Eds.). (1998). *Teaching science for understanding. A human constructivist view*. San Diego: Academic.

National Research Council. (1997). *Science teaching reconsidered*. Washington, DC: National Academy Press. 88 p.

National Science Board. (2002). *Science and engineering indicators*. Arlington: National Science Foundation.
National Science Board. (2003). *The science and engineering workforce: Realizing America's potential*. Arlington: National Science Teachers Association (NSTA), National Science Foundation. http://www.nsta.org/
National Science Foundation. (1996). *Shaping the future: New expectations for undergraduate education in science, mathematics, engineering, and technology*. Arlington: National Science Foundation. 76 p.
Patterson, C., Tilton, G., & Inghram, M. (1955). Age of the Earth. *Science, 121*, 69–75.
Philips, W. (1991). Earth science misconceptions. *The Science Teacher, 58*(2), 21–23.
Piaget, J. (1967). *Logique et Connaissance scientifique (Logic and scientific knowledge)*. Paris: Encyclopédie de la Pléiade.
Rudwick, M. J. S. (1992). *Scenes from deep time* (294 p.). Chicago: University of Chicago Press.
Tobin, K. (1990). Research on science laboratory activities: In pursuit of better questions and answers to improve learning. *School Science and Mathematics, 90*(5), 403–418.
Trend, R. (2001). Deep time framework: A preliminary study of UK primary teachers' conceptions of geological time and perceptions of geoscience. *Journal of Research in Science Teaching, 38*, 191–221.
Trend, R. (2002). Developing the concept of deep time. In V. J. Mayer (Ed.), *Global science literacy* (pp. 187–202). Heidelberg/Berlin: Springer.