

# 7.00x Introduction to Biology: The Secret of Life MITx on edX – 2013 Spring



“7.00x lets you explore the secret of life by guiding you through the basics of biochemistry, genetics, molecular biology, recombinant DNA, genomics, and rational medicine.” – edx.org

**Introduction.** 7.00x is an introductory level biology course hosted by professor Eric Lander; one of the leaders of the Human Genome Project. The course content reflects the topics taught in introductory biology courses at MIT and many biology courses across the world. Students first focus on the structure and function of macromolecules such as DNA, RNA and proteins. Students discover the association between structure and function of the macromolecules and implications for human health. As students continue in the course, they apply an understanding of heredity and information flow within cells to human health and disease, learning about techniques in molecular biology and their potential impact in changing the world.

7.00x was offered as a MITx course on edX in the Spring of 2013. Joining Professor Lander on the teaching team were Professor Graham Walker (MIT) and Professor Brian White (UMASS Boston), Dr. Michelle Mischke, and Dr. Mary Ellen Wiltrout.

Many participants in MITx on edX courses come with different intentions and levels of interest and many resemble shoppers and auditors in their course activity. This report identifies four subpopulations of participants (numbers below): those who registered, viewed some course content, explored the course, and were certified<sup>1</sup>.

## Inside

1. Introduction
2. Course structure
3. Demographics
4. Participants
5. Activity
6. Persistence
7. Special Features

Important Dates	
January 30, 2013	Registration Open
March 5, 2013	Course Launch
June 6, 2013	Course Close
September 8, 2013	Date of Report Data Collection

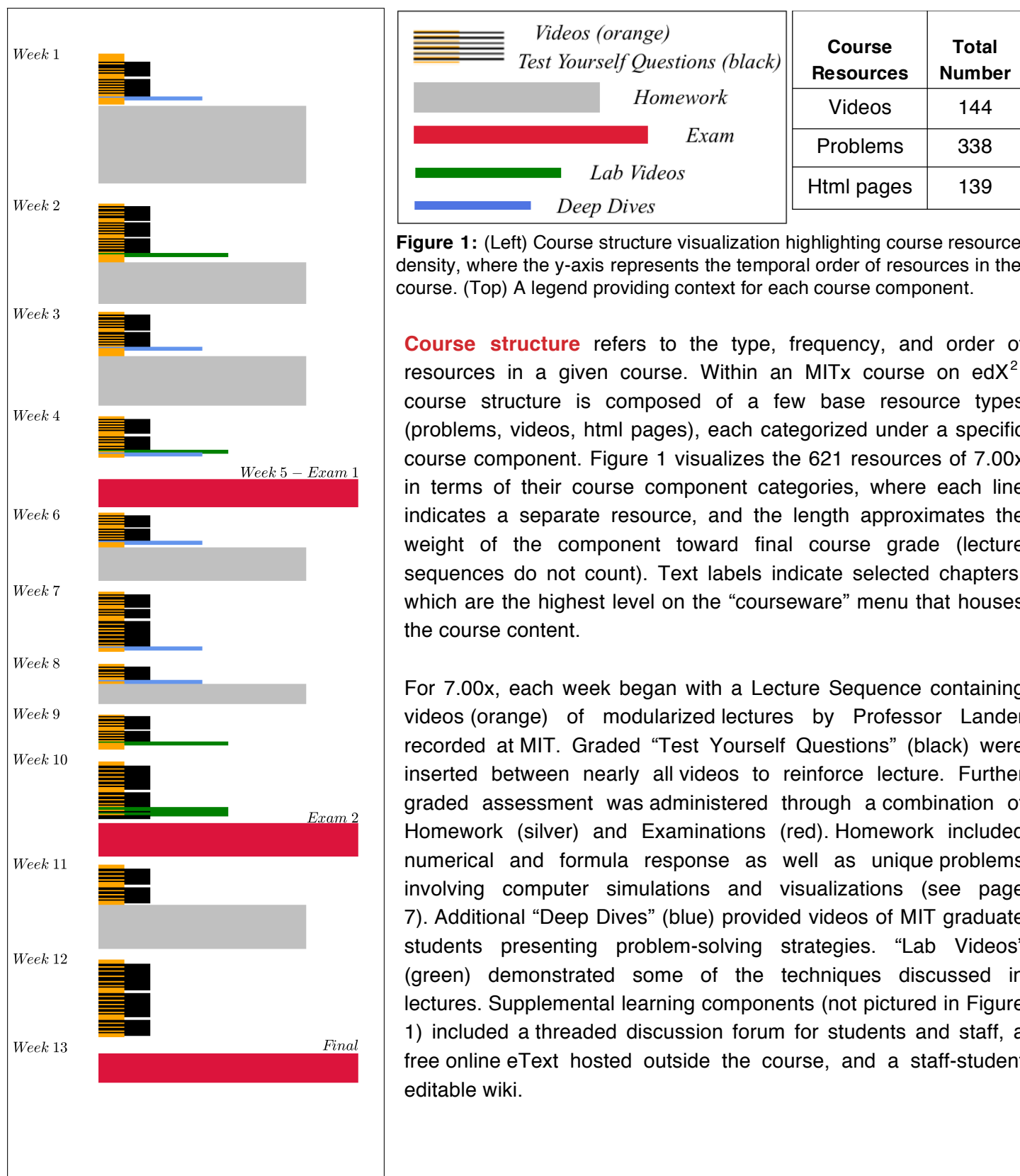
Participants	
Registered	37,977
Viewed	26,592
Explored	5,432
Certified	3,235

<sup>1</sup> See page 4 for definitions

This report was produced jointly by the MIT Office of Digital Learning, and the HarvardX Research Committee.

Seaton, D. T., Reich, J., Nesterko, S, Mullaney, T., Waldo, J., Ho, A., & Chuang, I. (2014). 7.00x Introduction to Biology: The Secret of Life - Spring 2013 MITx Course Report (MITx Working Paper #9)

Electronic copy available at: <http://ssrn.com/abstract=2382325>



**Figure 1:** (Left) Course structure visualization highlighting course resource density, where the y-axis represents the temporal order of resources in the course. (Top) A legend providing context for each course component.

**Course structure** refers to the type, frequency, and order of resources in a given course. Within an MITx course on edX<sup>2</sup>, course structure is composed of a few base resource types (problems, videos, html pages), each categorized under a specific course component. Figure 1 visualizes the 621 resources of 7.00x in terms of their course component categories, where each line indicates a separate resource, and the length approximates the weight of the component toward final course grade (lecture sequences do not count). Text labels indicate selected chapters, which are the highest level on the “courseware” menu that houses the course content.

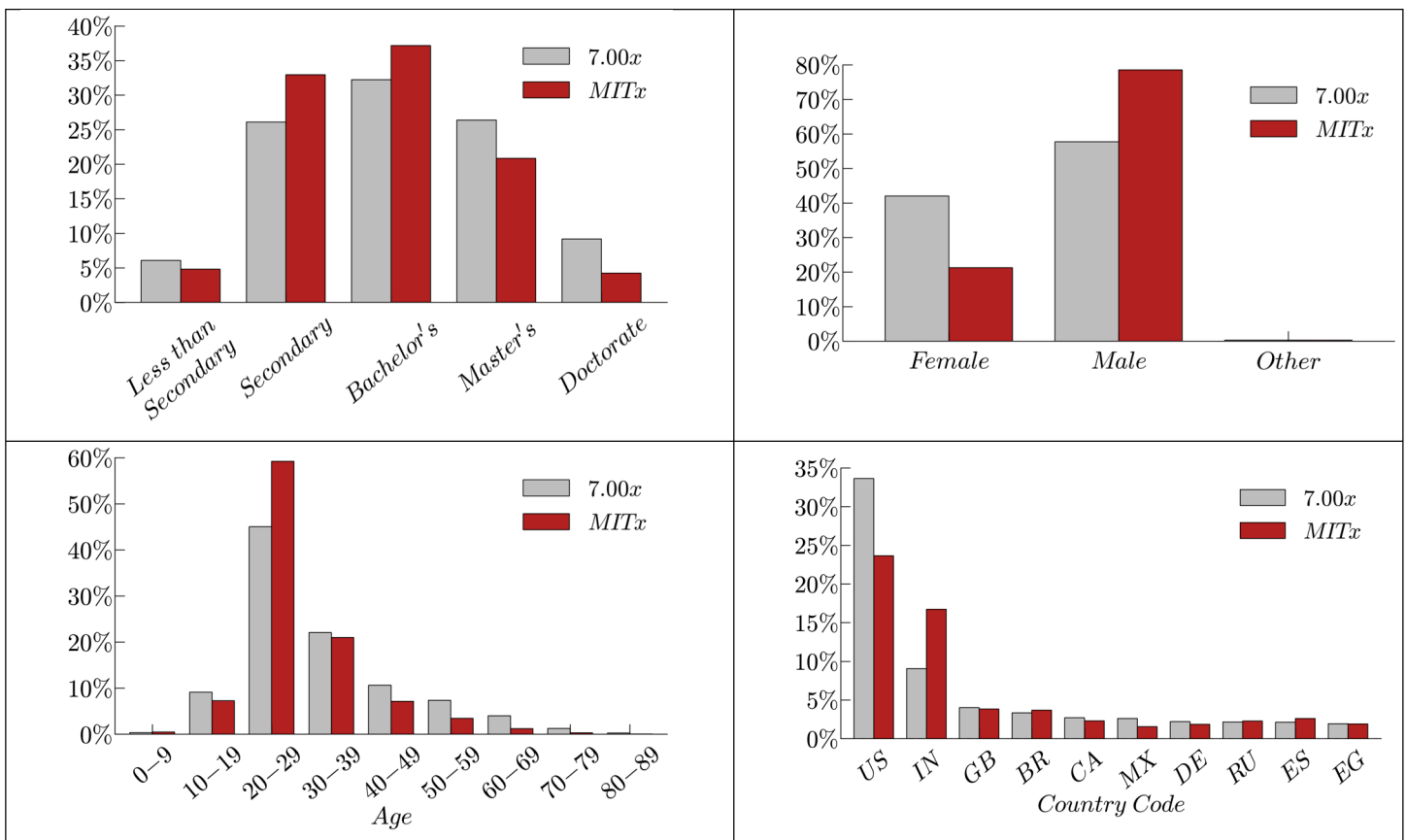
For 7.00x, each week began with a Lecture Sequence containing videos (orange) of modularized lectures by Professor Lander recorded at MIT. Graded “Test Yourself Questions” (black) were inserted between nearly all videos to reinforce lecture. Further graded assessment was administered through a combination of Homework (silver) and Examinations (red). Homework included numerical and formula response as well as unique problems involving computer simulations and visualizations (see page 7). Additional “Deep Dives” (blue) provided videos of MIT graduate students presenting problem-solving strategies. “Lab Videos” (green) demonstrated some of the techniques discussed in lectures. Supplemental learning components (not pictured in Figure 1) included a threaded discussion forum for students and staff, a free online eText hosted outside the course, and a staff-student editable wiki.

<sup>2</sup> Often colloquially referred to as a Massive Open Online Course (MOOC)

The weekly release of chapters in 7.00x occurred over 13 calendar weeks. Final course grades were determined by six homework sets (20%), two midterms (50%), and a final exam (30%). Certification was granted to students whose final grades were 58% or greater.

To date, MITx courses on edX have had similar schedules of regularly released resources and periodic due dates. Like 7.00x, courses have used quantitative assessment supported by lecture sequences and textual material, as well as a few special learning components (see end of this report). In the figures that follow, we compare 7.00x data with those from other MITx courses on edX, from the Fall of 2012 to the Summer of 2013.

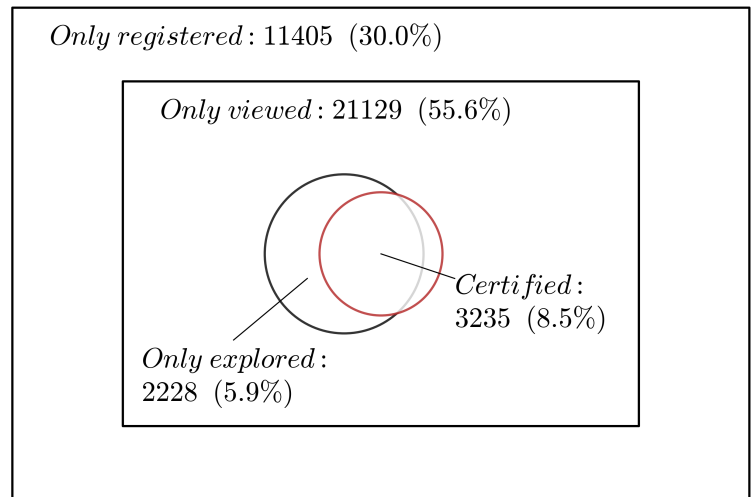
**Demographics.** For 7.00x, we have records of 37,997 students who registered for the course by the cutoff date of this report, September 8, 2013. According to self-reported demographics, registrants were 57.7% men, and as a group, highly educated: 32.2% have earned bachelor’s degrees, 26.4% have masters or professional degrees, and 9.2% of registrants had earned a doctorate (for certified students, percentages are 28.2%, 35.0%, and 13.7%, respectively). Distributions of level of education and gender are plotted below in the top half of Figure 2, along with the average distributions for all MITx courses<sup>3</sup>. Note, roughly 8% of 7.00x registrants did not self-report one or more of the demographic categories.



**Figure 2:** Distributions of demographic variables collected at registration for all registrants reporting data, with MITx-wide comparisons for reference: (top left) education attainment, (top-right) gender, (bottom left) age, and (bottom right) geolocation via IP look-up.

<sup>3</sup> See also Ho, A. D., Reich, J., Nesterko, S., Seaton, D. T., Mullaney, T., Waldo, J., & Chuang, I. (2014). *HarvardX and MITx: The first year of open online courses.* (HarvardX and MITx Working Paper No. 1).

As with other MITx courses on edX, 7.00x students came from all over the world. Based on geolocation of student IP addresses, 33.6% of registrants came from the United States (this was the highest rate across the first year of MITx courses on edX), but many students also came from over 150 other countries around the world. Consistent with other MITx courses on edX, the median age of registrants and certified students were higher than those of typical MIT undergraduates; namely 28 and 31 years, respectively. Distributions of geolocated country and age are plotted for 7.00x in Figure 2, alongside the average distributions for all MITx courses.



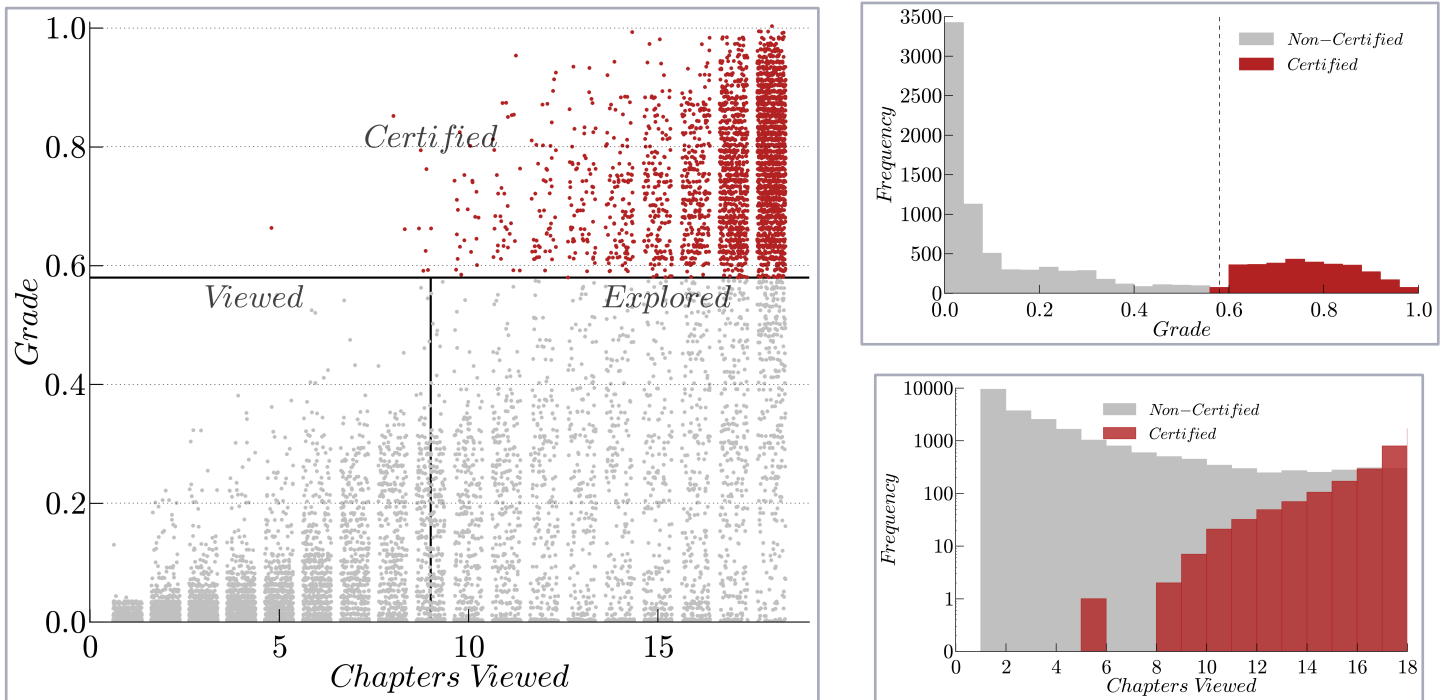
**Figure 3:** Participants separated into four mutually exclusive and exhaustive categories (not to scale).

**Participants in 7.00x.** Who are the students taking this course, and how much did they participate? The degree and kind of participation in the course varied considerably among those who signed up for 7.00x. Furthermore, like many MOOCs, this course remained open to new registrants, so that enrollment continued to rise through the semester. This asynchronicity is a key feature of open online courses. One feature of low barriers to registration, for example, is a large number of students who registered but never actually viewed any course material.

To illustrate the considerable variability in participants by their actions, we identify four subpopulations of interest within this course: those who (1) only registered, and never accessed the courseware, (2) only viewed, i.e. non-certified registrants who accessed at least one chapter but less than half of chapters, (3) only explored, i.e. non-certified registrants who accessed half or more of the chapters, and those who were (4) certified, by earning a final grade at or higher than the cutoff.

Certificate earners are the most conventional subpopulation, defined by crossing a certain threshold of achievement on quantitative assessments (well defined for 7.00x). The “only explored” subpopulation seeks to represent students having interacted with a considerable amount of course material, while the “only viewed” subpopulation highlights participants having only sampled the course. Both the “only explored” and “only viewed” subpopulations contain participants with highly varying levels of activity (as highlighted in the following two sections on persistence and activity). The fourth sub-population of “only registered” contains users that registered, but never accessed the courseware. This is a substantial number in 7.00x and many of the MITx courses on edX.

For 7.00x, the numbers of students who only registered, only viewed, only explored, or were certified, are shown in Fig. 3. The categories are defined as disjoint sets by using the “only” terminology, but the figure highlights the nested nature of registration, viewing, exploring, and certification. Students who were certified

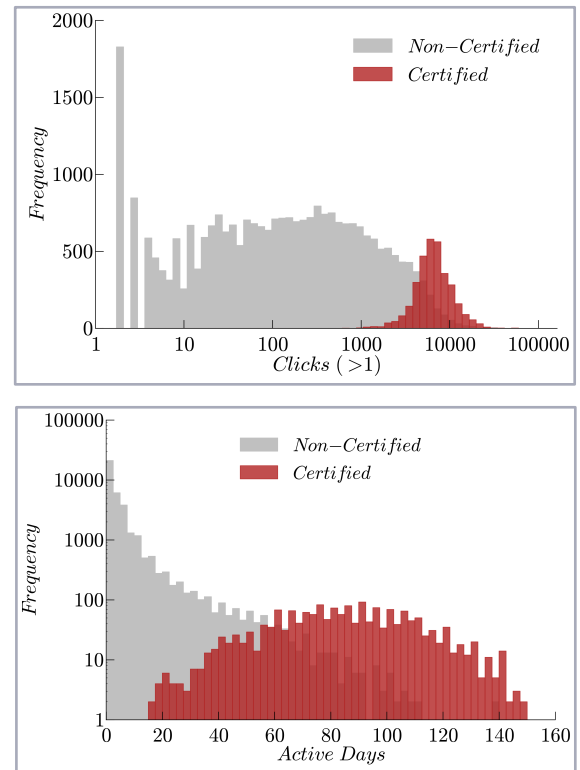


**Figure 4:** Scatter plot of grade versus chapters viewed (left), highlighting student sub-populations; certified students are red points and all points are jittered. Histograms of grades and number of chapters viewed (right) distinguished by student certification status.

certainly viewed the course, but not all viewers were certified. Similarly, “explored” and “certified” are both subsets of viewed.

**Activity in 7.00x.** The activity of students in MITx courses on edX is logged; activity level can provide perspective into which courseware components are of interest to students, and which specific activities students are undertaking in a course, e.g. to earn certification.

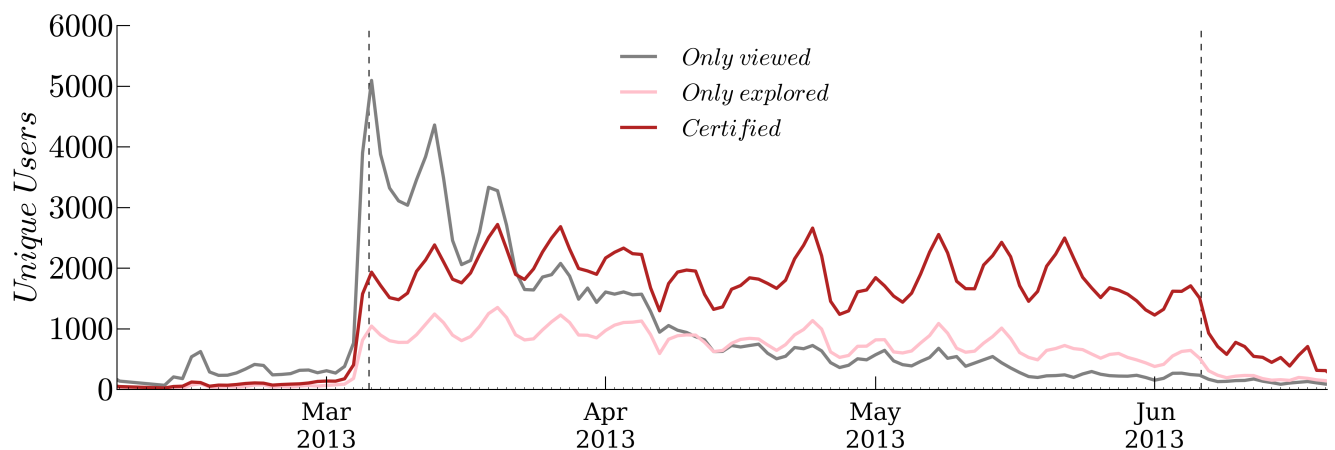
How diverse is the activity of students in the course? Consider the grade earned by students, which is determined in 7.00x by quantitative assessment with regular due dates and periodic exams. A scatter plot (Fig. 4) depicting earned grade versus the number of chapters visited, for each registered student, illustrates the distinctions between the sub-populations of students who only viewed, only explored, or became certified in the course. This plot also illuminates some possible registrant types: “completionists” (perfect grade, viewed all chapters), “optimizers” (occasional



**Figure 5:** Distribution of clicks (total course interactions, above) and active days, (below) distinguished by certificate status.

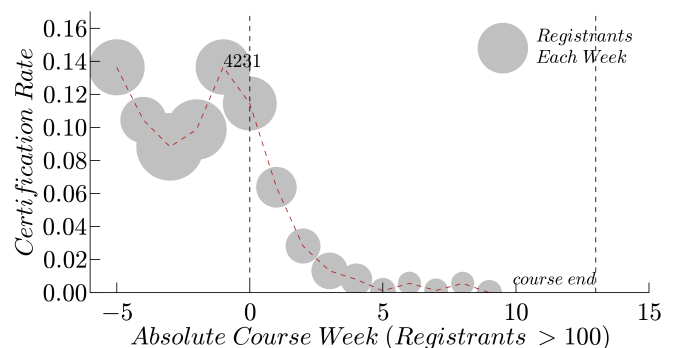
registrants that view minimal chapters but earn a certificate), and “listeners” (viewed all chapters, zero grade – submitted no work for credit).

For 7.00x, edX recorded 42,785,737 events in its tracking logs, relating to students interacting with the course contents, through the cutoff date of this report. These events included interactions indicating when students played videos, attempted problems, browsed through text pages, read or posted entries in the forum, and other activities. Fig. 5 gives a glimpse into these data, presenting a histogram of the total number of events (“clicks”) recorded for all registrants with at least one event, and highlighting the population of students who earned a certificate in comparison with all other students. The figure also presents a histogram of the number of active days, where “active days” represents any day where a course interaction was logged. The distribution of active days for certificate earners is broad, further reflecting the variety of approaches by which a participant can earn a certificate.



**Figure 6:** Daily number of unique participants accessing courseware or course-site (pre-launch) via our described participant categories. Vertical dashed lines indicate the start and end of the course. Due to data loss, some information about user activity between January 21 and Feb 13 is missing and not included in the plot.

**Persistence in 7.00x.** How does the activity of students change as the course progresses, particularly with students who persist to earn certification? Figure 6 highlights courseware and course-site (pre-launch) interactions for three of the disjoint participant types: only viewed, only explored, and certified. In 7.00x, there is a noticeable periodicity in each curve that is highly correlated with content release and due dates. Certified students show a relatively stable weekly periodicity, while the periodicity of the viewed and explored curves disappears after about one-third and two-thirds of the course, respectively. It is important to note that these curves do not account for registration date or for large times between user interactions. However, they do provide insight into the collective behavior



**Figure 7:** Certification fraction versus registration week relative to course launch. Bubble size represents the number of unique users registering each week.

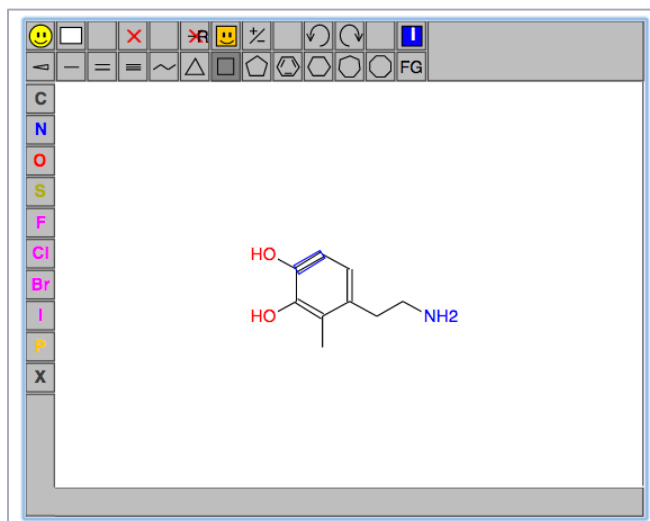
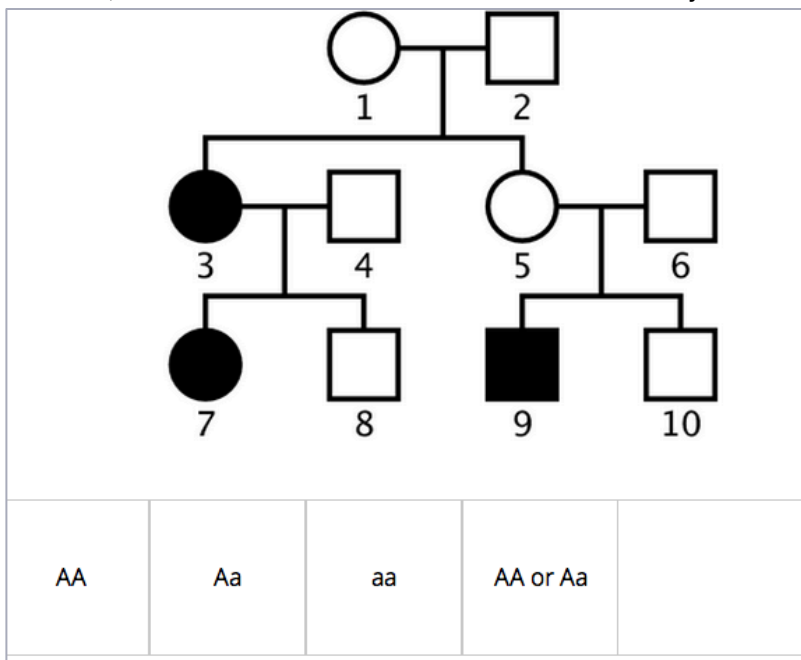
of our defined populations. It is particularly interesting to formulate hypotheses about behavior early in the course, e.g., attempting to target struggling students that may benefit from intervention.

Do students who register early have a higher likelihood of earning a certificate than students who register late? Figure 7 depicts the certification fraction versus registration week relative to course launch (bubbles indicate total enrollment each week). Across MITx courses, certification rates do not depend greatly on registration dates prior to launch, although some courses have higher registration rates very early as well as just prior to launch. For 7.00x, between 8% and 14% of users registering prior to launch earn a certificate regardless of enrollment week. After launch, the rate drops dramatically. Because of the highly structured due dates in 7.00x, it is reasonable to attribute most of the post-launch drop in certification rates to registrants who have missed due dates and find it difficult or impossible to catch up.

### Special Features

The 7.00x course was particularly unique; from the start, Professor Lander envisioned a wide variety of new assessment problem types, and insisted on the edX platform being shaped and adopted to fit his vision. As a result, students taking the course were able to experience nine new kinds of interactive problems (this doubled the number of types of problems available on the edX platform, as of the Fall of 2012). These new kinds of problems included:

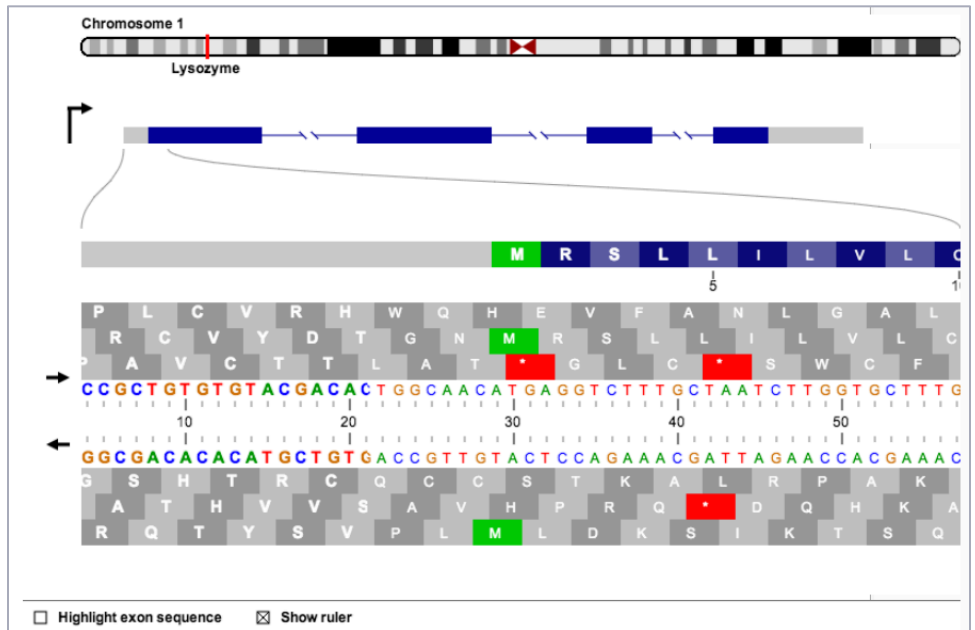
- Drag-and-drop:** Students given an image or diagram with empty boxes to be filled, and asked to drag icons from a list, and drop the icons into the boxes, to correctly answer a question. For example, the diagram to the right represents the heredity of humans with students asked to drag the appropriate genotype onto the empty boxes.
- Molecule Editor:** Students are presented an interactive editor with which students can draw molecule structures. Students are asked to edit a given molecular diagram, and change it so that the molecule has specified desired properties. For example, a dopamine molecule is shown to the right, and students are asked to change the molecule so it can make ionic bonds. Correctness of the diagram is evaluated by python code written by the course team, executed by the edX platform.



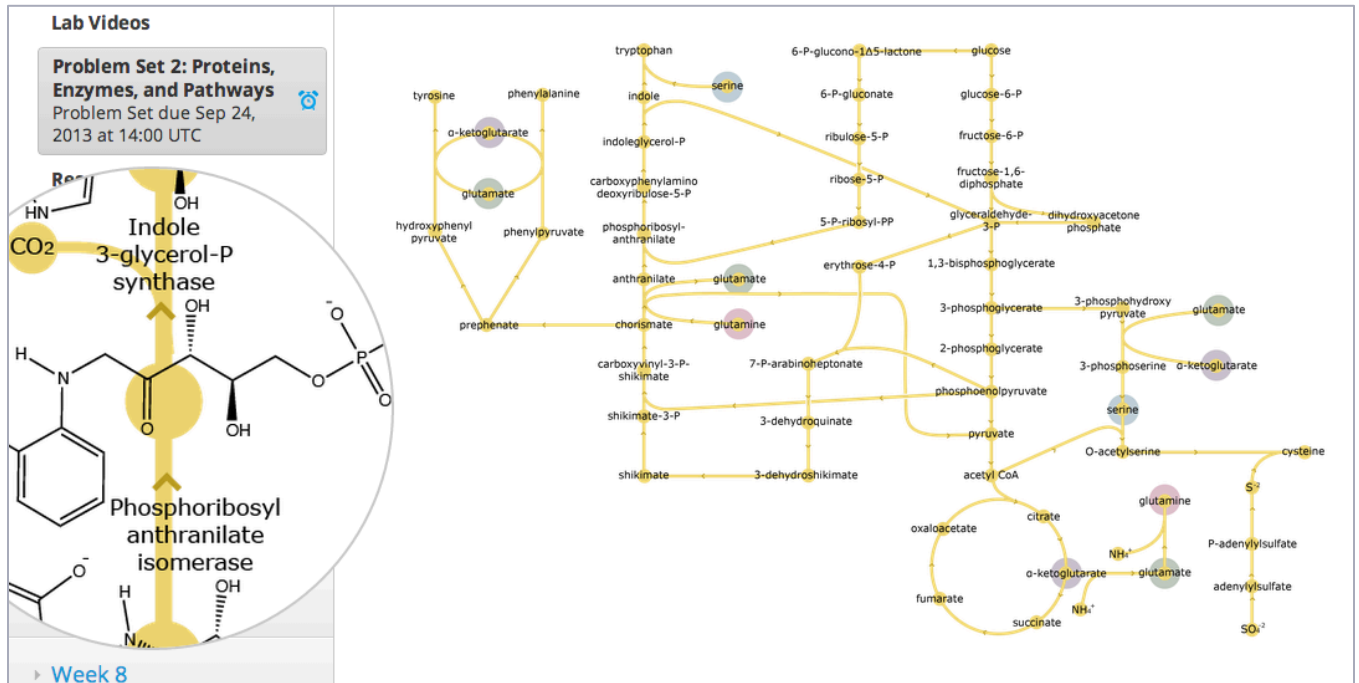




- Explore-a-gene:** Students are given an “Integrative Genome Viewer” tool, which allows a user to zoom in on a selected section of a genome, and interactively explore the sequence of DNA base-pairs. This tool enables problems that help students understand eukaryotic gene structure, expression, and mutation. The IGV is a custom version of a research tool developed at the Broad Institute and used by researchers worldwide.

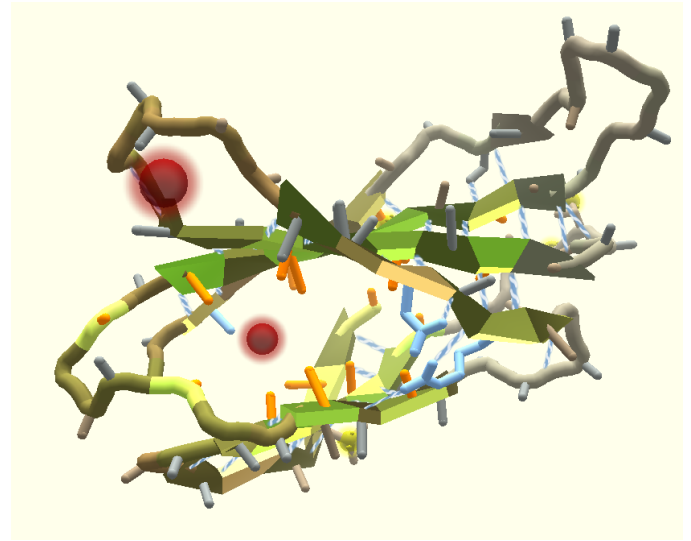


- Loupe and lightbox:** A common challenge in biology is depicting the entirety of a complex system, while allowing access to its details. The loupe and lightbox tool allows students to zoom in on large



diagrams, e.g. biochemical pathways, to access detailed information which is normally invisible on the large diagram, while also seeing the large diagram for perspective. This enables many kinds of exploration problems, e.g. about the synthesis of amino acids and function of enzymes within a depicted complex biochemical system.

- Foldit:** Foldit is a game where students fold simulated proteins in three dimensions. Points are earned for folding structures that properly take into account the many forces governing protein structure. Students used Foldit as part of one problem set and high scores were posted to a 7.00x Leader Board. Foldit was developed by the Center for Game Science at the University of Washington and customized for 7.00x.
- GeneX:** A simulation of gene expression where students are asked to make mutations with desired effects and design a functional gene “from scratch”. The simulation expresses their gene and determines if the goal has been achieved. Student used GeneX in a graded problem set and one of the exams. Professor Brian White developed GeneX.



**DNA: Promoter Terminator**

```

      10      20      30      40      50      60      70      80      90      100     110     120
5' -CAAGGCATAACCGAGATTGATGCCTTGTGGCGATAAGGTGTGCCCCCCCAAAGTGTCCGGATGTCGAGTGC CGCGTGC AAAAAAAAAA CAAAGGCGAGGACCTTAAGAAGGTGTGAGGGGGCGCTCGAT-3'
3' -GTTCCGATATTGGCTCTAACTACGGAACACGCTATTCCACACAGGGGGGGTTTCACAGCCTACAGCTCACGCGCACGTTTTTTTTTTGTTTCGGCTCCTGGAATTTCTCCACTCCCCGCGAGCTA-5'
    
```

**pre-mRNA: Exon Intron**

```

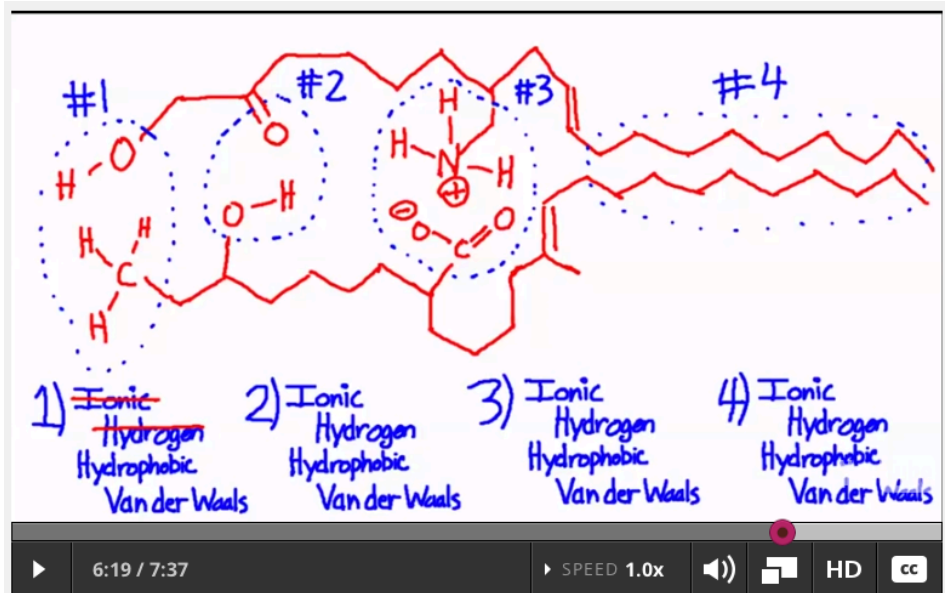
5' -CCGAGAUGAUGCCUUGUGCGAUAAGGUGUGUCCCCCCCCAAAGUGUCGGAUGUCGAUGUCGCGUGUCAAAAAAAAAA CAAAGCGAGGACCUUAGAAGGUGUGA-3'
    
```

**mature-mRNA and Protein (previous):**

```

5' -CCGAGAUGAUGCCUUGUCGGAUGUCGACCGAGGACCUUAGAAGGUGUGAAAAAAAAAAAAAA-3'
N-MetProLeuSerAspValGluArgGlyPro-C
    
```

In addition to these nine new problem types, 7.00x employed “Deep Dives” and “Lab” video segments that presented a discussion of a concept, explanation of problem solving, or a demonstration of basic lab experiments. Many of the deep dive videos were presented by MIT graduate students from the Biology department, offering a very personable connection to students, from students.



Throughout 7.00x, the course team put in special effort to keep student motivation high, offering various “Easter egg” videos, treasure chests of supplemental course resources and readings (e.g. from the New York Times, with recent news connected to course topics), and a segment on “Meet the Biology Faculty at MIT,” with a focus on diversity. In addition, the course offered a collaboration with the National Science Teachers Association, to facilitate practical classroom application of the courseware. And as a final touch, the course invited all of the students who qualified for a certificate to contribute their photo to a “DNA Mosaic,” as a joint product from the students and the course team.

