ASTROBIOLOGY
Volume 18, Number 1, 2018

DOI: 10.1089/ast.2016.1550

Education Article

Habitable Worlds: Delivering on the Promises of Online Education

Lev B. Horodyskyj,[1] Chris Mead,[1] Zack Belinson,[2] Sanlyn Buxner,[3] Steven Semken,[1] and Ariel D. Anbar[1,4]

Abstract

Critical thinking and scientific reasoning are central to higher education in the United States, but many courses (in-person and online) teach students information *about* science much more than they teach the actual process of science and its associated knowledge and skills. In the online arena specifically, the tools available for course construction exacerbate this problem by making it difficult to build the types of active learning activities that research shows to be the most effective. Here, we present a report on *Habitable Worlds,* offered by Arizona State University for 12 semesters over the past 6 years. This is a unique online course that uses an array of novel technologies to deliver an active, inquiry-driven learning experience. Learning outcomes and quantitative data from more than 3000 students demonstrate the success of our approach but also identify several remaining challenges. The design and development of this course offers valuable lessons for instructional designers and educators who are interested in fully capitalizing on the capabilities of 21st-century technology to achieve educational goals. Key Words: Online education—Active learning—SETI—Astrobiology—Teaching. Astrobiology 17, 86–99.

1. Introduction

As societal problems and their solutions become more strongly dependent on a firm grasp of science and its methods, it becomes critically important to provide students with authentic and meaningful science-learning experiences. For this reason, Arizona State University (ASU), like most of its peer institutions, requires each of its degree-seeking undergraduates to complete at least one four-semester-hour laboratory science course. For many students, this course is a terminal science course (their last formal science instruction for the rest of their lives). This final experience with science will influence their personal attitudes and public engagement with scientific issues (*e.g.,* Hobson, 2008).

Universities are offering an increasing number of online courses, many of which rely solely on learning management systems (LMSs) such as Blackboard or Moodle to present course material. These tools provide limited templates for delivering content, favoring a passive, teacher-centered mode of instruction using video lectures and quizzes rather than enabling discovery and exploration. Research shows that this is inadequate for teaching the authentic nature of science (Songer and Linn, 1991; Freeman *et al.,* 2014). In building *Habitable Worlds,* we aimed to address deficiencies in current online science education offerings to help ASU deliver on its mission of teaching science and scientific literacy to its students. Recently, the course has also been adopted by astrobiology faculty at other institutions who have similar needs. This highlights the need for a complete review and analysis of the course and our lessons learned. In this report, we describe our design decisions and rationale as well as present evaluation data that identify the strengths and weaknesses of this endeavor, with the aim of allowing others in the community to learn from our process and approach.

1.1. What's been done before?

The lecture-lab course structure common in undergraduate introductory-level science courses is designed to teach students both the *products* of scientific investigation (conveyed via lecture) and the *process* of scientific investigation (via lab). The lecture and lab components of this structure have been criticized as ineffective (Hofstein and Lunetta, 2004; Freeman *et al.,* 2014). Too often the lecture portion of the class employs teacher-centered, passive learning (Freeman *et al.,* 2014), while the lab portion focuses on verification

[1]Center for Education Through eXploration and School of Earth and Space Exploration, Arizona State University, Tempe, Arizona, USA.
[2]Smart Sparrow, Sydney, Australia.
[3]Department of Teaching, Learning, and Sociocultural Studies, University of Arizona, Tucson, Arizona, USA.
[4]School of Molecular Sciences, Arizona State University, Tempe, Arizona, USA.

(Hofstein and Lunetta, 2004). In addition, the credit load of the lecture versus lab (often 2–3 credits for the lecture and 1 for the lab) reinforces the idea that the body of knowledge is more important than the process used to generate that knowledge. As a result, students typically leave introductory-level science courses with the inaccurate perception that science is primarily about the discovery of information by experts and that their role as students, and eventually as citizens, is to accept and memorize these discoveries. Ultimately, the typical lecture-lab modality does not teach the practice of science so much as it teaches decontextualized facts about science.

Despite the inherent problems of the lecture-lab paradigm, this suboptimal design has been translated with little change into digital formats in recent years (Toven-Lindsey *et al.,* 2015). Many online science courses convert lectures into videos accompanied by quizzes and replace the physical lab with a virtual facsimile, ultimately delivering a less effective experience for students (Fig. 1). Beginning in 2012, public awareness of online education received a significant boost via high-publicity offerings of massively open online courses (MOOCs) by the university alliances Coursera and edX (Friedman, 2012). Both initiatives boasted enormous enrollment numbers, showing high demand for convenient and free access to higher education. However, despite the fanfare and eye-popping enrollment numbers, the success of MOOCs has been equivocal. Completion rates for these courses are low, around 5–10% (Jordan, 2014), and have not fundamentally altered the traditional method of science instruction.

Some in-person science courses do deliver an authentic view of science (*e.g.*, Deslauriers *et al.,* 2011; Holmes *et al.,* 2015). Such courses emphasize active learning, including scientific inquiry or problem-based learning. Prior research is overwhelming in its support of active learning (Hake, 1998; Springer *et al.* 1999; Lorenzo *et al.,* 2006; Haak *et al.,* 2011; Ruiz-Primo *et al.,* 2011; Freeman *et al.,* 2014), so much so that recent commentaries have asked whether it is even ethical to teach without active-learning techniques (Wieman, 2014; Waldrop, 2015). The remaining challenge is that courses of this nature are inherently difficult to offer at scale due to the resources needed to supply continuous feedback as students navigate complex tasks. Although a better way of teaching science, this paradigm had not been previously translated into the online environment at large scales.

1.2. Philosophy of digital design

At its best, computer-based education allows for educational experiences that would be impossible or impractical otherwise. With *Habitable Worlds* we sought to meld new technologies with research-based inquiry techniques to apply the principles of active learning at scale (*e.g.,* Schwartz *et al.,* 2004). To achieve this, we designed *Habitable Worlds* so that students learn important science knowledge primarily through exploration and experimentation, the way a scientist would. Therefore, during the course students

- Make observations;
- Build models to explain observations;
- Use their models to make predictions;
- Adjust models accordingly; and
- Integrate models across disciplines to solve complex problems.

Iteratively, students refine and strengthen their models. With the addition of equations and supporting information, we help students flesh out their models to become a good approximation of the current scientific model of the same phenomenon. Rather than describing a concept to students and asking them to master it, we give students room to discover the concept themselves through observations and modeling. This procedure is concordant with inquiry learning-cycle models such as those of Lawson *et al.* (1989) and Bybee *et al.* (2006).

Traditional science labs are intended to teach scientific reasoning and lab techniques. In *Habitable Worlds,* we dispensed with teaching laboratory skills since they are mostly irrelevant to a general-education audience. Instead, we focused on teaching scientific reasoning. This is easier to do in an online environment than a physical one because we can use complex simulations based on real scientific models, huge data sets, and randomization to create activities that require critical thinking and concept mastery to solve. This design is heavily inspired by video games, in which players master the game universe rules to solve complex problems. In *Habitable Worlds,* the rules of the game are the actual rules of the Universe as we currently understand them.

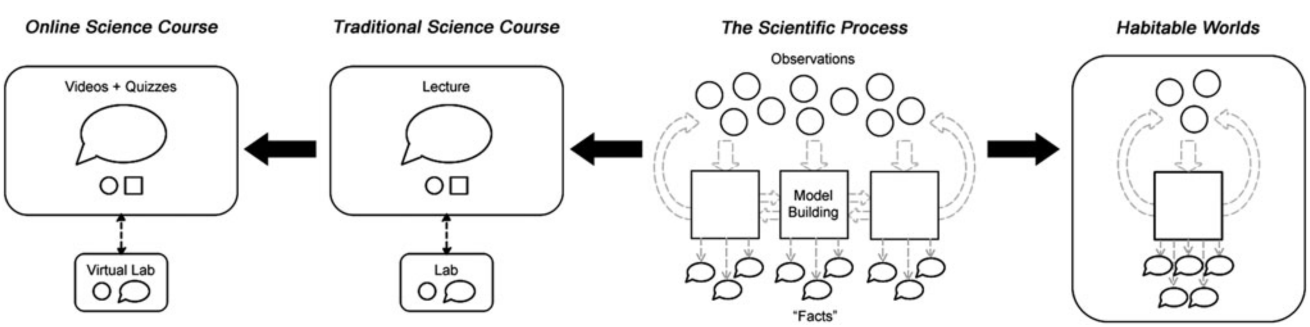


FIG. 1. The process of science focuses on observations (circles), constructing logical models to explain observations (squares), and using those models to predict future observations. The public perception of science is that it generates ''facts'' (speech bubbles). General education science courses and their online derivatives focus predominantly on the *products* of the process to the detriment of the *process* itself. *Habitable Worlds* focuses on observations and model building, allowing students to participate in the scientific process in addition to learning fundamental concepts. This approach more closely replicates the scientific endeavor.

Table 1. Digital Design Philosophy and Implementation in *Habitable Worlds*

Digital design element	Habitable Worlds *implementation*	*Provider*
Adaptive feedback Adaptive pathways Adaptive design Easy authoring and updates	Intelligent tutoring system	Smart Sparrow
Observing and modeling	Simulators, virtual field trips	Custom-built by Smart Sparrow (sims), in-house (VFTs)
Problem-based learning	The project	Custom-built by Smart Sparrow
Asynchronous instructor-student interactions	Discussion forum	Piazza
Synchronous instructor-student interactions	Live chats	Adobe Connect
Community building	Discussion forum and live chats	Piazza, Adobe Connect

We also sought to prevent rampant cheating in the course. Cheating in many online courses is trivial because the inflexibility of authoring tools restricts course design to question-and-answer systems that are easily gamed. To minimize cheating in *Habitable Worlds,* we used adaptive design, whereby problem sets, initial experimental conditions, and simulator setups are randomly generated to create a unique version of an activity for each student.

Students cannot be expected to learn the process of science in a completely unstructured and unguided environment. Nor will they quickly master new material if their mistakes are only noted hours or days after they are made. Therefore, *Habitable Worlds* provides

- Automated feedback specific to the mistakes students make when they make them;
- Alternative pathways to remediate poor performance;
- Instructor interactions (synchronous and asynchronous) for student problems that cannot be addressed through automation; and
- Community building to extend the topic of study beyond the required activities.

Finally, creating, maintaining, and improving such a course requires tools that give a considerable amount of control to the instructor. Many existing solutions are limited. Publisher-provided content often needs to be taken as is, giving instructors little ability to adapt content for their particular contexts. Custom-built solutions are expensive and likewise give instructors little ability to update or adapt content after the initial build. For *Habitable Worlds,* we built the majority of the course in an authoring environment (presented in more detail below) that provides us with full control over text, graphics, layout, and adaptivity and also captures analytical data about student performance. This empowers us to make informed edits and updates to our content.

Habitable Worlds implements this digital design philosophy as detailed in Table 1, Fig. 1, and below.

2. Course Design

2.1. Application to astrobiology

Astrobiology is an ideal subject for a terminal science lab course for nonmajors, as it allows students to see how complex problems can be solved by using knowledge from multiple disciplines. To that end, we used the principles summarized above to develop *Habitable Worlds,* a four-semester-hour, online-only lab course built around a fundamental scientific and humanistic question: ''Are we alone?'' To explore this question, astrobiologists synthesize concepts and practices from diverse scientific disciplines that include astronomy, geology, and biology.

The curriculum for *Habitable Worlds* (Table 2) uses the Drake equation (Fig. 2) as an organizing theme (Drake, 1961). Although it is presented using mathematical language, the Drake equation is probably best characterized as a way of organizing our uncertainties about the search for intelligent, communicating, alien life, with the better-quantified terms on the left and the less-constrained ones on the right. This makes the equation a perfect framework for *Habitable Worlds* because it requires understanding of many scientific disciplines and because the relative lack of certainty of the later terms provides an authentic illustration of uncertainty in science.

2.2. Course structure

Habitable Worlds consists of a project, training activities, and application activities, each of which count for approximately one-third of the course grade.

2.2.1. Project. The capstone activity for the course is a project that consists of a specially developed simulator. In the project, students are tasked with identifying a habitable world in a field of 500 randomly assigned artificially generated star systems, pulled from a background database. The project is treated as the final examination. Week by week, students learn the techniques necessary to solve the project via training activities. They are then tasked with synthesizing this knowledge into a methodology to solve the project. The project is deliberately designed to be large enough that students are not tempted to brute-force the project and are instead prodded toward developing a strategy. This approach mimics a genuine scientific experience. Scientists are often confronted with overwhelming problems that are only approachable once broken down into smaller, more addressable components. Because the project is randomized for each student, it is not possible for one student to complete the project and tell the others where the habitable worlds are, but successful strategies can be shared, which we encourage. This type of active, collaborative learning is known to be effective in greater learning and often greater self-efficacy, an important predictor of future learning success (Bandura, 1982; Robbins *et al.,* 2004).

TABLE 2. THE TYPICAL SYLLABUS FOR *HABITABLE WORLDS*, INCLUDING THE KEY TOPICS COVERED

Unit	*Title*	*Key topics*
1	Introduction	Introduce students to cognition, the scientific process, the scale of time and space, and early attempts at answering the question ‘‘Are we alone?’’
2	Stars (R*)	Students explore stellar properties and their relationships with each other through experimentation, including parallax, luminosity, color, spectra, and stellar lifecycles.
3	Planets (f_p)	Students discover planets and basic physical relationships via transit and radial velocity data and explore their composition through meteorite analysis and modeling.
4	Water (n_e)	Students explore water stability and the conditions that allow liquid water to remain stable on a planet’s surface, including surface pressure, energy balance, greenhouse effect, and climate.
5	Habitability (n_e)	Students explore plate tectonics and geochemical cycles to understand the habitability of Venus, Earth, and Mars, then use geology to explore the past habitability of Earth and how it has changed.
6	Life (f_l)	Students investigate the components of life and their formation, as well as redox chemistry and biosignatures.
7	Conclusion (f_i, f_c, L)	Students apply techniques learned in earlier units (Planck function, light curves) to the search for intelligence and technosignatures. Students also learn about sustainability and the ultimate end of habitability on Earth.
2–7	Project	Students apply combined skills from previous units to find habitable worlds in a field of 500 randomized stars.

2.2.2. Training activities. The bulk of learning takes place in training activities. In these activities, students explore a given topic (*e.g.,* phase equilibria of water, ancient climates) through observational and experimental pieces, immersive virtual field trips, short video ‘‘lecturettes,’’ images, and text. There are no limits on attempts or time on an activity, allowing students to make as many mistakes as necessary to learn the concept. Many of these training activities are designed in anticipation of student misconceptions, using mistakes as learning moments rather than penalizing students for making them. Students earn course points by passing certain checkpoints through a sequence of activities, creating motivation for students to continue to the end while also incentivizing necessary failures, asking questions, or exploring the parameter space and limits of a model.

2.2.3. Applications. The application activities serve as evaluations. Students are awarded points based on their mastery and synthesis of concepts taught in the training activities. Students have full access to training activities during any evaluation piece, which reinforces the idea that in the information age, memorizing information is less important than being able to solve problems with that information. For evaluation activities, students earn more points for successfully completing activities on their first try (regardless of time) than if they take more tries to complete. Activities in an application can include calculations, evaluations based on criteria learned in training activities, setting up simulators, or a combination of these.

$$N = R^* \times f_p \times n_e \times f_l \times f_i \times f_c \times L$$

where

N = number of communicating civilizations in our galaxy
R^* = rate of star formation
f_p = fraction of stars with planets
n_e = number of habitable worlds per star system
f_l = fraction of habitable worlds with life
f_i = fraction of life that becomes intelligent
f_c = fraction of intelligent life that communicates
L = lifetime of communicating civilizations

FIG. 2. The Drake equation (Drake, 1961).

3. Specific Technology Elements

In building *Habitable Worlds,* we converged on several technological solutions to implement our digital design philosophy. These include an intelligent tutoring system, simulators, immersive virtual field trips, and online platforms for discussion groups and teleconferencing. Below we detail the specific technology elements we utilized and how they allowed us to implement our digital design philosophy.

3.1. Authoring environment: Adaptive eLearning Platform (AeLP)

Over the past few years, we have worked closely with the educational software company Smart Sparrow (www.smartsparrow.com), developer of the Adaptive eLearning Platform (AeLP), an intelligent tutoring system (ITS), built specifically for science instruction (Ben-Naim, 2011). The AeLP is structured similarly to a PowerPoint presentation, with a stage on which objects can be placed, and a list of slides (Fig. 3). Objects that can be added to the stage include text, images, simulations, and a variety of input fields. The instructor can also add one or more conditional statements that will trigger feedback in response to student actions.

In the simplest example of adaptivity, an instructor can determine the correct states of the inputs on the stage, and/or a range of acceptable answers. A correct response will trigger appropriate feedback and send the student to the next question, while an incorrect response will trigger alternative feedback that asks the student to try the task again. In more complex adaptivity setups, the instructor creates custom *trap-states* that trigger specific feedback for anticipated

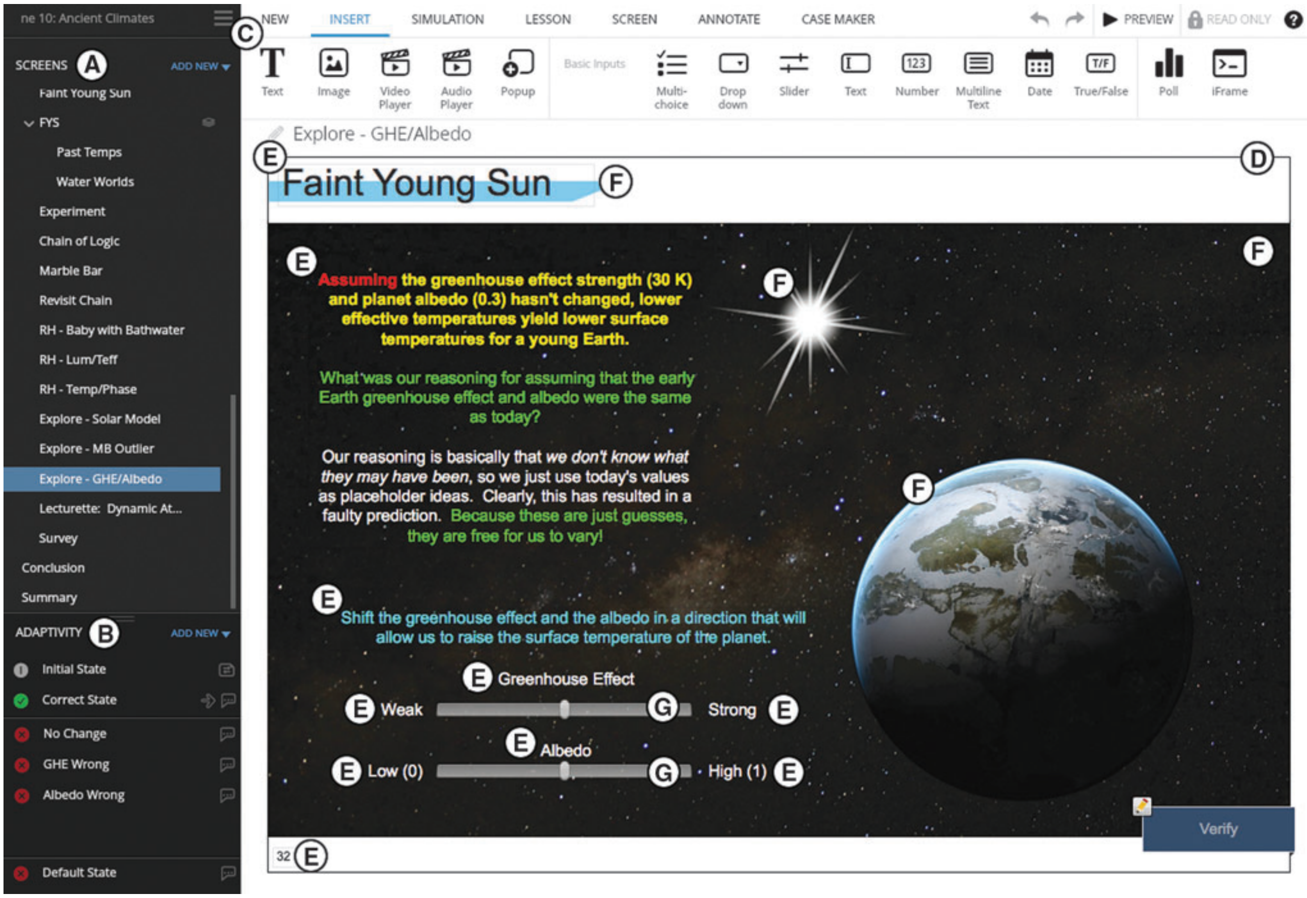


FIG. 3. The AeLP author environment, showing (**A**) the list of screens in the lesson, (**B**) adaptivity panel where trap-states are set up, (**C**) top ribbon, which shows available authoring actions and tools, and (**D**) stage where the lesson is set up. In this example, the page consists of (**E**) textboxes, (**F**) images, and (**G**) slider inputs. The adaptivity panel (B) allows us to determine what feedback/redirection student will see/experience in response to actions involving the slider inputs (G). The lesson pages, page adaptivity, and elements on the stage are easy to move, rewrite, and replace.

mistakes (''traps''). This allows for more fine-grained responses and guidance for students based on their specific errors, rather than a simple correct/incorrect dichotomy. In expert adaptivity setups, the instructor can create complex logic that provides individualized responses, manipulates objects on the stage, and customizes lesson flow based on student abilities (either remediation activities or bypasses of certain content). The AeLP is sufficiently powerful and flexible that it can allow for students to participate in true experimentation by allowing instructors to (a) identify the conditions of success, (b) define trap-states that look for flaws in students' experimental structure, (c) allow students the freedom to approach and complete the experiment as they see fit (under any conditions specified by the instructor), and (d) randomize aspects of the experiment to force students to complete their own work.

Most significantly for teaching and research on student learning, the AeLP includes powerful analytical tools that allow the instructor to see how students interact with activities in near-real-time. Every time a student submits his or her work on a particular page, a variety of properties are automatically documented, including timestamp, answers, and the trap-states that were triggered. The instructor can monitor, for example, the median time students spend on a page, the average number of attempts, and sequence of errors that students make (Fig. 4). The AeLP allows for easy extraction of data for further analysis and curriculum improvement. For example, frequent incorrect answers by many students may reflect confusing instructions or structure, or they may indicate conceptual difficulty.

Over the years, the course has been improved by using the analytical data captured by the AeLP, which has allowed us to identify areas of weakness. Problems that students experience can often be solved by adding additional feedback or clarifying existing instructions. Each ''intelligent'' trap-state must be programmed beforehand, meaning that unanticipated wrong answers will receive a generic response. By examining each semester's response data, we have found common wrong answers that were not being trapped and added new granular feedback. We have also added optional pathways designed to help students who appear to be struggling with prerequisite material or, in some cases, students who try to guess their way through an exercise. In addition, we are currently investigating the patterns of answers that some of our weaker students display with the goal of identifying these students earlier and supporting them through material with which they have more difficulty.

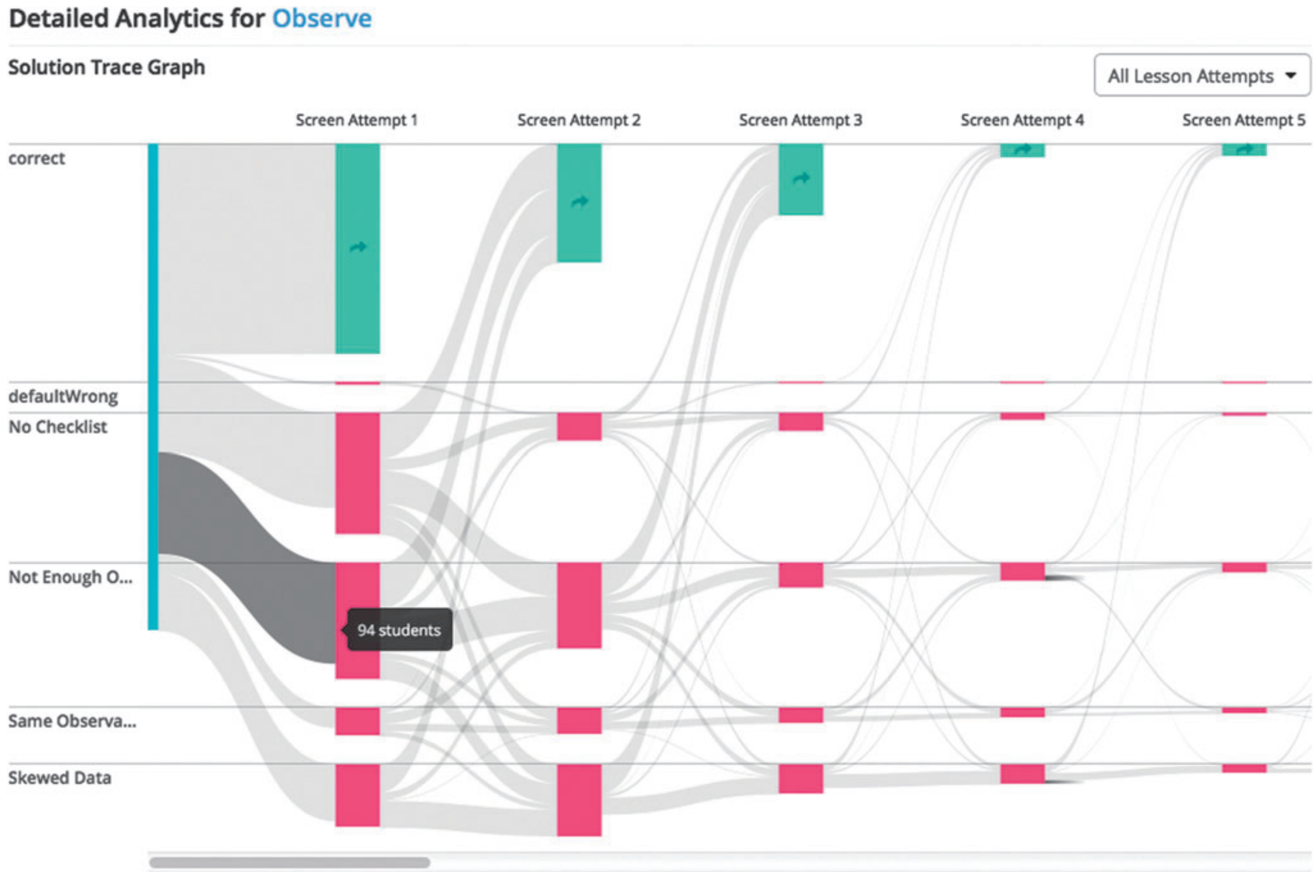


FIG. 4. Sample analytics from an experiment activity in *Habitable Worlds*. In this activity, students were required to establish an observation methodology, then execute it. This visualization shows how the class as a whole progressed through the activity, including the trap-states they triggered and the pattern of mistakes they made from attempt to attempt. Many students made no methodological errors and continued on to the next page of the activity (''correct''). A number of students failed to indicate that they had completed their observations (''No Checklist''), did not make the number of observations they indicated they would make when establishing their methodology (''Not Enough O[bservations]''), or did not make a good distribution of observations (''Skewed Data''). Some students attempted to cheat the system by making all their observations in one place, hoping that we were only checking the number of clicks, rather than their locations (''Same Observa[tions]'').

3.2. Adaptive feedback, pathways, and design: Adaptive eLearning Platform

The AeLP allows instructors to create activities that include complex, adaptive behavior with relative ease. From the students' perspective, the course provides instant feedback on their work. For simple tasks, like an opinion survey, a failure to submit an answer to the posed question will trigger a helpful response (*e.g.,* ''Please answer the question.''). For questions with a definitive correct answer, incorrect responses can receive a range of feedback, from general to specific. For example, feedback on a calculation page could trigger a response such as ''One or more of your answers is wrong'' or ''Your answers are off by a factor of ten. Make sure you are using the correct units.'' Initial offerings of *Habitable Worlds* often provided more general responses, like the first one, in part due to a lack of insight into what mistakes students would make. However, after each offering, a careful dissection of submitted answers to particular questions allowed us to quickly and easily add more fine-grained responses, like the second one, for subsequent offerings. For activities where students work with simulators and data sets that have no single correct answer, students are free to approach the analysis however they choose, which triggers feedback if they make methodological errors. Once these problems are rectified, a student proceeds, with feedback that provides summary information or additional information building on what he or she has just completed. Students experience adaptive pathways based on their responses as well. Two students working side-by-side may find that they are completing different work in the same activity based on the specific mistakes that they have made.

Finally, because much of what we teach is numerical, we use equations with randomized variables, question banks with similar-looking questions but different solutions, and the combination of the two to create unique problems sets for each student. Two students working on the same activity find themselves completing the same task but with different numbers and correct solutions. Rather than calculating a solution and sharing with friends, students can only assist by explaining to each other how to determine the solution, which is the skill we want students to learn anyway.

3.3. Observing and modeling: simulations (sims) and immersive virtual field trips

Although the AeLP has powerful scripting abilities, the heart of the course lies in the simulators (Fig. 5). Smart Sparrow has developed several simulators for *Habitable Worlds,* including a stellar nursery, an operational carbonate-silicate cycle, and an electron transport chain with swappable reductants and oxidants. These simulations are embedded in the AeLP so that their properties and states can be accessed by the AeLP. This allows us to directly track each student's interactions with the sims. In most online courses, simulations and activities are decoupled. This often results in activities where students are instructed to manipulate a sim and answer a question, with no way of determining whether the answer they are submitting resulted from their own work with the sim or from looking at a friend's screen or from prior knowledge or a lucky guess. With integration of sims and the AeLP, we can require students to manipulate the sim in a way that demonstrates mastery of concepts. We can also provide adaptive feedback to correct mistakes and use adaptive design to ensure that students begin with different initial conditions resulting in different final solutions, even if the skill they are practicing is the same (*i.e.,* in an activity where students need to find a star's habitable zone, the star's luminosity is randomized, resulting in a different location for the habitable zone for each student). Multiple sims can be strung together in sequence to give students a taste of how scientists use models to understand complex phenomena and make predictions about observations they will make.

In addition, our team has developed a suite of freely available *immersive virtual field trips* (iVFTs, vft.asu.edu; Bruce *et al.,* 2014) that showcase astrobiologically significant locations such as Shark Bay and the Pilbara regions of Australia. An iVFT consists of a set of linked, spherical images that students can navigate and explore. Further details of the virtual field locality can be explored through high-resolution gigapan imagery, videos, photos, and rotatable 3-D objects. These iVFTs are fully integrated into the AeLP platform, allowing us to script an inquiry-based learning experience at a virtual field site, observe what students are doing, and create appropriate responses. Combined with the ability to ask questions, show videos, and even overlay sims, an iVFT in the AeLP becomes a powerful experience that not only allows students to observe and comment on their surroundings but also to build and manipulate reconstructions, feed these reconstructions into other sims, and use results from their sims to make predictions about what a deeper investigation of the field site may reveal.

3.4. Instructor support and community building: Piazza and Adobe Connect

When we piloted *Habitable Worlds* in the fall of 2011, we realized that the presence of bugs and unoptimized lesson design could potentially constitute significant stumbling blocks for students and success of the project. In an attempt to mitigate these problems, we made considerable use of the discussion board within the LMS, which proved helpful to the students but cumbersome to navigate. Subsequent to that initial course offering, we switched to the freely available Piazza (www.piazza.com) system as a discussion board, owing to its superior functionality. Piazza allows students and instructors to construct collaborative answers and responses to questions with embedded screenshots, videos, and equations. This makes troubleshooting and instruction much easier. All posts are searchable and accessible via smartphone apps, which allows more flexibility in how and when students and instructors interact. As reported in end-of-term surveys, interactions on Piazza are of significant value to our weaker students and provide an opportunity for stronger students to gain greater confidence by explaining their own solutions.

Our prominent use of Piazza also highlights the evolving role of instructors in online courses. In an in-person course, much of an instructor's day-to-day work is spent preparing for each class, attending class to lecture or supervise activities, and grading. In an online course, this work is done in advance or automatically by the course software. As a result, the time saved on repetitive course infrastructure can be invested in interacting with students. With faster and more in-depth responses from course staff to student difficulties and commentary, students gain a richer experience.

As instructors, we often miss the personal interaction with students. To mitigate this problem, in certain semesters we incorporate Adobe Connect (www.adobe.com) into the course. When desired, we conduct once-weekly group meetings with interested students to talk more about the philosophy of the class, the weekly activities, astrobiology in the news, and even bring in guest scientists to interact with the students. The class often overlaps with major scientific conferences attended by one or more of the authors (*i.e.,* American Geophysical Union, Astrobiology Science Conference, Lunar and Planetary Science Conference), so we can link *Habitable Worlds* students directly into the conference, allowing them to converse with a number of scientists, many of whom have results that are making headlines in the news. Typically we chat live via video, and students interact with us via text chat. About 5% of the students in a given class attend these optional interactions, and they return week after week. This is excellent value-added content for our more advanced students.

4. Course Outcomes and Evaluation

Habitable Worlds has been offered 12 times between 2011 and 2017, to more than 3000 students. Data analysis here will focus on the Fall 2014, Spring 2015, Fall 2015, and Spring 2016 offerings, after the majority of the course development was complete and offerings became directly comparable to each other. This work was conducted in accordance with an approved institutional review board protocol to protect the rights of human subjects.

4.1. Learning outcomes

4.1.1. Final grades. Final grades are a high-level overview of student performance in the course (Table 3). During the four semesters studied, 50% of students earned A grades while 28% of students earned a D or E or withdrew from the course (note that a ''withdraw'' is distinct from ''dropping'' a course in terms of when the event takes place during the semester). This distribution is different from other equivalent general education courses at ASU, which tend to have fewer A grades and fewer very low grades. However, grades alone are weak evidence of learning (*e.g.,* Marzano, 2000). Here, we use detailed student data from the course project to

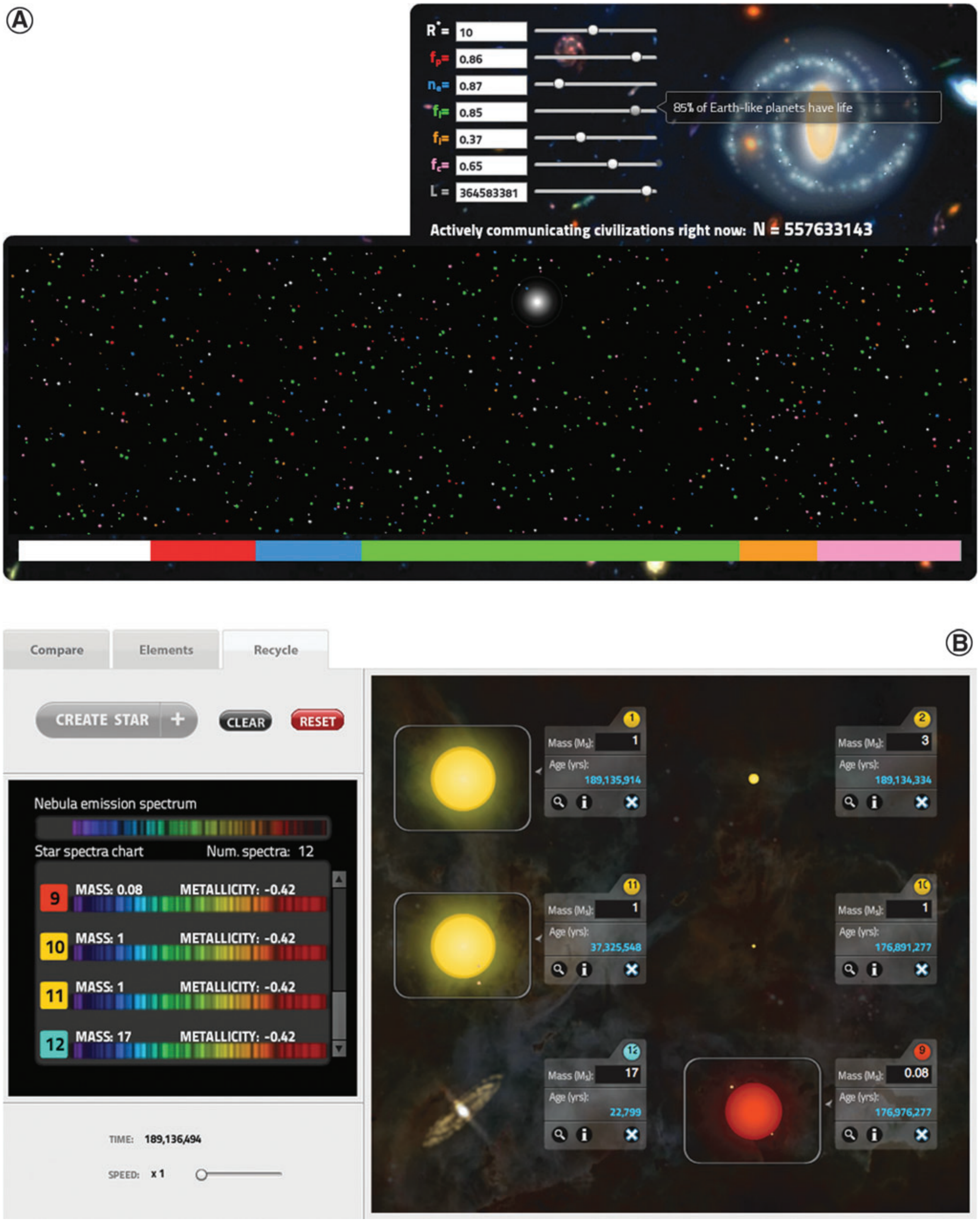
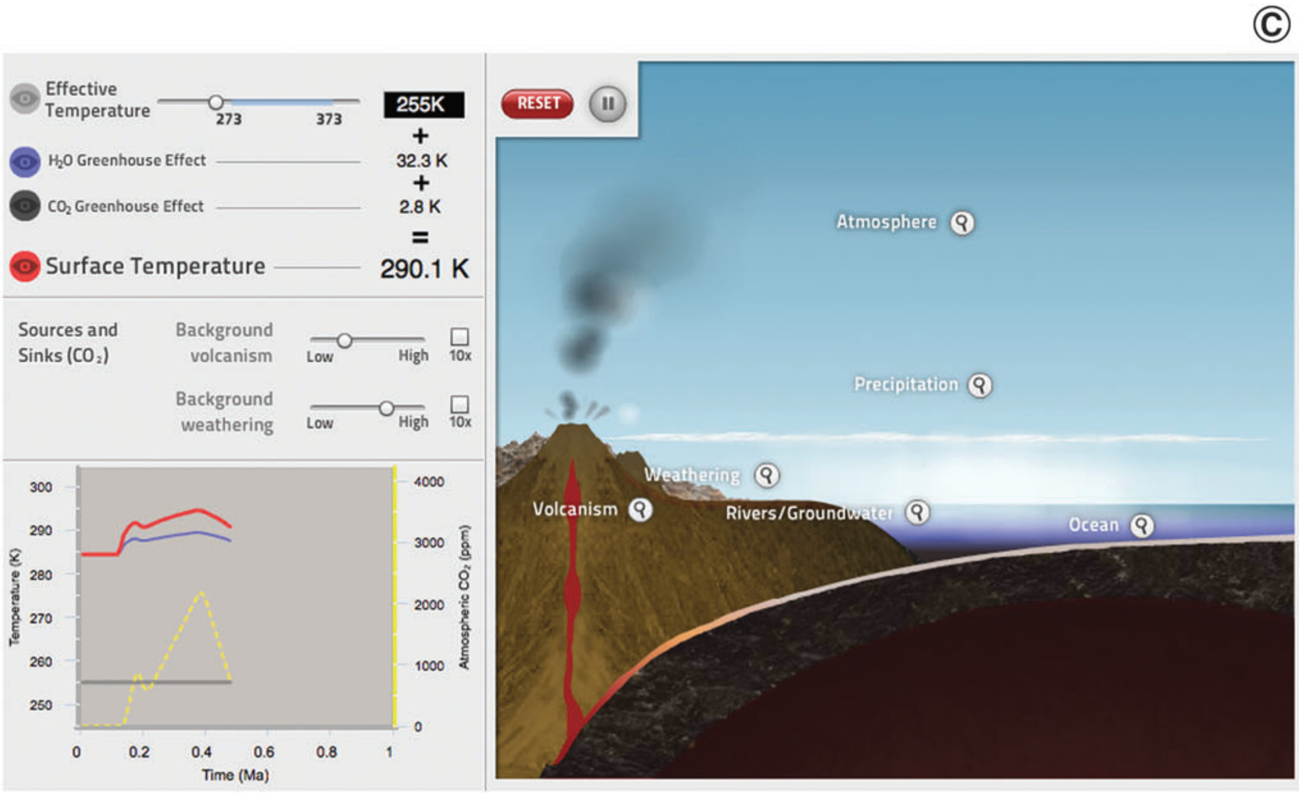


FIG. 5. Sample simulators from *Habitable Worlds,* including (**A**) the Drake Equation, which allows students to manipulate the terms of the equation and observe how the distribution of planets with life changes; (**B**) the Stellar Nursery, in which students use various stars to manipulate the metallicity of the host nebula in order to generate star systems with planets; (**C**) the Inorganic Carbon Cycle, in which students manipulate a planet's effective temperature and background volcanism and weathering rates and can trace the movement of carbon through the atmosphere, geosphere, and hydrosphere; and (**D**) the Redox Sim, in which students can select their reductants and oxidants and observe whether they react to generate energy for the cell.



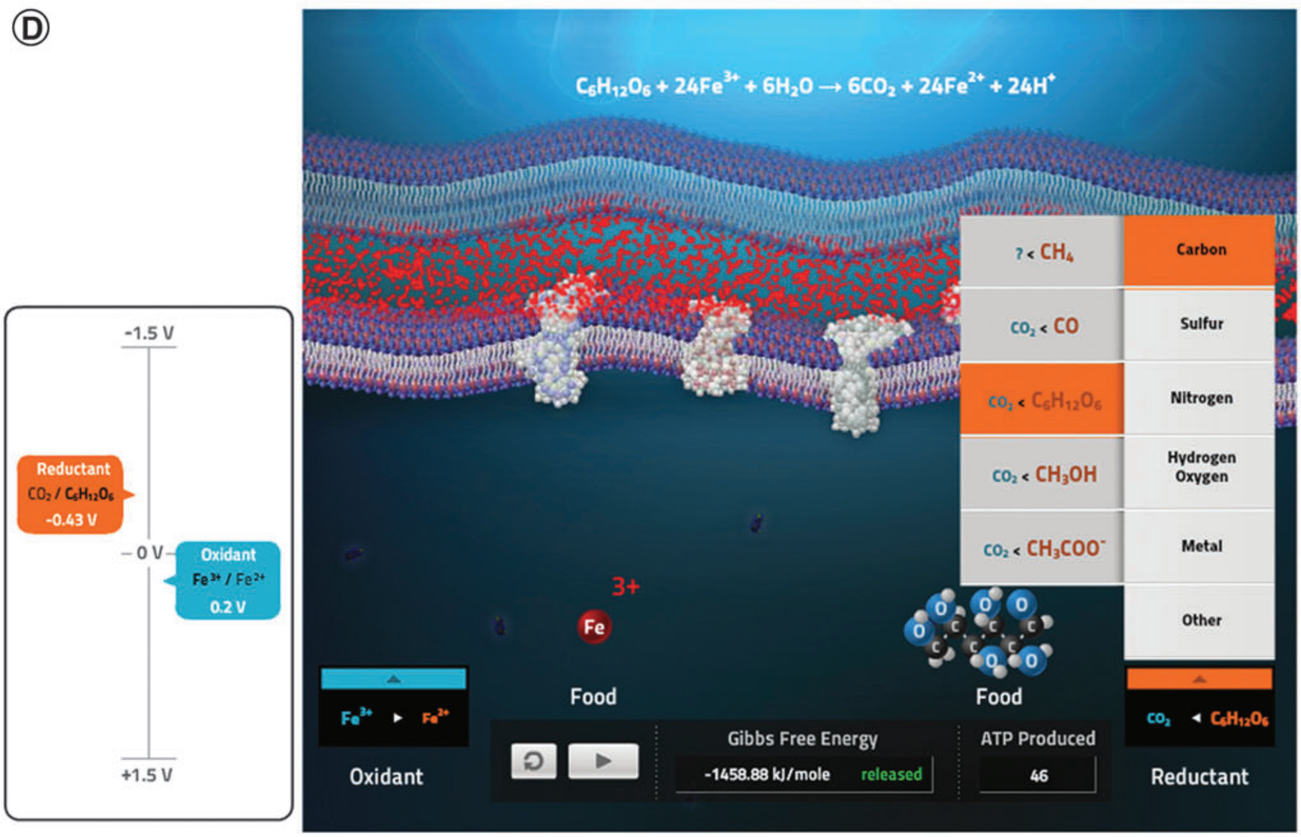


FIG. 5C and 5D. (*continued*).

Table 3. Grades in *Habitable Worlds* (Fall 2014–Spring 2016)

Grade	*All*	*Men*	*Women*	*Online*	*Campus*	*Part-time*	*Full-time*	*White*	*Hispanic/Latino*	*Black/African American*	*Asian*
A	535 (50%)	279 (58%)	256 (43%)	272 (51%)	263 (49%)	158 (46%)	377 (52%)	383 (54%)	77 (39%)	19 (33%)	16 (62%)
B	158 (15%)	69 (14%)	89 (15%)	68 (13%)	90 (17%)	32 (9%)	126 (17%)	106 (15%)	33 (17%)	7 (12%)	2 (8%)
C	86 (8%)	32 (7%)	54 (9%)	39 (7%)	47 (9%)	18 (5%)	68 (9%)	58 (8%)	16 (8%)	5 (9%)	1 (4%)
D or E	180 (17%)	68 (14%)	112 (19%)	103 (19%)	77 (14%)	59 (17%)	121 (17%)	103 (14%)	41 (21%)	17 (30%)	3 (12%)
W	115 (11%)	37 (8%)	78 (13%)	55 (10%)	60 (11%)	76 (22%)	39 (5%)	65 (9%)	28 (14%)	9 (16%)	4 (15%)
Total	1074	485	589	537	537	343	731	715	195	57	26

A indicates excellent performance, B indicates above-average performance, C indicates average performance, D indicates passing performance, E indicates failing performance, and W indicates a withdrawal from the class during the course of the semester. Drops were not included, as the add/drop period is typically one day, and students who dropped the class engaged with little to no course content.

argue that higher grades in *Habitable Worlds* can be considered evidence of learning.

4.1.2. Success and efficiency in completion of the course project.

The project presents an opportunity for students to demonstrate their understanding of the essential learning objectives of *Habitable Worlds,* including basic content knowledge, interpreting data, and executing a scientific methodology to solve a problem. In the project, students are required to find a habitable world in a field of 500 stars and do so by completing calculations relevant to the search for a habitable world. Their grade in this project is based principally on two factors: (1) the average accuracy of their calculations and (2) the number of specific star and planet types they were able to identify among their 500 stars. Students who earn a C or better in the class typically earn high grades in the project (average score 92%), demonstrating proficiency in key learning objectives, whereas D and E students generally show little progress in the project (average score 21%).

To encourage students to approach the project systematically, the project provides a funding mechanism that allows students to automate repetitive calculations (such as distance from parallax data, or planetary effective temperature from star luminosity, planetary albedo, and planetary distance). To ‘‘unlock’’ this feature, students must demonstrate competence in a chosen calculation. Students do this by submitting 10 stars and ‘‘paying’’ for a review using their funding. If they pass the review (by correctly calculating the chosen property for each star), then that calculation becomes automated, and they do not need to complete it again during the project. If they fail the review, they lose their money without gaining any benefits. Because the funding is nonreplenishable, students are incentivized to check their work before submitting for review, or they risk running out of funding, which then requires completing the hundreds of required calculations for the project manually. It is important to note that the automated calculations do not reveal correct answers; they simply automate repetitive work. Automated calculations use whatever data a student supplies, whether correct or not. This further incentivizes a careful approach to the project, as partial automation can lead to erroneous conclusions as a result of error propagation (see Fig. 6). Students receive enough funding to allow for up to three errors on each automatable field (or dozens of errors on one or two fields if a student chooses to spend their funding that way). Students can also use this funding to pay for progress checks, which update their overall score but do not identify specific mistakes.

The amount of funding used in the project provides more detailed information than grades alone in terms of how well a student understands the course material. We found that students who earn high grades are more efficient with their spending (Fig. 7). Students who earned A or A+ grades in the course typically used close to the minimum amount of funding necessary to automate all allowable fields, indicating that they submitted a correct set of calculations on their first attempt. B and C students generally spent at least the minimum but were equally likely to spend all their money, indicating that they made multiple errors in calculation. Students who received a D or E grade tended to spend very little of their funding. This indicates that low-achieving

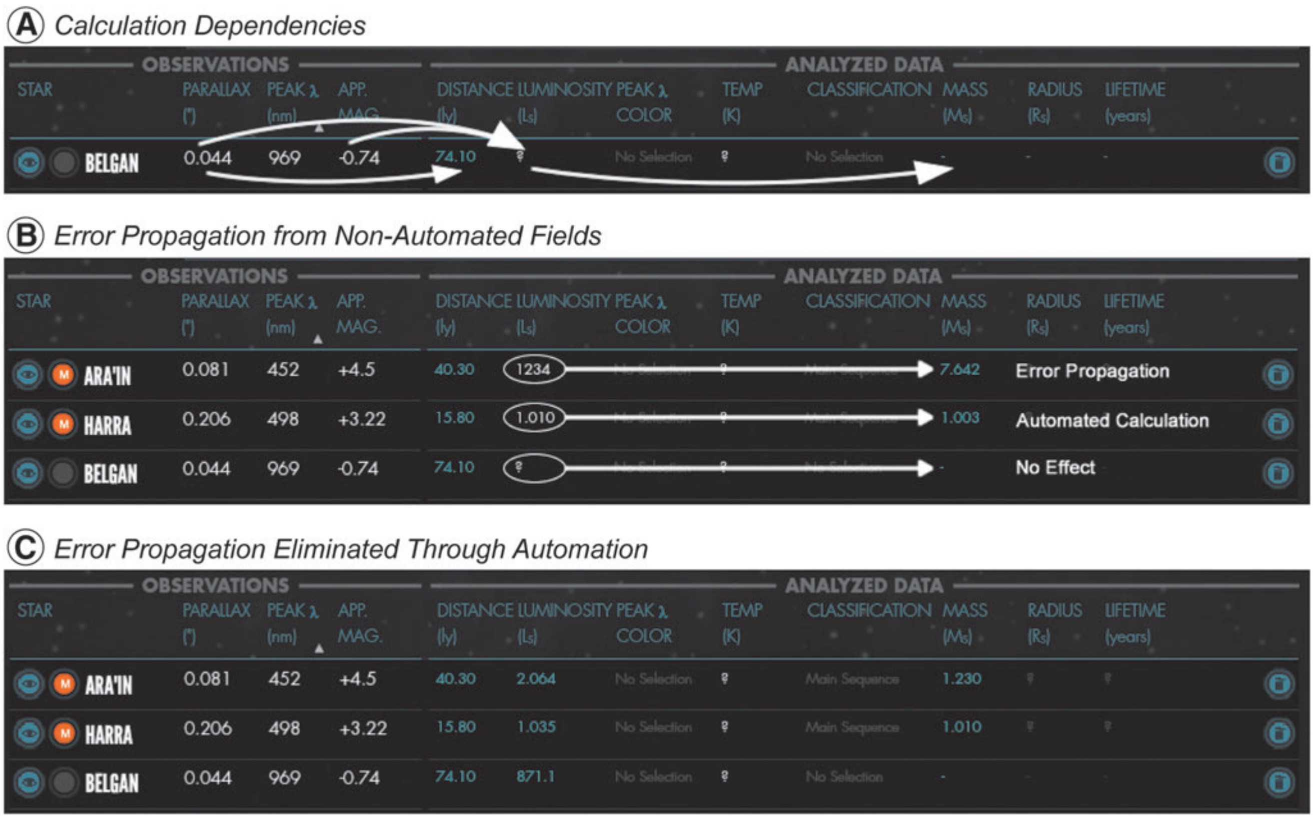


FIG. 6. Calculations can be automated in the project if a student shows competence in completing those calculations. Calculations depend on other data, as indicated by the arrows in (**A**), which allows for error propagation. This feature incentivizes students to automate columns sequentially. In the first example (**B**), the student has already automated distance and star mass but has failed to automate luminosity. As a result, the student must manually compute the luminosity. Any errors, such as for the star Ara'in, are amplified by the automation, which makes subsequent calculations, such as star radius and star lifetime, more difficult to complete correctly. In the second example (**C**), the student has automated the columns in sequence and eliminated error propagation.

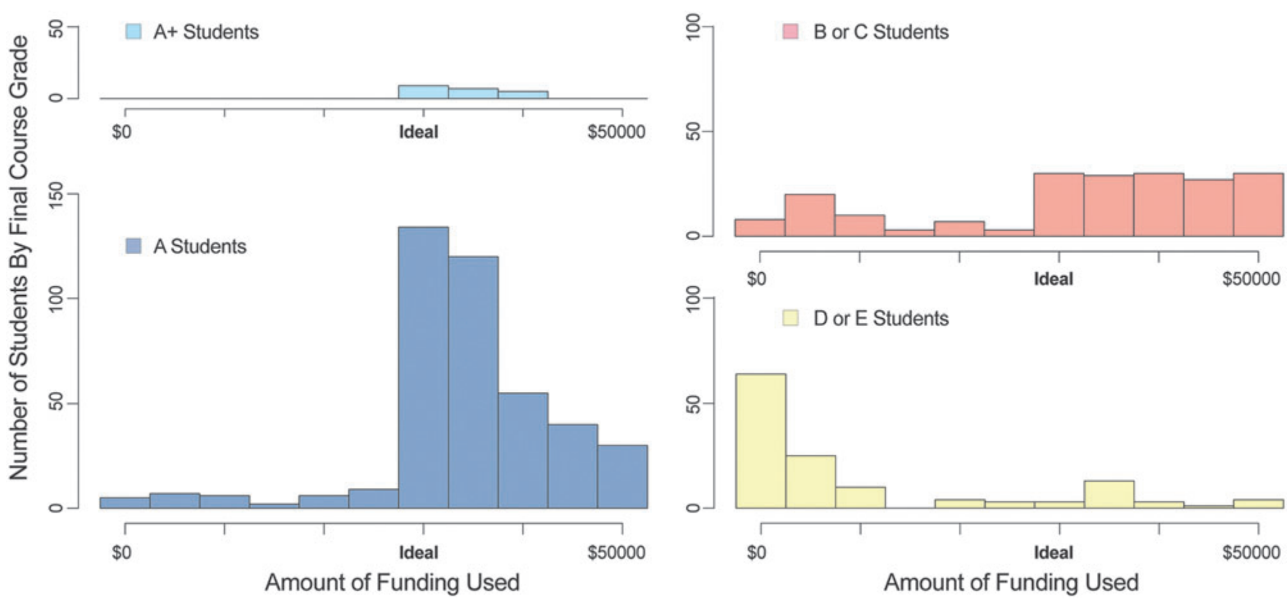


FIG. 7. Comparison of total virtual dollars spent by students separated by final course grade. The minimum cost to achieve all automation unlocks and to check overall score once is $27,825, marked as **ideal**. A and A+ students typically spent close to this ideal number, while B and C students were equally likely to spend all their funding. D and E students were more likely to spend little or nothing. This demonstrates the relative efficiency of higher-performing students on this task as compared to their lower-performing classmates.

students often did not engage with the project, and when they did, did not take advantage of tools to streamline and reduce their workload. These patterns are statistically distinguishable.

The project requires students to apply the skills they have learned throughout the course. Thus, any student who completes all the required tasks within the project has shown an acceptable level of proficiency in the course material. The funding data show the gradations of achievement beyond that minimum proficiency. Students spending near the ideal amount (~ $30,000) demonstrate their ability to execute their methodology accurately. Students spending near the maximum funding have a more haphazard approach toward problem solving, which may include trial and error. The significant relationship between grade and project efficiency supports our claim that the high number of A grades indicates a high degree of student learning in the course, rather than easy grading. Further, these kinds of data highlight the potential of *Habitable Worlds* and computer-based courses to yield sophisticated assessments of students even on higher-order skills, such as problem solving or scientific reasoning.

4.2. Student attitudes and feedback

We have collected student feedback in several ways including the standard university post-course surveys, pre- and post-course surveys using items from the CURE survey (Classroom Undergraduate Research Experience, Denofrio *et al.,* 2007; Lopatto *et al.,* 2008), and two targeted approaches: a short survey for students who had dropped or withdrawn from the course and interviews with students who reached out to the instructors with positive comments, both implemented and analyzed by an external evaluator (coauthor Buxner). Overall, students have responded positively to the course.

In the post-course survey, students were asked to reflect on the value of their educational experience. All values are on a 5-point Likert scale ranging from *strongly disagree* (1) to *strongly agree* (5). They rated their agreement with the statement ''This course was a good way of learning about the subject matter'' at 4.4 ($n = 818$) compared to a national benchmark of 4.0. Similarly, they rated their agreement with the statement ''This course had a positive effect on my interest in science'' at 4.2 ($n = 816$) compared to a benchmark of 3.7. Comparing the pre- to post-course surveys, we observed an increase from 3.8 to 3.9 ($p < .001$), with a benchmark score of 4.0, in students agreeing with the statement: ''I can do well in science courses.'' We also saw an increase from 4.1 to 4.2 ($p = .06$), with a benchmark score of 4.1, in agreement with the statement ''Even if I forget the facts, I'll still be able to use the thinking skills I learn in science.'' The benchmark data are courtesy of the survey developer (Lopatto, personal communication) and were collected over a period of one year, with students completing the post-course survey at the end of their academic term. The 8960 students supplying data represented 101 institutions of higher learning, most frequently from large research universities and small liberal arts colleges. The median number of student responses from a course was 28. Student participants were predominantly female (62%) and white (61%). Because our survey results are above the benchmark values and because attitudes toward learning frequently decline during science courses (Redish *et al.,* 1998; McConnell and van der Hoeven Kraft, 2010), these results indicate that students perceive *Habitable Worlds* to be a positive experience.

Of the 54 respondents who completed a follow-up survey after dropping or withdrawing from the course, 14 indicated that they had since completed the course or planned to take the course in the future. The other 40 respondents indicated that they did not complete the course nor did they plan to take the course in the future. Students' reasons for dropping the course included the following: the course took too much time (15); they did not need the course for their major (12); they were not interested in the course topics (12); they did not like the delivery of content or format of the course (8); they did not feel they had sufficient math knowledge to complete the course (8); they were getting a lower grade than they wanted (7); they had financial barriers to paying for the course (5); they were overloaded with courses (5); they felt the course was too difficult (3); and they had left school (2).

During the winter of 2015, eight students who had previously provided incredibly positive feedback about the course were interviewed about their experiences in the course. All these students reported overall positive experiences despite encountering barriers with content or technology. These students highlighted the importance of instructor support, interaction with other students via Piazza, and applying the concepts that they were learning to later parts of the course as essential to their success and enjoyment of the course.

4.3. Shortcomings

Habitable Worlds engages a wide swath of students, who frequently highlight that the format and the level of instructor interaction have made learning the concepts easier and more engaging. Although most students feel that the material can be difficult and frustrating at times, most also feel that they can easily receive the help they need to successfully complete the course. The course grades highlight that success is attainable for the majority of our students.

Nevertheless, there are some noticeable shortcomings. Analysis of the grades by demographics shows that not all groups of students perform equally well in *Habitable Worlds* (Table 3). There is a statistically significant difference between the performance of men and women in the class ($p < .001$), with men earning higher grades than women. There are also statistically significant differences between ethnicities, with White and Asian students earning higher grades and Hispanic and African American students earning lower ones ($p = .004$). Students enrolled in on-campus degree programs earn marginally higher grades ($p = .09$). There is also a statistically significant difference between part-time (<12 credit hours) and full-time (12 or more credit hours) students, with full-time students performing better ($p < .001$). Data from other courses at our university that fulfill the same general education requirement as *Habitable Worlds* show similar grade differences. A regression analysis indicates that much of this apparent demographic bias can be attributed to differences in overall GPA between demographic groups. This is consistent with prior work showing that preparedness is the underlying driver of achievement and persistence gaps (*e.g.,* Roderick *et al.,* 2009). That said, previous studies have noted grade gaps of around half a letter grade for women and minorities in introductory-level college science (*e.g.,* Miyake *et al.,* 2010; Haak *et al.,* 2011), so we consider this to be an area where *Habitable Worlds* can be improved.

5. Applications Beyond *Habitable Worlds*

Online education is fundamentally different from traditional in-person instruction. What works for synchronous, in-person instruction may not be appropriate for asynchronous, online instruction. Therefore, we believe the best outcomes require new tools and new methods and cannot be mere translations of existing in-person courses to the online realm. In this work, we have described the process of designing and developing an online astrobiology lab course informed by ideas in active and inquiry-based learning. The evaluation data show that students have responded positively to the course and its structure and that they are generally successful in the course's summative project.

Our experiences and decisions are thus specific to our institutional setting. However, the goal of providing active, inquiry-based learning online and the pedagogical and technological choices inherent to that goal are common to many institutions. Thus, our results will prove valuable to other researchers and practitioners building their own digital learning experiences, whether fully online or as supplements to existing in-person classes. For all these potential designers and educators, we suggest that a 21st-century online course should

- Be built in an adaptive and responsive platform that provides students with immediate feedback and adaptive pathways to address persistent misconceptions;
- Include rich and interactive media content (like simulators and virtual field trips) that interfaces with an ITS to engage students and allow manipulation of concepts for greater depth of understanding;
- Include an overarching ''big question'' project to test student comprehension of course topics, critical thinking skills, and ability to apply course concepts; and
- Include a robust discussion or interaction platform for student-student and student-instructor interactions.

A key aspect of our approach is in transforming the role of the instructor in an online classroom. By automating many conventional responsibilities of the instructor (lecturing and grading), we have opened our time to interacting more meaningfully with students. In a traditional video-and-quiz online course, an instructor may not be necessary. However, in a media-rich virtual environment with simulators, virtual field trips, and complex projects, an expert's guidance becomes critical to student success, as our own experience has shown and as our students have reported on end-of-term evaluations.

Using an ITS provides considerable benefit for instructors, both in initial development and ongoing refinement. Because *Habitable Worlds* has undergone 5 years of design and development, the end results look overwhelming to new instructors who are interested in developing their own content. However, the majority of *Habitable Worlds* content began as worksheets that were simply translated into the online environment. Student interactions and reactions guided the development and transformation of this content over the years. The flexibility of an ITS allows the development of content that integrates video, simulators, feedback, and images into combinations that are often not possible in the rigid formatting required by most LMSs.

Finally, and most importantly, the detailed records of student actions, answers to specific questions, and time spent on tasks allow for rapid and effective iteration from semester to semester, a process that increases the depth and granularity of feedback over time. In traditional courses, instructors may be in the dark about what went wrong when students fail an exam. With deep analytics captured throughout a course developed in an ITS, instructors can dissect their class in considerable detail and improve existing lessons over time by adding new pathways, new feedback, remediation content, and bonus pathways.

Use of any or all of these tools and techniques can help improve course offerings, both online and in-person. Through our partner institutions (currently at 22), elements and sometimes the entirety of *Habitable Worlds* have been incorporated into in-person courses as supplementary learning modules, homework assignments, or lab replacements. We are currently applying the lessons we have learned from the development of *Habitable Worlds* to continue improving the ASU offering as well as to compare student results from the core offering to those in our partner institutions. At the same time, we continue to develop analytical methods to study how and why students struggle with the course, so that we can fully deliver on the promises of online education in the 21st century.

Acknowledgments

Funding for development of *Habitable Worlds* was provided by the NASA Astrobiology Institute, National Science Foundation (TUES grant number #1225741), and ASU Online. We would like to thank the instructors, teaching assistants, and developers who assisted greatly in the development and refinement of course content and running the course over the past six years: Matthew Kellom, Marc Neveu, David Copeman, Shaowei Ho, Jason Raymond, Svetlana Shkolyar, Nikhil Monga, Anusha Kalyaan, Kirt Robinson, Adam Monroe, Jason Lai, Bryce Carande, Aleisha Johnson, and Audrey Horne.

Author Disclosure Statement

Coauthor Belinson is a cofounder of Smart Sparrow. Author Horodyskyj and coauthors Mead, Buxner, Semken, and Anbar have no competing financial interests.

References

Bandura, A. (1982) Self-efficacy mechanism in human agency. *Am Psychol* 37:122–147.

Ben-Naim, D. (2011) A software architecture that promotes pedagogical ownership in intelligent tutoring systems. Doctoral dissertation, University of New South Wales, Sydney, Australia.

Bruce, G., Anbar, A., Semken, S., Summons, R., Oliver, C., and Buxner, S. (2014) iVFTs: immersive virtual field trips for interactive learning about Earth's environment [abstract ED52A-02]. In *AGU Fall Meeting 2014,* American Geophysical Union, Washington, DC.

Bybee, R.W., Taylor, J.A., Gardner, A., Van Scotter, P., Carlson Powell, J., Westbrook, A., and Landes, N. (2006) *The BSCS 5E Instructional Model: Origins, Effectiveness and Applications,* BSCS—Biological Sciences Curriculum Study, Colorado Springs, CO. Available online at http://www.bscs.org/bscs-5e-instructional-model

Denofrio, L.A., Russell, B., Lopatto, D., and Lu, Y. (2007) Linking student interests to science curricula. *Science* 318:1872–1873.

Deslauriers, L., Schelew, E., and Wieman, C. (2011) Improved learning in a large-enrollment physics class. *Science* 332:862–864.

Drake, F.D. (1961) Project Ozma. *Phys Today* 14:40–46.

Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., and Wenderoth, M.P. (2014) Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci USA* 111:8410–8415.

Friedman, T.L. (2012, May 15) Come the revolution. *New York Times.* Available online at http://www.nytimes.com/2012/05/16/opinion/friedman-come-the-revolution.html

Haak, D.C., HilleRisLambers, J., Pitre, E., and Freeman, S. (2011) Increased structure and active learning reduce the achievement gap in introductory biology. *Science* 332:1213–1216.

Hake, R.R. (1998) Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am J Phys* 66:64–74.

Hobson, A. (2008) The surprising effectiveness of college scientific literacy courses. *The Physics Teacher* 46:404–406.

Hofstein, A. and Lunetta, V.N. (2004) The laboratory in science education: foundations for the twenty-first century. *Sci Educ* 88:28–54.

Holmes, N.G., Wieman, C.E., and Bonn, D.A. (2015) Teaching critical thinking. *Proc Natl Acad Sci USA* 112:11199–11204.

Jordan, K. (2014) Initial trends in enrollment and completion of massive open online courses. *The International Review of Research in Open and Distributed Learning* 15:134–160.

Lawson, A.E., Abraham, M.R., and Renner, J.W. (1989) *A Theory of Instruction: Using the Learning Cycle to Teach Science Concepts and Thinking Skills,* National Association for Research in Science Teaching, Manhattan, KS.

Lopatto, D., Alvarez, C., Barnard, D., Chandrasekaran, C., Chung, H.-M., Du, C., Eckdahl, T., Goodman, A.L., Hauser, C., Jones, C.J., Kopp, O.R., Kuleck, G.A., McNeil, G., Morris, R., Myka, J.L., Nagengast, A., Overvoorde, P.J., Poet, J.L., Reed, K., Regisford, G., Revie, D., Rosenwald, A., Saville, K., Shaw, M., Skuse, G.R., Smith, C., Smith, M., Spratt, M., Stamm, J., Thompson, J.S., Wilson, B.A., Witkowski, C., Youngblom, J., Leung, W., Shaffer, C.D., Buhler, J., Mardis, E., and Elgin, S.C.R. (2008) Genomics education partnership. *Science* 322:684–685.

Lorenzo, M., Crouch, C.H., and Mazur, E. (2006) Reducing the gender gap in the physics classroom. *Am J Phys* 74:118–122.

Marzano, R.J. (2000) *Transforming Classroom Grading,* Association for Supervision and Curriculum Development, Alexandria, VA.

McConnell, D. and van der Hoeven Kraft, K. (2010) GARNET: Geoscience Affective Research NETwork; what we've learned & where we're going. Presented at the AAAS/NSF Meeting on Transforming Undergraduate Education in Science, January 26–28, Washington, DC.

Miyake, A., Kost-Smith, L.E., Finkelstein, N.D., Pollock, S.J., Cohen, G.L., and Ito, T.A. (2010) Reducing the gender achievement gap in college science: a classroom study of values affirmation. *Science* 330:1234–1237.

Redish, E.F., Saul, J.M., and Steinberg, R.N. (1998) Student expectations in introductory physics. *Am J Phys* 66:212–224.

Robbins, S.B., Lauver, K., Le, H., Davis, D., Langley, R., and Carlstrom, A. (2004) Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychol Bull* 130:261–288.

Roderick, M., Nagaoka, J., and Coca, V. (2009) College readiness for all: the challenge for urban high schools. *Future Child* 19:185–210.

Ruiz-Primo, M.A., Briggs, D., Iverson, H., Talbot, R., and Shepard, L.A. (2011) Impact of undergraduate science course innovations on learning. *Science* 331:1269–1270.

Schwartz, R.S., Lederman, N.G., and Crawford, B.A. (2004) Developing views of nature of science in an authentic context: an explicit approach to bridging the gap between nature of science and scientific inquiry. *Sci Educ* 88:610–645.

Songer, N.B. and Linn, M.C. (1991) How do students' views of science influence knowledge integration? *J Res Sci Teach* 28:761–784.

Springer, L., Stanne, M.E., and Donovan, S.S. (1999) Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. *Rev Educ Res* 69:21–51.

Toven-Lindsey, B., Rhoads, R.A., and Lozano, J.B. (2015) Virtually unlimited classrooms: pedagogical practices in massive open online courses. *Internet High Educ* 24:1–12.

Waldrop, M.M. (2015) Why we are teaching science wrong, and how to make it right. *Nature* 523:272–274.

Wieman, C.E. (2014) Large-scale comparison of science teaching methods sends clear message. *Proc Natl Acad Sci USA* 111:8319–8320.

Address correspondence to:
Lev B. Horodyskyj
School of Earth and Space Exploration
Arizona State University
781 S Terrace Road
Tempe, AZ 85281

E-mail: LevH@asu.edu

Submitted 17 June 2016
Accepted 27 April 2017

Abbreviations Used

AeLP = Adaptive eLearning Platform
ASU = Arizona State University
ITS = intelligent tutoring system
iVFTs = immersive virtual field trips
LMSs = learning management systems
MOOCs = massively open online courses
sims = simulations