Assessment with E-Textbook Analytics

Combining Multiple Sources of Data for Effective Assessment of Educational Technology

ECAR RESEARCH BULLETIN | February 15, 2015

Sam Van Horne, University of Iowa
Jae-eun Russell, University of Iowa
Kathy L. Schuh, University of Iowa

Overview

Higher education administrators and faculty are increasingly interested in determining how large amounts of student data in electronic systems can be used to promote student success. For example, in their recent report of the results from the 2014 EDUCAUSE faculty survey, Dahlstrom and Brooks reported that faculty are increasingly interested in learning analytics, specifically, how it can support students who are struggling to succeed in their courses. Whereas educational-technology applications are often adopted based on a set of criteria related to how well the application assists students or instructors, administrators now want to go further with educational technology and use these data to inform models that predict student outcomes. But this process is often made more difficult by a lack of understanding of what data are typically stored in such systems, how the data may be predictive of student success, and how the data can be used to directly intervene in educational settings.

The purpose of this research bulletin is to provide concrete strategies, grounded in research with e-textbook analytics, on how to use data from interactive platforms to inform decisions about supporting student learning with educational technology. Given the very real difficulties of leveraging complex analytics data to increase student success, higher education administrators must decide whether similar actionable information can be obtained from more subjective measures, such as surveys and ethnographic methods, including focus groups and semistructured interviews. In a climate of dwindling resources for higher education, administrators must determine how to conduct effective assessment that is both cost-effective and rigorous enough to inform decision making.

Our recommendations are grounded in a data-intensive study of student usage of educational technology. We conducted a research project to examine the effectiveness of an interactive e-textbook platform in a set of courses across a variety of academic disciplines. We used a control group of students who used a paper textbook in a course that was in a similar field to a course that had adopted the e-textbook platform. To learn how students used their e-textbooks, we administered surveys at the beginning and the end of the semester to students who used an interactive e-textbook. In addition to collecting from the vendor all of the analytics related to student and instructor usage of the e-textbook, we administered weekly electronic reading journals to students to understand their reading behaviors.

Thus, for the students in the e-textbook group, we had both objective and subjective measures of e-textbook usage. Combining these multiple sources of data, we effectively assessed the implementation of the e-textbook platform. We also learned important lessons about integrating student-level data from
an interactive e-textbook system. We believe the findings from our analysis can be applied to the greater field of learning analytics research because both areas involve the issue of how to translate research with large data sets into effective practices with real benefits. This bulletin explores the following recommendations:

1. Analytics can be useful for determining the amount of engagement but may not be able to provide insight into the quality of interaction.

2. Analytics can be biased if analysts fail to introduce procedures to account for the abandonment (or simply nonusage) of educational technologies.

3. Data gathered from users prior to the semester can provide important context for usage patterns later in the term and lead to the development of interventions.

Highlights

We conducted a mixed-methods research study to investigate how students and instructors used a free, interactive e-textbook platform. Because it was free and was accessible only in the learning management system (LMS), all students had access to the e-textbook. Our particular research questions centered on how students used e-textbooks during the semester, and we designed a study to investigate both students’ reading behavior and how much they accessed and used the markup features in the e-textbooks. Previous research has shown mixed results for student preference for e-textbooks. Some researchers have found that students preferred e-textbooks. Others have found that students preferred a paper textbook.

We recruited 287 students in 8 courses that used an e-textbook and 263 students in 9 matched courses that used a paper textbook. This bulletin focuses on the e-textbook users. To investigate students’ reading behaviors with the e-textbook, we administered electronic reading journals during the semester. Each week we asked students whether they had access to the e-textbook, how satisfied they were with their access, and how much time they spent reading the e-text. After weeks 3, 6, 9, and 12, we asked students to answer a series of questions adapted from the Reading Behavior Questionnaire. For the other weeks, we asked students to describe their reading processes in a brief narrative.

From the e-textbook vendor we collected all measures of students’ and instructors’ usage of the e-textbook platform. This included data about online and offline reading and about how the participants interacted with the e-textbook. We collected data about the markup features, including the number and text of highlighted sections, notes, questions, tags, and annotations.

Working with Vendors and Obtaining Analytics

In our research, it was vital to negotiate access to and obtain definitions of the variables contained in the database of student interactions in the e-textbook platform. All of these e-textbook analytics were generated from students’ reading and usage of markup tools in the platform. Obtaining this information required developing relationships with the vendor to learn about available data and securing approval from our institution to use these data in research.
Developing Relationships with a Technology Vendor. In an analytics project with data from an external provider, a positive working relationship with the vendor is essential. We initiated conversations about obtaining the analytics data before the start of the semester by describing our project in detail. Through phone and e-mail discussions, we obtained a list of the data that would be available, provided that we obtained informed consent from the students. This is an important aspect of our research: Vendors may have particular procedures that apply to the sharing of data collected in their systems. In our case, it was important to identify the data types early and ensure that our institutional review board (IRB) could approve the collection of the available data. Most importantly, we incorporated a procedure in our study to receive a partial download of data during the semester. This procedure was essential to our ability to understand and immediately use the full data set we received at the end of the semester.

Asking Many Questions about the Data. Because we had not worked with e-textbook analytics before this research study and were unfamiliar with the data, it was important to ask questions. Some questions were basic. We learned through interacting with the vendor that a page was measured as “read” after it had been open for 10 seconds. In addition, in checking the final data sets, we discovered that 13 of the 287 students in our study were not present in the reading data. We followed up with the vendor to ensure this was not a mistake, and we concluded that these students had never accessed the e-textbook. In the database of markups (which included almost 30,000 records of student behavior), we identified duplicate records that we needed to remove. It was imperative for us to compute descriptive statistics and be in conversation with the vendor to ensure that we were comfortable with analyzing the analytics.

Separating the Analysis of Analytics from Self-Reported Measures

The main purpose of this research study was to determine whether students who used an e-textbook achieved better learning outcomes and had a better overall reading experience than students who used a paper textbook in a matched course. Our institution was interested in exploring whether to promote e-textbook adoption on a larger scale, so our research was originally designed to focus on the comparison of users of e-textbooks and paper textbooks.

Overall Findings. Compared to students who used a paper textbook, students in the e-textbook group had significantly lower average satisfaction with their access to the textbook. Students in the e-textbook group indicated in their reading journals that they did not always have reliable Internet access and that they could not always access the e-textbook in the university’s LMS. Figure 1 shows that students’ satisfaction with access to paper textbooks was higher than students’ satisfaction with access to e-textbooks throughout the entire semester. At the end of the study, students in the e-textbook group, on average, indicated that they would prefer to use paper textbooks in the future. In general, our findings complement those of a recent EDUCAUSE study that reported students who used e-textbooks encountered difficulties with access and did not fully adopt the learning tools.
Accounting for Missing Data. In our process of examining and making sense of the data, we determined that we needed to properly account for the students in the e-textbook group who did not read or use the e-textbook. Almost all of the students in the research read the e-textbook, but we found, on average, that few students engaged in using the markup tools. That is, instead of using the interactive tools that were designed to boost student learning, most students opted only to use it as a substitute for a paper book. Although students received a brief orientation on how to use markup tools, most instructors did not integrate those tools (e.g., ability to share notes) into their course design. Because the analytics data did not include information about the many students who did not use the tools, we needed to add this information into our database. We found that, except for the highlight tool, less than half of the students used the markup tools (notes, annotations, questions, or bookmarks). We used the median of each markup tool because the usage does not have a normal distribution; rather, the distribution was right-skewed (indicating that a small number of students were enthusiastic adopters).

An even more important concern for educational administrators, however, is how to treat the number zero (0). Table 1 indicates that when we looked only at the vendor’s database of markup behavior, which did not include users who never used the tools, we saw an inflated estimate of usage. For example, as shown in the left side of table 1, when we include all of the 287 students in the study—all of whom had access to the e-textbook—the mean number of highlights per student was 93. The right side of the table, however, shows that just 153 students used the highlight tool (53% of the students in the study). Only counting the participants who actually used the highlighting tool, the mean number of highlights per
When studying educational technology systems that may not be adopted by all learners, integrating the “nonadopters” into the analysis is critical. Interestingly, because almost all students in the e-textbook group read the e-textbooks, the statistics about reading do not differ much when all students are included, whether they read the e-text or not. However, when factoring in all students, the median usage for the highlight tool decreased, and the medians for annotations, notes, and bookmarks fell to zero.

**Combining the Analytics and Self-Reported Usage**

We combined our analytics and survey data to create a database of e-textbook usage. Our weekly reading journals included questions about whether students had access to the e-textbook, how satisfied they were with that access, and how much time students spent reading. By combining the analytics and self-reported usage, we investigated whether the usage that students reported was recorded in the analytics.

**Using Analytics to Examine Data about the Reading of the E-Textbook.** Over the course of the semester, a growing number of students reported reading the e-textbook who did not have any record of reading in the analytics (see figure 2). This gap exposes a weakness in only using self-reported measures of usage of an educational technology system. Combining students’ reports about e-textbook access with vendor analytics can illuminate trends that may not be apparent with access to only one type of data.

Another finding concerned access to the e-textbook. At week 12, only 84% of the students in the e-textbook group reported having access to the e-textbook, despite the fact that all students had access to the e-textbook only through the LMS. These students who claimed not to have access may have been reading a paper version of the e-textbook or some other material because students in the e-textbook group had the option to purchase a paper version of the textbook.
We also examined the relationship between the amount of time spent reading (as measured by the e-textbook system) and students’ self-reported time spent reading. In our initial analysis, we assumed that a student who reported reading five times as long as another student would have read five times as many pages. But when we modeled the number of pages read during week 3 (a time when most students in the e-textbook group still indicated they had access to the e-textbook), we found a curvilinear relationship that is illustrated in figure 3. This suggests that it’s important for educators not to assume a perfect linear relationship between how much time students say they read and the number of pages recorded as being read in the analytics.

The figure suggests that a student who was engaged in trying to learn information on a rather limited number of pages may have reported spending a lot of time reading in a process that may have been inefficient. Indeed, one recent study of e-textbook users found that their learning processes were more...
inefficient compared with users of paper textbooks. In our study we found that e-textbook analytics could not tell us about the quality of these reading experiences; to do that, we needed to gather more information from the reading journals. There, we found that students exhibited a range of reading behaviors, but the most frequent category was what we labeled “purposeful reading,” in which students “read” their e-textbooks to locate information that they needed for a homework assignment, quiz, or test. Other reading categories included rereading, skimming, skipping, and reading for understanding.

Importantly, our survey question about whether students had purchased the paper textbook also was a significant, negative predictor of the amount of reading recorded in the database. Indeed, this simple question from our first survey was helpful in illuminating the amount of e-textbook usage we observed during the semester because it helped determine who was less inclined to use an e-textbook. Once students bought a paper textbook, their usage of the e-textbook decreased.

What It Means to Higher Education

Our project to examine e-textbook analytics as well as self-reported data has important implications for the adoption of educational technology in general and learning analytics in particular. In this section we summarize the benefits and limits of learning analytics and provide recommendations on how to effectively use learning analytics in assessment of student learning.

Learning Analytics Cannot Always Indicate the Quality of Engagement

Administrators and faculty can use learning analytics more effectively when they understand the limits of such data. In our example, the e-textbook analytics indicated how much students accessed different pages for reading, but we found that there was not a linear relationship between time reported reading and the number of pages recorded as having been read. Researchers must go further to determine the quality of such reading behavior by using other methods (such as focused questions in reading journals) that can be used to illuminate the quality of such interactions.

We should not eschew different research methods that put the usage of educational technology in context. In our case, interviews with faculty members provided information about why they did not adopt the e-textbook markup tools or promote usage among students. In addition, controlling for random effects due to students being clustered in classes helped solidify the quantitative findings and make them more reliable. Instructors, we found, had primarily adopted the e-textbook for their students to use as a free substitute for a paper textbook. Thus, cost was a key motivating factor. Qualitative research methods can provide insights about quantitative findings and can also be translated into programming for faculty.

Can the Results of Analytics Research Be Translated into Practice?

Perhaps one of the most critical areas of learning analytics is the translation of research into practice. We must be wary of a research methodology that risks becoming an obsession with inputs rather than a more rigorous examination of how usage of educational technology is connected to the outputs of education. In a follow-up study, we collected more data about e-textbook usage in the context of the individual courses to determine whether usage was more frequent in the week prior to the due date of a major assignment. Our initial results suggest that students still made minimal use of the interactive tools and that the use of
those tools was not scaffolded by instructors, who had often assumed that their students would pick up those markup tools on their own.

Learning analytics can be used to develop early-warning systems or other interventions designed to implement at the beginning of the semester. We want students to effectively use the educational technology that is part of their courses. Learning analytics can be informative about the obstacles that students face to such effective adoption and can be combined with other research methods to identify successful interventions. In our case, learning analytics gave us information about who was not using the technology of note-taking, and the journals gave us more information about why they may not have been using it (e.g., they were using a paper textbook). We contend that learning analytics can provide some answers to educational technologists, but additional data may be necessary to take effective action.

For learning analytics research to be possible and robust, we need to promote the effective adoption of tools. Our assumption that students would enthusiastically adopt the markup tools on their own was not valid, and only later in the study did we learn that most faculty members only wanted a substitute for the paper textbook. Educational technologists and instructional designers should engage faculty in the instructional design process to determine how these learning tools become integrated into course activities. This includes a process of extensive user testing to determine how the tools work and what problems students may face. We must avoid the pitfall of assuming that as "digital natives," undergraduate students are able (and willing) to automatically adopt any technology for rich learning on their own. This assumption has been called into question by researchers, and educators are likely to have better outcomes if they scaffold student adoption of learning technologies.

Learning Analytics Requires Effective Partnerships

Learning analytics researchers must harness a variety of institutional data to effectively model student outcomes and create interventions designed to assist students in an instructional context. In many ways, an educational technologist or educational assessment specialist works within the sphere of a course—the instructors and students—to examine the effectiveness of some use of educational technology. But learning analytics researchers must develop a working knowledge of the institutional data that are available to use for assessment or research purposes by collaborating with the staff who manage these data. Measures of prior learning and demographic characteristics are important variables to integrate into a study of learning, and institutional data stewards can provide important insights into the usage of educational technology.

Mapping the Landscape of Available Data

Do you know which systems your instructors are using and the data that these systems collect? With an ever-increasing amount of software and number of mobile applications that instructors may select from, keeping track of this information can be a daunting task. Educational technologists and administrators can engage campus partners with cataloging the programs, the student data they collect, and the processes and procedures for obtaining such data for learning analytics projects. Several research groups have put together effective models and services (such as DataShop) for collecting, storing, analyzing, and sharing learning analytics data. Some instructors may negotiate access to educational software for their students but not access to important data about student usage. Not implementing an effective process may result in missed opportunities for understanding student behavior and seeking ways to improve how students use
educational technologies. We know that educational technology has the best impact on students when it is used to promote cognition and deeper learning, so it is incumbent on us to facilitate that kind of usage.

**Become Skeptical of Data to Avoid Costly Misunderstandings or Analysis Built on Faulty Assumptions**

One critical takeaway from our work has been a healthy skepticism of large data sets of student usage of educational technology. Administrators and educational technologists must accept that these data sets cannot—of and by themselves—answer the important questions, for it takes significant time and labor to prepare the data for proper analysis. It is vital for researchers to take the time to interrogate large data sets so that they understand the variables and the structure of the data. In our case, it meant asking the vendor basic questions, such as how long a student needed to be on a page in the e-textbook before it would be counted as read.

**Key Questions to Ask**

- What questions about students' or instructors' successful adoption of educational technology can be answered by studying learning analytics? (And given that there are questions, how can the outcomes be translated into practices or principles that affect teaching and learning at your institution?)
- How can you effectively negotiate with educational technology vendors about obtaining access to student-level data for the purposes of better understanding how students and instructors use their products?
- Can you have conversations early with your institutional research staff and IRB about the proper ways to gather and use analytics for internal purposes or research?
- What kinds of expertise (educational research, statistics, and computer programming) can you leverage on your own campus to conduct effective work with learning analytics?
- When engaging faculty on the adoption of instructional interventions grounded in learning analytics, how can you make connections between student learning and the goals instructors have for their students?

**Where to Learn More**

- [Student Learning and Analytics at Michigan (SLAM)](https://www.its.umich.edu/academic/learning-analytics)
- [PAR Framework](https://www.parc坊.org)
- [Society for Learning Analytics Research](https://www.slr-site.org)
- [The Simon Initiative at Carnegie Mellon at University](https://www.simonsite.org)
- [DataShop at the Pittsburgh Science of Learning Center](https://www.pslc.org)

**About the Authors**

At the University of Iowa, Sam Van Horne holds a joint appointment in assessment research in the ITS Office of Teaching, Learning, & Technology and in the Office of Assessment; Jae-eun Russell is an instructional designer in the ITS Office of Teaching, Learning, & Technology; and Kathy L. Schuh is associate professor in the Department of Psychological and Quantitative Foundations in the College of Education.
Acknowledgments

The authors would like to acknowledge the generous support of the Consortium of College and University Media Centers, who provided a grant for our study of e-textbook usage.

Citation for This Work


Notes

6. A more detailed explanation of the results of the research study is found in Van Horne, Schuh, Russell, in progress.