Tracking student use of web-based resources for chemical education

Authors

Robert Bodily and Steven Wood

Abstract

We have developed an analytics system for use by general chemistry students that tracks video and quiz interaction data and reports it back to students in real-time. To this end, we have developed a learning analytics dashboard for students. When students access the dashboard, they can easily ascertain the gaps in their knowledge. Furthermore, they can click on a flagged concept to access video resources, practice questions, or web resources to remediate their lack of knowledge. We have tracked the use of our initial dashboard version by students in the first semester of a two-semester general chemistry sequence. We report on the student use of web resources in this course and provide recommendations for researchers and practitioners in chemistry education based on our results.

Background

As online and blended learning in chemistry education continues to grow (Allen & Seaman, 2014), it is increasingly important to understand how students interact with resources online. Once we begin to understand how students use web-based resources, that data can be used to improve the design, the access, and the content with the object in mind to improve the learning potential of these resources. This is the focus and objective of learning analytics. The learning analytics process includes selecting data, capturing data, using data, and acting on data (Lias & Elias, 2012). We set out to develop a learning analytics system, in the form a class dashboard, to capture student use data as they interact in real-time with course content such as videos and quizzes. We had three major research questions that we initially set out to answer.

Research questions

R1: How can student use of web-based resources be tracked in open learning environments?

R2: How are students using the videos, quizzes, and dashboard provided to them in the course?

R3: What is the relationship between student use of web-based resources and overall performance in the class?

Design and Development of the Learning Analytics System or Class Dashboard

Technical Infrastructure

The following technical term definitions are provided to help the reader understand the analytics system that we constructed:

- **Learning Record Store (LRS)**: a database to store student interaction events
- **Learning Management System (LMS)**: an online platform to host content and keep track of student grades
- **Experience API (xAPI)**: a standardized data format that is consistent across applications. It uses a “verb”, “actor”, and “object” framework to identify what has happened in what context.

A graphical description of our analytics system can be seen in Figure 1.

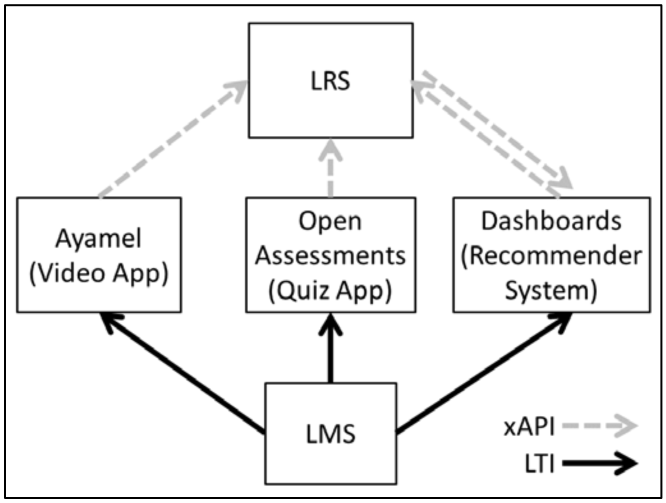

Figure 1. Our web-based resource student use tracking system.

Using LTI, students can launch to external learning applications from the LMS. This allows a student to be tracked across systems. It also provides a streamlined experience for the student because they do not have to sign in again when they use another learning application. Experience API is a data format standard that allows data to be passed in the same format to our LRS from multiple applications. The dashboard can then pull data out of the LRS in real-time to provide students with their performance data in real-time in the form of a visual report. This system enables us to collect student interaction data as students interact with quizzes, videos, and the dashboard in the course. Table 1 provides a description of the types of data we collect in our system.

Table 1.
Data points collected or calculated in our analytics system.

Video Analytics	Quiz Analytics	Dashboard Analytics
# of plays	# of question attempts	# of dashboard views
# of pauses	Time spent on a quiz	Time spent in dashboard
# of skip forward	# of quizzes attempted	# of video suggestion clicks
# of skip backward	Average confidence level	# of quiz suggestion clicks
# of play rate changes	Max number of attempts	# of unique visits to dashboard
# of volume changes	Max time on a quiz	

In order to provide a better understanding of our system, we have included screenshots of the dashboard tool (Figure 2), quiz application (Figure 3), and video application (Figure 4). The videos used in the class can also be accessed at http://chempath.byu.edu/.

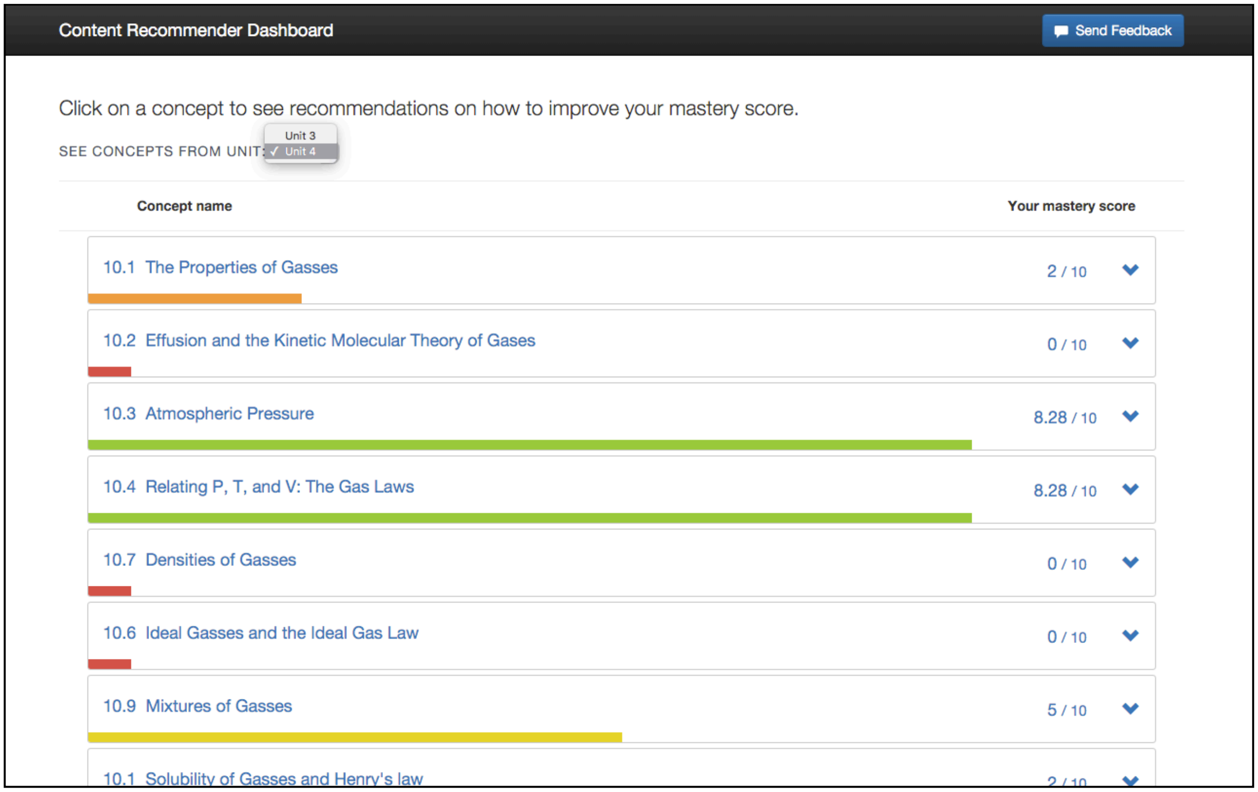

Figure 2. The content recommender dashboard displaying student mastery scores.

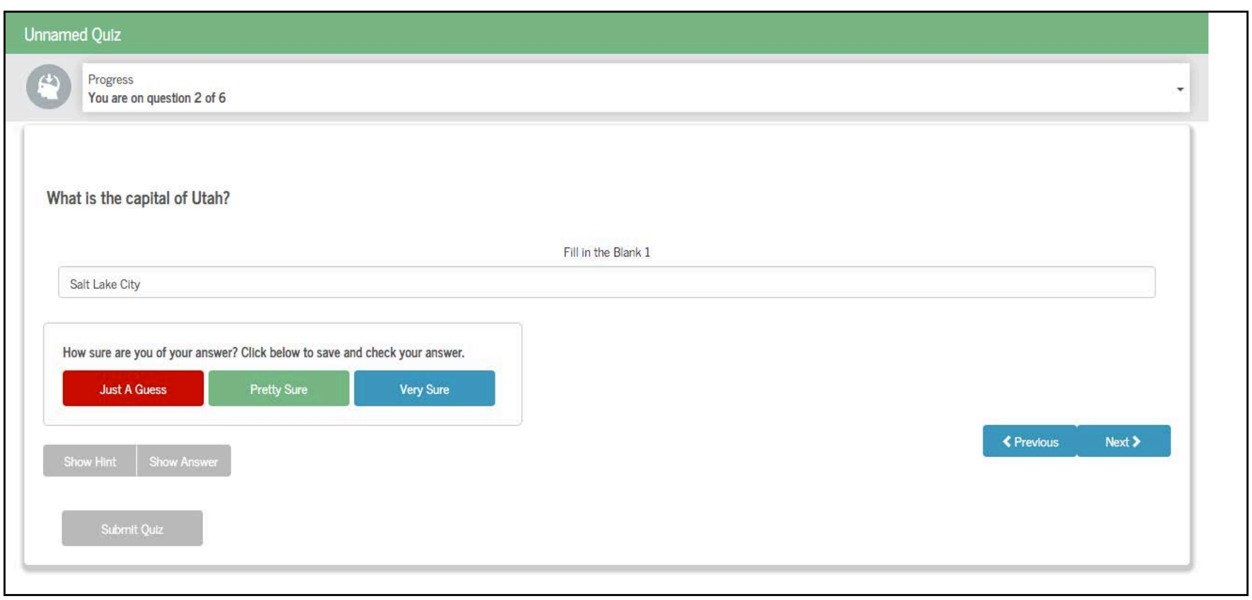

Figure 3. A demo quiz question illustrating our assessment system.

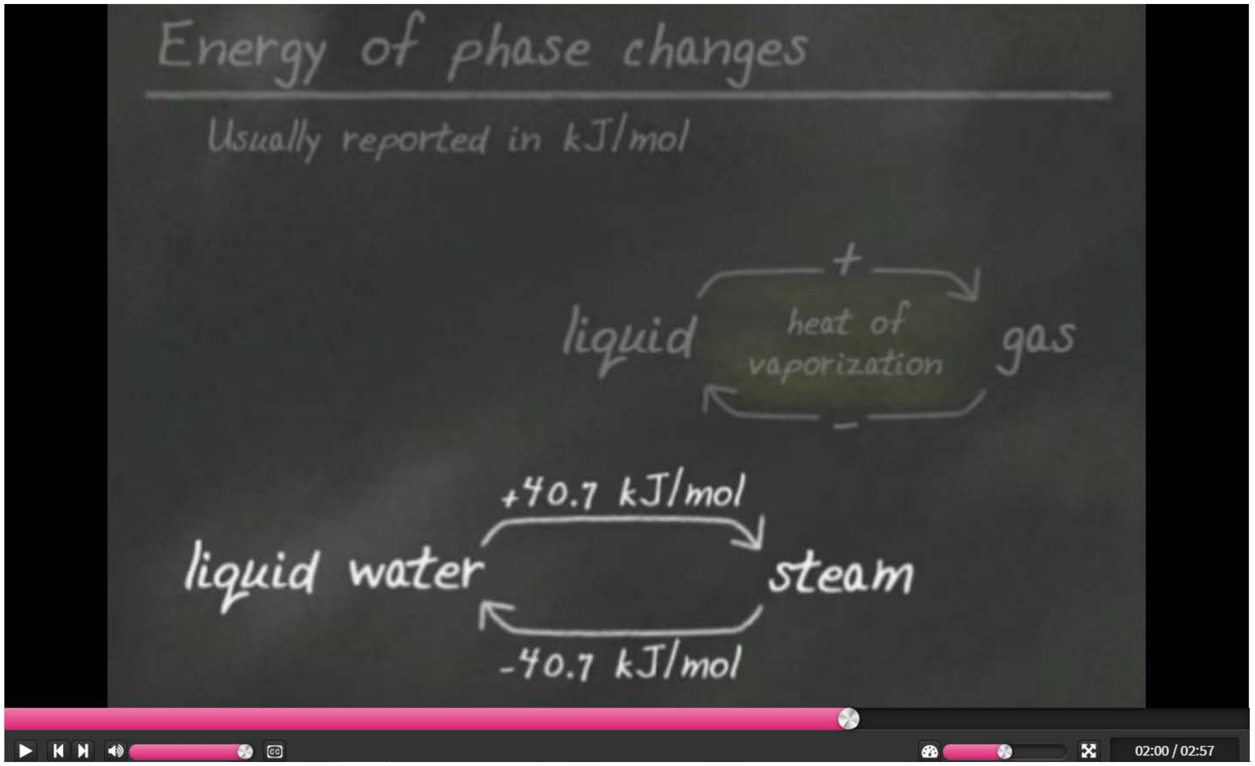

Figure 4. An example of one of the course videos in our video player.

Methods

During the Fall semester of 2015, we initially implemented our web-resource tracking system in a general level blended chemistry course at a large western US university. There were 200 students that consented to participate in the study who allowed us to collect their data. The chemistry course content that was available to these students via the system we had developed consisted of (1) high quality videos developed with animations and audio, (2) short formative quizzes related to the videos in the course, and (3) a dashboard that provided students with information on their knowledge gaps and metacognitive skill abilities based upon the use and performance data generated on their performance and use.

Relationship between Online Resource Use and Overall Performance in the Course

In order to identify which use elements were predictive of student grade in the class, we ran a linear regression. The variables used as independent variables in the class were taken from the student use of web-based resources. However, it should be noted that some of the variables were removed from the model as some of the variables were too similar to each other (multicollinearity). The final list is defined in Table 2 below.

Table 2
Descriptive statistics for variables use in the analysis

Variable	Mean	Std	Definition
total (Overall) Percent	81.34	12.92	Grade in the class
suspended	117.21	69.96	# of times navigated away from the quiz
paused	105.91	164.1	# of times paused a video
jumped	84.81	133.47	# of times clicked to go to another place in video
changed play-rate	20.09	44.27	# of times changed the video play-rate
showed-hint	13.44	14.42	# of times clicked show hint in a quiz
confidence level	1.28	0.42	Average of self-reported confidence in answers
play-rate	1.09	0.19	Average play rate in videos
first Attempt Quiz Score	57.37	24.85	Score on quizzes based on the first attempt only
follow Feedback	0.75	2.62	# of times followed feedback in dashboard
duration	402.83	353.75	Time it takes to take a quiz
late night	0.04	0.06	Percent of events after 11pm and before 5am
consistency	0.15	0.06	Percent of days a student accesses online material
knowledge awareness	0.34	0.11	Percent of high confident correct answers
persistence	1.55	0.35	# of attempts without clicking to see answer
deep learning	0.12	0.1	# of questions without clicking show answer

We first discuss the results of some general resource use average trends among video, quiz, and dashboard use, and then we discuss the results of a multiple-linear regression predicting student final grade using these resource use variables.

There were 215 videos available to students in the course. However, the optional video use was low (see Figure 5).

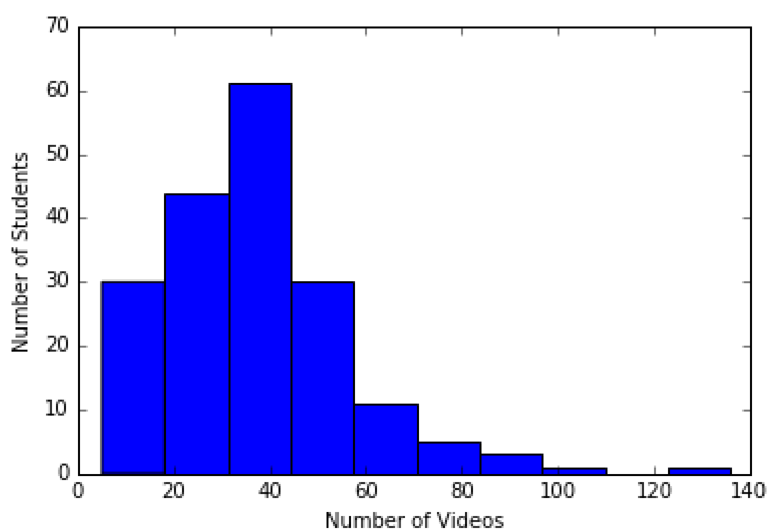

Figure 5. The number of videos watched by each student.

Despite the fact that there were over 200 videos available for students, the median student number of videos accessed was about 40. These videos were high quality videos developed from an NSF grant with animations synchronously delivered with a concise audio text. It was disappointing for us to see that even when providing excellent resources for students, many of them chose not to use the videos. Our initial thought in creating the short quizzes that were correlated with the videos was that the quizzes would encourage students to view the videos prior to attempting the quizzes. Even with quizzes related to the videos, video use was still low. We suspect that this low use could have resulted from the quizzes being too easy and students were able to do well on the quizzes without having to watch the videos. This could also happen because students had a number of other resources available to them that they could have used instead of the videos.

Student use of the dashboard paralleled video viewing. We found that only about 40% of students accessed the dashboards, and even among those that accessed the dashboard, they typically only accessed it a few times.

These video and dashboard use data highlight the importance of identifying how best to support student engagement with web-based resources and feedback. If time and effort is invested in the creation of resources that students do not utilize, that effort is wasted. Tracking student use is an important means for evaluating not only the effectiveness of web-based materials, but also to determine the best ways to incorporate these materials as an integral part of the course such that student benefit.

We used the collected data for the analytic system to see if student use of the resources was correlated with a student's overall grade in the course. We used linear regression to determine if elements of student resource use are predictive of student final grade. The results of our regression can be seen in Table 3 below.

Table 3

Linear regression results predicting final class grade

Variable	Coefficient	Std Error	t	p-value
suspended	-5.24E+04	2.28E+04	-2.295	0.023
paused	-3.06E+04	1.10E+04	-2.781	0.006
jumped	2.42E+04	1.42E+04	1.703	0.09
changed play-rate	2.38E+04	4.22E+04	0.564	0.574
showed-hint	1.38E+05	9.97E+04	1.38	0.169
confidence level	1.61E+07	3.47E+06	4.631	0
play-rate	5.89E+05	9.36E+06	0.063	0.95
first Attempt Quiz Score	9.52E+04	8.59E+04	1.109	0.269
follow Feedback	6.09E+05	4.97E+05	1.225	0.222
duration	-1923.9176	3679.677	-0.523	0.602
late night	-4.55E+07	2.03E+07	-2.246	0.026
consistency	1.11E+08	3.00E+07	3.685	0
knowledge awareness	-6.93E+07	1.37E+07	-5.071	0
persistence	2.64E+05	3.90E+06	0.068	0.946
deep learning	-2.97E+07	1.70E+07	-1.752	0.082

Note: Coefficients are large because the dependent variable had to be transformed to satisfy the assumptions of linear regression.

As can be seen in the table, pausing, changing tabs frequently during homework, and working during the middle of the night were negative predictors of student achievement. Jumping around in the video and being confident in responses were positive predictors of student achievement. Although these results are not causal, they do highlight indicators that may be useful in identifying students struggling in the course.

Implications for Practice and Research

Based on the results of this limited study, we provide recommendations for chemistry education researchers and practitioners. First, when investing time and money in educational resources, take the time to collect data on how students are actually accessing and utilizing the resources created. This information provides instructors or designers the actual and not the self-reported feedback required to redesign certain aspects of their online material and their courses. Second, tracking student use of web-based resources results in information that is unobtrusively collected that can be used in prediction algorithms to help identify who is struggling in the course. We envision that these variables can then be included in some sort of instructor dashboard to help instructors in the student remediation process. Third, additional research is needed to understand how to support students in metacognition, and taking advantage of feedback provided by systems like the dashboard of our analytics system. Perhaps the major question raised by our findings is best started as, why are students choosing to not take advantage of resources we are providing for them?

Conclusion

This article has described the technical infrastructure needed to collect student use of web-based resources, trends and issues encountered by analyzing this student use, and a possible predictive model to predict student grade in the course using a learning analytics system that collects actual student use of online resources. Implications for practitioners and researchers based on the results of this paper have been provided.