Authentic Science Inquiry Learning at Scale Enabled by an Interactive Biology Cloud Experimentation Lab

Zahid Hossain [1,2,*] , **Engin Bumbacher** [3], **Paulo Blikstein** [2,3], **Ingmar Riedel-Kruse** [1,*]
[1]Bioengineering, [2]Computer Science, [3]Education
Stanford University
*{zhossain@cs.stanford.edu, ingmar@stanford.edu}

ABSTRACT

National guidelines advocate for a more sophisticated STEM education that integrates complex and authentic scientific practices, e.g., experimentation, data collection, data analysis, and modeling. How to achieve that is currently unclear for both presential and distance education. We recently developed a scalable cloud lab that enables many online users to perform phototaxis experiment with real, living Euglena cells (opposed to just simulations). Here we iteratively designed and deployed an open course on the edX platform including suitable user interfaces that facilitates inquiry-based learning on this cloud lab: Online students (>300) run real experiments (>2,300), performed data analysis, explored models, and even formulated and experimentally tested their own hypotheses. Platform and course content are now suited for global adaptation in formal K-16 education. We will demo our cloud lab at the conference.

ACM Classification Keywords

C.2.4 Computer-Communication Networks: Distributed Systems; K.3.1 Computers and Education: Computer Uses in Education; I.6.7 Simulation and Modeling: Simulation Support Systems; H.5.2 Information Interfaces and Presentation (e.g. HCI): User Interfaces; I.4.7 Image Processing and Computer Vision: Feature Measurement

Author Keywords

inquiry-based learning; cloud lab; remote experimentation; life science; biology; interactive biotechnology; Euglena; phototaxis; modeling; data analysis; user interface; user studies; learning analytics; MOOC; edX; education

INTRODUCTION

An unsolved challenge for presential and distance learning is the proper design and integration of tools that enable authentic scientific practices [2], especially with real experimentation [4], e.g., living organisms, as opposed to mere simulations.

Inquiry-based practices in which students construct knowledge like professional scientists [9] are at the core of national guidelines such as the Next Generation Science Standards of the United States [10, 3], and the benefits of these approaches for science education have been demonstrated [9]. Many *remote labs* [4] with real experimentation have been developed but these labs are usually not suitable for large scale deployment due to lack of concurrent access design, and none of them includes living matter. We recently developed a *cloud lab* that allows remote users to perform real biology experiments (with live organisms, opposed to just simulations) at high-throughput with concurrent user access (akin to cloud computation, Fig. 1) [6]. The contributions of this paper are: (1) We developed and deployed an open online course that enables for the first time inquiry-based learning at scale over the Internet that combines real experimentation, augmented data analytics and modeling tools (in line with bifocal modeling [1]). (2) We report on how to scaffold authentic inquiry tasks online and at scale with such cloud labs.

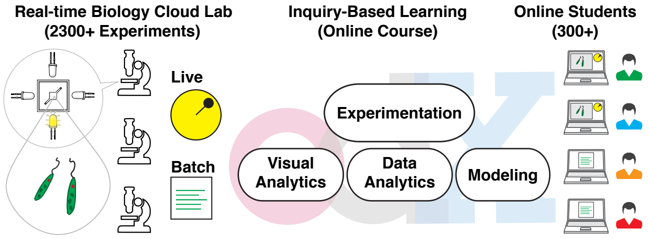


Figure 1. We integrated a scalable biology cloud lab, designed as a distributed system, into an open online course to enable authentic science practices. Real biology experiment (with live organism) can be executed in real-time interactive manner (*live*) as well as with offline (*batch*) processing. We explored the general affordances and design rules of online experimentation science labs to enable inquiry-based learning.

RESEARCH AND ONLINE COURSE DESIGN

We adopted an iterative design-based research approach building on our previous pilot studies [6]. We converged on a mini-course (∼ 4h over 1 week) on the Open edX platform that is suitable for a diverse MOOC audience (middle school and university students, science teachers). The course theme centered on the core scientific practices (Fig. 2A), which the students executed while investigating the phototactic behavior of the single-celled organism *Euglena gracilis* (Fig. 2B,C). Importantly, our fully automated cloud lab technology enables students to execute real experiments with living cells from across the world 24/7 (Fig. 3).

The course had six highly scaffolded units [5]. Each unit introduced a new *scientific practice*, a new *biological target concept*, and a new *UI tool* as described below:

L@S 2017, April 20 - 21, 2017, Cambridge, MA, USA

ISBN 978-1-4503-4450-0/17/04. . . $15.00
DOI: http://dx.doi.org/10.1145/3051457.3053994

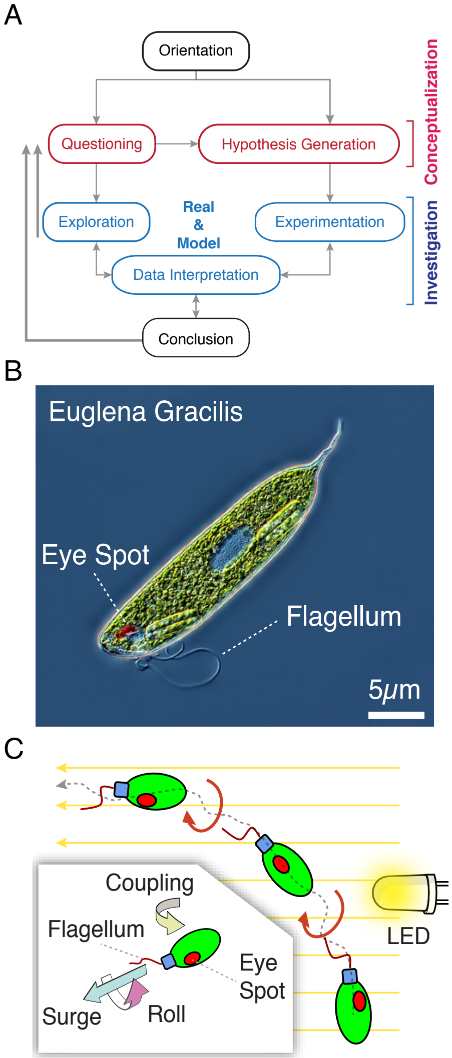


Figure 2. Cloud-based biology experimentation with living cells. (A) Inquiry-based learning emphasizing exploration and experimentation with real specimen and models (adapted from [9]). (B) Single-celled *Euglena gracilis.* (C) Negative phototaxis mechanism of Euglena.

Unit 1: Observation (online microscope)
Students were introduced to the cloud lab dashboard interface (Fig. 4A), i.e., how to select an online microscope, how to observe the Euglena cells in real time, and how to watch the resulting experimental video afterwards either directly on the website via streaming or after downloading. Students were tasked to describe their observations in a free form manner. We deliberately started scaffolding with a passive observation, rather than asking students to already explore the light responses because our earlier pilot studies have shown that premature interactivity without the proper foundation being established could overwhelm students, especially when working with a noisy biological system. This passive observation was then followed by a short description on the basic biology of Euglena: they are photosynthetic organisms that detect light using an eye-spot; there was no mention of phototaxis yet.

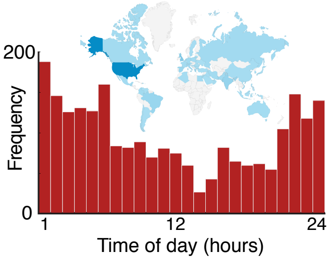


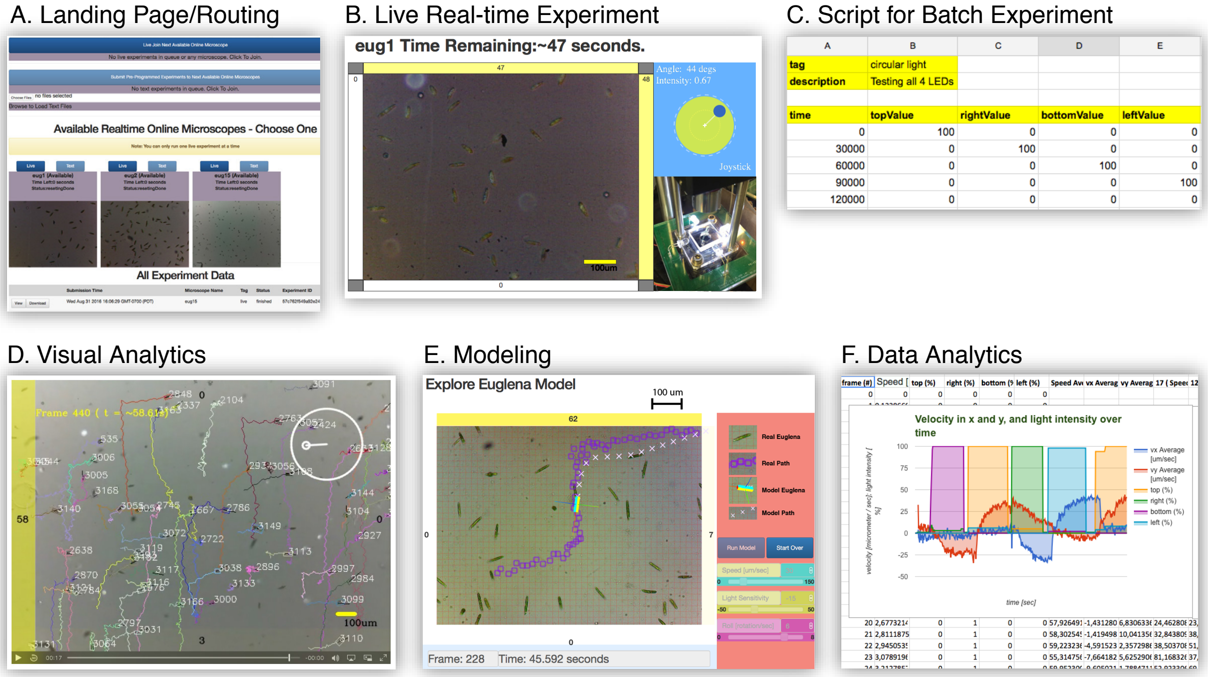
Figure 3. Number of experiments run over 24 h period over the 6 weeks of deployment. The inset map shows the distribution of incoming traffic from different geographic locations.

Unit 2: Experimentation (interactive live and scripted batch)
Students were introduced to the virtual joystick to actuate directional light stimuli on the interactive online microscope (*live* mode) (Fig. 4B). Students were prompted to then run experiments to explore how Euglena reacts to light stimuli. In order to eradicate misinterpretation of the instrument usage early on, we primed students with simple test questions, e.g., *"In which direction does the light shine when you pull the joystick in this direction?"* We also introduced the *batch* mode (Fig. 4C), which is suitable for scripting controlled and repetitive experimentation, but importantly also counters low Internet bandwidth that is unsuitable for the *live* mode.

Unit 3: Visual analytics and qualitative data interpretation
Students were instructed to analyze and explain their movie data more closely (Fig. 4D), where Euglena exhibit a wobbling, meandering motion as apparent via the overlaid tracks. Students then performed simple, direct measurements regarding speed and rolling frequency - solely based on visual analytics using these overlaid tracks, the timer and the scale bar.

Unit 4: Model exploration and evaluation
Students explored models (Fig. 4E) that provide a mechanistic explanation of the phototaxis phenomena in the absence of real-life noise. Students were tasked to find the best parameter values that fit the model to a real Euglena path and then explore how to accomplish both positive and negative phototaxis. Students were introduced to the relevant sub-cellular structure of Euglena and the mechanistic explanation of Euglena phototaxis, i.e., the coupling of the eye spot with the flagellum (Fig. 2B,C), which causes rolling around the long axis and side way turning. These activities then also provided a deeper explanation for the wobbling motion analyzed in the previous unit. Students were also asked to go back to the real experimentation and compare.

Unit 5: Quantitative data processing and analysis
Students engaged in the process of exporting the numerical data into a google spreadsheet, graphing that data, and interpreting the graphs. Students first worked through a highly scaffolded example to analyze how the Euglena speed depends on Light-On versus Light-Off, which typically has a weak effect. Then the students were asked to perform a similar analysis on their own, but now to determine graphically whether the average velocity vectors changed with directional light stimulus from the LEDs (Fig. 4F). Here the students received



	A	B	C	D	E
	tag	circular light			
	description	Testing all 4 LEDs			
	time	topValue	rightValue	bottomValue	leftValue
	0	100	0	0	0
	30000	0	100	0	0
	60000	0	0	100	0
	90000	0	0	0	100
	120000	0	0	0	0

Figure 4. Enabling the key components of inquiry-based learning at scale. (A) Landing page to route students to a suite of online microscopes. (B) Real-time Euglena biology lab in *live* interactive mode. (C) Experiment script in CSV format for *batch* mode. (D) A playback movie viewer for automatic tracking of Euglena cells. (E) Modeling applet simulating Euglena overlaid on a pre-recorded video. (F) Google Sheets for data analytics.

the data in a format where the velocity of each cell was already decomposed into its cardinal directions (x and y). Depending on the direction of the light origin, one velocity component would average to zero, and the other would be either negative or positive. Students had to design and run a new set of experiments to generate the data for this analysis. During pilot studies we had offered units 4 and 5 in reverse order, but user feedback revealed that it is more effective to do the interactive modeling activity before the comlex data analysis.

Unit 6: Open and self-guided investigations

Students were prompted to carry out a self-guided research activity, where they proceeded through the main parts of the inquiry cycle (Fig. 2A) while applying all or most of the previously used tools (Fig. 4). They were asked to make an observation (specifically one that had not been stated by the course material previously), and transform this observation into a testable hypothesis with experimental designs. Students were then encouraged (optionally) to pursue the actual experimentation, analysis, and interpretation.

COURSE DEPLOYMENT AND RESULTS

The course (3.5 ± 1.1 h to finish on avg.) was offered six times with minor updates between successive offerings. Students (N=325, active) came from 46 countries (Fig. 3) with 47% female and a median age of 32 years (IRQ=19). A total of $\sim 2,300$ experiments were executed (*live* and *batch*) with little wait time (median: 4.8 s, IRQ=1.55 s) for *live* experimentation due to distributed geographic timezones (Fig. 3). The capacity of the cloud lab would have enabled >500 students/week.

As a case study, here we consider the work flow of a single student. This student ran 13 experiments (nine *live* and four *batch*) and completed the course in ~ 3 h. She ran five experiments, including one *batch*, before noting that Euglena moves faster upon light stimulus but without a clear direction of motion. Later, she executed three more experiments on different online microscopes and thereby experienced biological and system variabilities. She then ran three more experiments before formulating her hypothesis that *"Euglena respond when light intensity is above the threshold of* 50%*."* To test this hypothesis, she ran two more carefully designed *batch* experiments. Her data analysis in google sheets revealed that the vertical velocity component of the cells increased with increasing light intensity, while the horizontal component remained at 0 μm/s (Fig. 5). Hence instead of the hypothesized discrete threshold, she found a more gradual response. Thus, our cloud lab platform and the associated online course enabled students to perform realistic science-practices including self-formulating and experimentally testing hypotheses, and where the various interaction modalities (Fig. 4) brought forward the nuances and natural variability of real biology.

Our logged data from all participants revealed a $\sim 51\%$ completion rate, which is high even for a 1-week course. Here our user interfaces and inquiry activities appeared to have been effective, e.g., *"The tangibility that lab provided was one important reason that carried me on...[sic]"*, *"The way we could conduct experiments remotely was very cool!"*, *"Feeling like I was part of real research"*. Crucially, the course changed student attitudes toward science, e.g., 15 voluntarily responded on a scale of 1-9 (*"not at all"* to *"totally"*), before and after participating in our course, to the statements *"Ordinary people can be scientists"* $(8.7 \pm 0.6 \rightarrow 7.5 \pm 2.0, p < 0.05)$ and *"I can imagine myself as a scientist"* $(6.7 \pm 1.8 \rightarrow 8.5 \pm 1.3, p <$

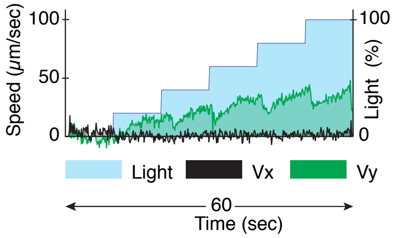


Figure 5. Students formulated and experimentally tested their own hypotheses. Here this student expected an intensity threshold for light responsiveness, instead a more complex relationship between the light intensity (increased by 20% every 10s) and the Euglena swarm velocity was found (Vx/Vy: average horizontal/vertical velocity components).

0.005). The vast majority of the students, ~ 72%, indicated great value in having real living organisms (opposed to pure simulations), e.g., to capture ground truth details that simulations would miss. Overall, students rated their experience between *"very"* and *"extremely"* positive (6.3 ± 0.6 on a 1-7 scale, N=34); difficulty was rated as between *"neutral"* and *"somewhat easy"* (4.6 ± 1.1 on a 1-7 scale, N=31). A more detailed analysis of all student activities and data will be published elsewhere.

DISCUSSION

Achievements and implications: We demonstrated, that our cloud lab enables concurrent online users at scale to experiment with real biological materials, which is a key technological advancement for online learning including MOOCs. We augmented this cloud lab with data analytics and modeling components, which enabled inquiry-based learning with authentic, high-dimensional, "interactive biology" experiments [7, 8]. The key design features of our platform not only mitigate non-deterministic and noisy biological behavior but actually exploit its educational value. Recognizing the inherent variability in biological systems empowers students to recognize differences between a determinist model and real biology (bifocal modeling [1]). The science practices that this cloud lab affords are a paradigm shift of what students currently can do in either presential or online life-science education. Multiple K-12 science teachers who took this course expressed interest to use these and related activities with their own students, the corresponding deployment studies are currently under way.

Future work: There are multiple important avenues for future research and development with cloud-based experimentation and pedagogy. (1) Utilizing platform data for learning analytics. (2) Refining and testing course content for specific learner groups. (3) Including other relevant scientific practices such as collaborative teamwork and model building. (4) Extending the platform to other experiment types.

Conclusion: We successfully deployed an open online course with an integrated biology cloud lab (with real live organism) in a scalable manner. Students could engage in the core activities of scientific inquiry while interacting with living cells, which goes significantly beyond current educational practices of passive observation through a microscope or using computer simulations or animations. Instead, the lab automation and ease of data collection and analysis leads to easier logistics and extended lab time for students, also when working from home. Ultimately, this approach could bring authentic science practices to millions of students annually across the globe.

Acknowledgements: We would like to thank A. Brauneis, M. Diaz, A. Saltarelli, J. Cash, and the Riedel-Kruse lab.

REFERENCES

1. P. Blikstein. Bifocal Modeling: Comparing Physical and Computational Models Linked in Real Time. *Playful Learning Interfaces*, pages 317–352, 2014.
2. C. A. Chinn and B. A. Malhotra. Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86(2):175–218, 2002.
3. N. R. Council. Guide to Implementing the Next Generation Science Standards. Technical report, Committee on Guidance on Implementing the Next Generation Science Standards, Board on Science Education, Division of Behavioral and Social Sciences and Education, Washington, DC: The National Academies Press, 2015.
4. R. Heradio, L. d. l. Torre, D. Galan, F. J. Cabrerizo, E. Herrera-Viedma, and S. Dormido. Virtual and remote labs in education: A bibliometric analysis. *Computers & Education*, 98:14 – 38, 2016.
5. C. E. Hmelo-Silver, R. G. Duncan, and C. A. Chinn. Scaffolding and achievement in problem-based and inquiry learning: A response to kirschner, sweller, and clark (2006). *Educational psychologist*, 42(2):99–107, 2007.
6. Z. Hossain, E. W. Bumbacher, A. M. Chung, H. Kim, C. Litton, A. D. Walter, S. N. Pradhan, K. Jona, P. Blikstein, and I. H. Riedel-Kruse. Interactive and scalable biology cloud experimentation for scientific inquiry and education. *Nat Biotech*, 34(12):1293–1298, Dec. 2016.
7. Z. Hossain, X. Jin, E. W. Bumbacher, A. M. Chung, S. Koo, J. D. Shapiro, C. Y. Truong, S. Choi, N. D. Orloff, P. Blikstein, et al. Interactive cloud experimentation for biology: An online education case study. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3681–3690. ACM, 2015.
8. S. A. Lee, A. M. Chung, N. Cira, and I. H. Riedel-Kruse. Tangible interactive microbiology for informal science education. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 273–280. ACM, 2015.
9. M. Pedaste, M. Mäeots, L. A. Siiman, T. De Jong, S. A. Van Riesen, E. T. Kamp, C. C. Manoli, Z. C. Zacharia, and E. Tsourlidaki. Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational research review*, 14:47–61, 2015.
10. H. Quinn, H. Schweingruber, T. Keller, et al. *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press, 2012.