Evaluating How the CourseSmart Engagement Index™ Predicts Student Course Outcomes

Reynol Junco
Associate Professor of Library Science
Purdue University
Fellow
Berkman Center for Internet & Society
Harvard University

with
Candrianna Clem
Graduate Student
Department of Sociology
Purdue University
The CourseSmart Promise: Greater Engagement Leads to Student Success

Educators have historically relied upon data such as attendance and class participation to identify at risk students who may be struggling. While observation and personal interaction with students are important aspects of ensuring learning outcomes, there is great demand in the higher education community for data that measure student engagement with digital course materials as a means to facilitate student success.

CourseSmart Analytics was developed specifically to address some of the most compelling challenges in higher education: improving retention, controlling costs, and improving learning outcomes. CourseSmart Analytics assesses a variety of usage statistics including session length within eTextbooks, pages viewed, and activities such as highlighting or taking notes, to provide meaningful metrics of student engagement with digital course materials.

A key pillar of the analytics solution is the CourseSmart Engagement Index™, an algorithm devised from a combination of third-party research and CourseSmart’s own proprietary formula, developed to more meaningfully measure engagement with course materials. All of the usage statistics, along with the CourseSmart Engagement Index™, are packaged into a user-friendly dashboard that can be accessed directly from learning and course management systems (LCMS). CourseSmart Analytics is designed to benefit a range of stakeholders:

**Faculty** are able to view aggregate and individual student engagement data, correlate that data to overall student performance, and use it as a means to intervene with “at risk” students to help them stay on track in their studies so they can graduate on time and on budget.

**Provosts, deans, and course designers** are able to use CourseSmart Analytics to assess the performance of adopted digital titles and ensure they are being used effectively and for optimum return on investment.
Publishers are able to use CourseSmart Analytics to assess the relative impact of various digital course materials so they can continuously update and improve their products.

CourseSmart Analytics was developed utilizing the rich set of Learning Tools Interoperability data (LTI data), which is available through the integration of the core product directly into an institution’s learning management system or online portal. As such, faculty and administrators are able to capitalize on both the eTextbook and analytics dashboards’ services within their existing institutional workflow. CourseSmart invested significant time and resources to develop the data provisioning processes and multiple reporting dashboards on top of the GoodData BI platform. All of this is delivered seamlessly to institutions, saving them time, money and manpower within their IT departments.

This white paper was developed specifically to assess the efficacy of CourseSmart Analytics. Relevant literature and previous research findings were factored into the development of this paper, including:

- Student reading comprehension in print vs. digital formats.
- The correlation between specific student reading activities—such as time spent reading the textbook, taking notes, and highlighting—on student comprehension.
- Effects of student engagement with LCMS and the impact of analytics within LCMSs.
- The efficacy of digital textbook analytics—specifically, the CourseSmart Engagement Index™—as an effective method of formative assessment.
The more students engaged with their digital textbook, the better their final course grade. The CourseSmart Engagement Index™ is as strong of a predictor of student academic outcomes as prior academic achievement.

**CourseSmart Analytics: From Pilot Program to Proven Effectiveness**

A pilot program for CourseSmart Analytics was announced in November 2012, comprising 76 faculty members, 26 administrators, and more than 3,700 students. This paper uses the *CourseSmart Engagement Index™* to evaluate how students read their digital textbooks and to evaluate how these behaviors can predict course outcomes.

**METHODS**

- This paper encompasses a subset of the original pilot program: 236 students from 11 courses who participated in CourseSmart’s pilot program.

- Students’ *CourseSmart Engagement Index™* data were evaluated against their final course grades. A linear regression analysis was conducted to determine whether the *CourseSmart Engagement Index™* predicted student final course grades.

- No effort was made to standardize grading across courses (i.e., professors chose how much of course grades relied upon textbook material).

**SIGNIFICANT RESEARCH FINDINGS**

- On average, students read 551 pages throughout the 16-week semester.

- They used the book an average of 11 days over the semester and engaged in an average of 17 reading sessions.

- The average cumulative total time spent using the digital textbook was 442 minutes.

- Students created on average 4 highlights, .42 bookmarks, and .16 notes.

- The *CourseSmart Engagement Index™* was a strong predictor of student course outcomes.

- The more students engaged with their digital textbook, the better their final course grade.

- The *CourseSmart Engagement Index™* is as strong of a predictor of student academic outcomes as prior academic achievement.
The CourseSmart Engagement Index™ is a transparent, unobtrusive, and effective barometer for student progress.

- The CourseSmart Engagement Index™ is a readily-available and easy-to-access indicator of student reading.
- Faculty can use the CourseSmart Engagement Index™ as an efficient method for formative assessment.
- Despite a student’s prior academic ability, with CourseSmart Analytics and the CourseSmart Engagement Index™, a course instructor can have an unobtrusive, real-time method to identify students at risk of academic failure that is not tied to student activity on a LCMS.

A thorough description of the literature and previous research impacting educational analytics and a detailed analysis of the CourseSmart Engagement Index™ is outlined on the following pages.
Such predictive modeling is the ultimate form of student formative assessment—educators can have information about how a student might fare in her/his courses even before the student takes an exam or a quiz.

LEARNING ANALYTICS IS THE COLLECTION AND ASSESSMENT OF STUDENT-PRODUCED DATA IN ORDER TO BEST TAILOR EDUCATION TO ENSURE STUDENT SUCCESS.

The 1st International Conference on Learning Analytics & Knowledge defined learning analytics as “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs.” Put another way, learning analytics is the use of large datasets collected in educational settings that allow for building predictive models of student success. Such methods may not involve direct input from the students but instead use information from technology such as learning and course management systems (LCMS). These data are collected and can be analyzed in real-time, giving educators the ability to identify students at risk of academic failure. Such predictive modeling is the ultimate form of student formative assessment—educators can have information about how a student might fare in her/his courses even before the student takes an exam or a quiz.

Presently, learning analytics tools and projects have collected limited data on student academic work. For instance, most learning analytics tools collect information solely from the LCMS and restrict the data collected to number of times students log on to the system, number of posts in online discussions, and assignment grades. Analytics platforms at some universities have begun to integrate additional data such as academic records from the student information system, student self-reported confidence level, and student involvement. However, the number of analytics platforms that collect additional data is limited, which also serves to restrict the predictive ability of their models. Data based on student technology use are not difficult to collect; however, many institutions either do not have access to these data or are too concerned about privacy issues to try to collect them.
Additional data can be collected in ways that protect student privacy, help improve predictive models, and improve educators’ ability to conduct early intervention with students whose data profiles suggest they will be at risk. One example of risk prediction is the Purdue Signals project, which mines data from the student information system and LCMS to flag at-risk students by posting a traffic signal indicator (i.e., red for high risk, yellow for moderate risk, and green for low risk) on the LCMS homepage. Based on the results of the system’s predictive student success algorithm, an intervention schedule is created consisting of either a traffic signal indicator on the student’s home page or an email or text messages referring the student to an academic advisor or to the instructor for face-to-face meetings. Such a system helps instructors become aware of students who might be at risk of failure and helps prioritize their limited time on the students who need help. While learning analytics applications to this point have focused on the LCMS, the growth of digital textbooks creates opportunities for educators to unobtrusively collect learning analytics data from students’ use of digital course materials. Textbook analytics can not only provide additional data to existing learning analytics models, but also can serve as stand-alone predictors of student success, since student textbook use should be predictive of course outcomes.
The growth of digital textbooks creates opportunities for educators to unobtrusively collect learning analytics data from students’ use of digital course materials.

READING ACTIVITY

Researchers have examined how students read in an attempt to uncover differences between digital and print reading. Digital reading does not affect reading comprehension, however, students who read digitally employ more reading strategies because digital text allows viewers to respond to others, collaborate on texts, and mark and highlight without damaging the book. Students who read digitally are more likely to use the strategies of: asking questions of the text, finding answers to these questions in the text, and guessing unknown words in context. The reading strategies that most successfully influenced comprehension were annotating (taking notes) and highlighting, rereading, matrix style note-taking and outlining; while the least successful was listing. However, highlighting is not always an effective strategy. While reading print materials, low-skill readers highlighted more text and more often than high-skill readers. Additionally, low-skill readers preferred to study previously marked material and were more likely to study only previously marked material than were high-skill readers. Although research has found variations in reading strategies and highlighting between digital and print reading, no differences have been found in GPA between electronic and print textbooks. This is important, as researchers have found that the act of reading a textbook is directly related to student success. Landrum et al. found that the percentage of reading completed in a textbook was positively correlated with quiz scores and final course grades. Students who read a chapter, regardless of whether it was in print or digital form, scored significantly better on a quiz of material contained in that chapter.
TIME AND READING

Research also shows that time spent reading course materials is a strong predictor of student success. On average, students spend fewer than three hours each week reading their print textbook, and they report using surface study strategies such as study guides more often than textbooks. In contrast, students spend more time reading electronic textbooks and PDFs than print textbooks. McEneany et al. tracked students’ navigation of hypertext links and found that reading rates varied over the course of each reading session. Earlier in the session, students spent more time on each page and used more complex navigation; later in the session, they used more linear navigation. Alternatively, Woody et al. found that there were no differences in the rates that students read section summaries and answered study questions. However, students spent more time reading at home than in controlled lab conditions. Furthermore, average reading time varied by the type of course the student was enrolled in. These differences may be explained by discipline-specific textbook usage patterns or the role, expectations, and involvement of the instructor.

Researchers have found that the act of reading a textbook is directly related to student success...the percentage of reading completed in a textbook was positively correlated with quiz scores and final course grades.
Progress has been made in the sophistication of digital textbooks, not just in their interactivity but also in their ability to collect data on how students are interacting with them.

Even though current implementations of learning analytics collect limited amounts of data, research has shown that analytics can help predict student outcomes. Specifically, research has shown a positive relationship between LCMS activity and course outcomes. Macfadyen and Dawson\(^6\) found that the number of discussion messages read as well as the number of posts and replies were significantly related to final course grade. A similar study by Smith, Lange, and Houston\(^24\) found that log-in frequency, site engagement (viewing the course syllabus, viewing an assessment, completing an assessment, etc.), and points submitted were all correlated with successful course outcome (i.e., whether students earned a final letter grade of “C” or higher).

Digital textbook analytics is a new sub-category of learning analytics being pioneered by CourseSmart. Beyond a university’s LCMS, progress has been made in the sophistication of digital textbooks, not just in their interactivity but also in their ability to collect data on how students are interacting with them. Because, as is noted earlier, there is a direct correlation between student reading and academic success, digital textbook analytics could also suggest how a student will perform in a course.\(^19\) Given previous research on reading activity, collecting data on how often students highlight and annotate could add to textbook analytics predictive models.\(^13\), \(^16\) Analytics measuring students’ engagement with their textbooks allow for the collection of data that can act as proxies for students’ reading skills and learning strategies. For instance, data on highlighting activity might effectively discriminate between high- and low-skilled readers.

CourseSmart has developed a learning analytics platform for digital textbooks that produces the CourseSmart Engagement Index™, a real-time indicator of how students are interacting with their textbook, including multiple indices of reading frequency (time spent reading) and interaction (such as highlighting and annotating).

Of particular importance to educators is the fact that digital textbook analytics can serve as an effective method of formative assessment. The Association for Middle Level Education (AMLE) defines formative assessment as part of the
educational process and when, “incorporated into classroom practice, it provides the information needed to adjust teaching and learning while they are happening.”\(^{25}\) Furthermore they state, “formative assessment helps [educators] determine next steps during the learning process as the instruction approaches the summative assessment of student learning.”\(^{25}\) Therefore, faculty might use digital textbook analytics to keep their “finger on the pulse” of how students are performing, or are about to perform, in their courses. Since digital textbook analytics are a new tool, there has been no research showing how these data can be used to predict student outcomes. The current study examines CourseSmart Analytics to answer the following questions relevant to digital textbook analytics:

1. **WHAT ARE THE DIGITAL READING PATTERNS OF STUDENTS?**

2. **HOW IS THE COURSESMART ENGAGEMENT INDEX™ RELATED TO ACADEMIC PERFORMANCE IN A COURSE?**

**DESCRIPTION OF THE COURSESMART ENGAGEMENT INDEX™**

The *CourseSmart Engagement Index™* is a calculated linear function that includes the following data: number of pages read, number of sessions (number of times a student opens/interacts with the digital textbook), number of days the student uses their textbook, session length (time spent reading), number of highlights, number of bookmarks, and number of notes. The *CourseSmart Engagement Index™* is a score that ranges from 20 to 80, with 20 representing no engagement and 80 representing the highest level of engagement. The *CourseSmart Engagement Index™* is intended to serve as a single number that captures how much the student is engaged with the reading. In practice, faculty can access the *CourseSmart Engagement Index™* through CourseSmart’s Analytics dashboard on their institution’s LCMS. Given previous research showing that students who read and interact with their book perform better in class, it is hypothesized that there will be a relationship between the *CourseSmart Engagement Index™* and course outcomes.\(^5,11,23\)
METHODS

This study was conducted at a medium-sized, two-year, upper-level, Hispanic Serving Institution in the south. The institution was asked to identify at least three courses with willing instructors to take part in the study. These courses were required to use a book that was offered by CourseSmart, use the campus LCMS, and have at least 50% of the students using the CourseSmart eTextbook. The resulting sample size was 236 student participants from 11 courses. The course disciplines were accounting, biology, business, criminal justice and management. Data were collected using the CourseSmart Analytics eTextbook platform, which unobtrusively tracked student interaction with the digital textbook throughout the Spring 2013 semester. There was no effort made to standardize grading across courses—in other words, professors chose how heavily the course grades were impacted by textbook material. Students were aware that the eTextbooks tracked their usage. Engagement data were downloaded directly from the CourseSmart servers, checked for anomalies, coded, and analyzed using SPSS Statistics 21.

To answer the second research question, a blocked linear regression analysis was conducted to determine whether the CourseSmart Engagement Index™ could predict student final course gradesa. Using a blocked linear regression allows the researcher to insert variables into the model by grouping them based on a theoretical construct. The blocks were gender, race/ethnicity, and transfer GPA. Gender and race/ethnicity are demographic variables that strongly correlate to student academic achievement26. Transfer GPA was included as both a control variable and in order to compare other variables’ relative impact on student course grades.

aThe data were analyzed to evaluate whether they met the assumptions necessary for a regression analyses. Collinearity and evaluation of outliers was examined through collinearity diagnostics and by examination of residual plots. All tolerances were above .84 and all VIF’s were below 1.19. The curve estimation procedure of SPSS was used to examine the linearity of the relationships in the data. Functions were plotted for both linear and quadratic functions and examination of these functions revealed that linear functions were the appropriate fit for the data. Therefore, the curve estimation, collinearity, and outlier analyses showed that the assumptions for regression were met. However, the analyses also indicated there were three outliers who had Engagement Scores of over 4,900. These outliers were removed from further analyses reducing the sample size to 233 students. Categorical variables were dummy-coded for purposes of the regression analyses. The reference categories for these variables were female and White students.
RESULTS

**Research Question 1:** What are the digital reading patterns of students?

<table>
<thead>
<tr>
<th>Descriptive Statistics of Reading Behaviors</th>
<th>Mean (SD)</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pages Read</td>
<td>551 (782)</td>
<td>224</td>
</tr>
<tr>
<td>Number of Days</td>
<td>11 (10)</td>
<td>7</td>
</tr>
<tr>
<td>Number of Sessions</td>
<td>17 (20)</td>
<td>10</td>
</tr>
<tr>
<td>Session Length (min)</td>
<td>442 (664)</td>
<td>169</td>
</tr>
<tr>
<td>Highlights</td>
<td>.42 (2)</td>
<td>0</td>
</tr>
<tr>
<td>Bookmarks</td>
<td>.16 (1.2)</td>
<td>0</td>
</tr>
<tr>
<td>Notes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Research Question 2:** How is the CourseSmart Engagement Index™ related to academic performance in a course?

The predictive model using the CourseSmart Engagement Index™ was statistically significant, showing that the Index is a strong predictor of academic performance in a course.

The CourseSmart Engagement Index™ score was positively predictive of student course grade (see Figure 1). Additionally, confirming previous research that has shown that student past academic achievement is predictive of future achievement, transfer GPA was positively predictive of course grade. An examination of the two significant predictors of course grade shows the relative strength of the CourseSmart Engagement Index™ in predicting student academic outcomes: the β coefficient for the Engagement Index™ score was .192, stronger than the transfer GPA coefficient of .183, which shows that the CourseSmart Engagement Index™ is a stronger predictor of student success than prior academic achievement.

**Descriptive Statistics**

Sixty-four percent of students in the sample were female and 36% were male. The mean age of the participants was 32 with a standard deviation (SD) of 9. The age of participants ranged from 20 to 66, though 82% were between 20 and 39 years old. In terms of race and ethnicity, 65% of students identified as white, 59% as Latino/Hispanic, 8% as African American, 4% as Asian American, 3% as Native American, and 13% as "Other." The average transfer GPA in the sample was 3.02 (SD = .51) and the average course grade earned during the study was 2.84 (SD = 1.50), which equates to roughly a B-. 
DISCUSSION

The results from this study show that the CourseSmart Engagement Index™ serves as a transparent, unobtrusive, and effective barometer for student progress. The CourseSmart Engagement Index™ score was found to be a strong predictor of student final course grade, even when taking preexisting academic achievement into account. Therefore, the predictive ability of the CourseSmart Engagement Index™ cannot be explained by differences in student academic ability. In fact, the data from these analyses show that the CourseSmart Engagement Index™ is a stronger predictor than previous academic achievement, which has been shown to be the strongest single predictor of student success.27,28,29 While it is no surprise, given previous research, that reading and engaging with the textbook are predictive of student course performance, the Engagement Index™ is a readily-available and easy-to-access indicator of student reading.11 Consequently, the CourseSmart Engagement Index™ is an especially efficient method for formative assessment—providing the information needed to adjust teaching and learning by evaluating how students are engaging with their textbooks. With access to student-produced textbook engagement data, faculty can track student engagement at any given time in the course, even well before students take an exam or submit other gradable material. As such, the CourseSmart Engagement Index™ is an effective early warning system to identify students at risk of academic failure. It is worth noting that a faculty member need not be the person intervening to help the student: equipped with data from the CourseSmart Engagement Index™, they can make the appropriate referral to campus resources that may best help the student. From a curricular perspective, the CourseSmart Engagement Index™ can help faculty plan, adjust, and assess their teaching strategies in order to maximize student engagement with the text. This in turn will lead to improved outcomes because, as the research shows, further engagement is a predictor of ultimate student success.
The data from these analyses show that the CourseSmart Engagement Index™ is a stronger predictor than previous academic achievement, which has been shown to be the strongest single predictor of student success.

Even though previous research shows that reading patterns may differ based on type of course as well as the involvement of the instructor, this was not evident when using the CourseSmart Engagement Index™. On the contrary, the CourseSmart Engagement Index™ was an effective predictor of outcomes across all levels of readers, academic disciplines, and teaching styles, suggesting that these factors don’t play a role in the effectiveness of the Index. The CourseSmart Engagement Index™ was also an effective predictor of course grades even when grading was not standardized across courses (i.e., how much of the course grade relied upon textbook material) showing its flexibility in predicting learning outcomes with different course reading requirements.

Unlike previous research, this study finds that students did not use the highlighting feature in their digital textbook frequently. In fact, most students did not highlight, take notes, or use bookmarks in their digital textbooks. An additional analysis found that number of highlights was not predictive of course grades; however, further examination of these data showed that students who were in the top tenth percentile of number of highlights had significantly lower course grades than those in the lower ninetieth percentile. Although there were not many students in the top tenth percentile and this group produced few highlights. Therefore, further research will be needed to examine this issue with a sample of students more inclined to highlight. It might also be the case that future research will find that in practice and across different educational settings, students don’t highlight as much in their digital textbooks.

⁵ Students who were in the top 10th percentile of number of highlights had significantly lower course grades (M = 2.69, SD = 1.55) than students in the lower 90th percentile (M = 3.52, SD = 1.23; t(231) = -2.594, p < .01). Students in the top 10th percentile highlighted two or more times during the semester while those in the lower 90th percentile highlighted up to one time.
The CourseSmart Engagement Index™ is an especially efficient method for formative assessment—providing the information needed to adjust teaching and learning by evaluating how students are engaging with their textbooks.

CONCLUSION

The CourseSmart Engagement Index™ represents a significant leap forward in terms of student evaluation. Using a single straightforward number, faculty can easily assess how well a student is performing in their course at any given time. In practice, a faculty member can scan the CourseSmart Engagement Index™ of their students, discover how much students are reading, and identify students who may need additional help. No matter what a student’s prior academic ability, which may not be specifically known, the course instructor can have an unobtrusive, real-time method to identify students at risk of academic failure that is not tied to activity on a LCMS.

Adding data from the CourseSmart Engagement Index™ to learning analytics tools based on an institution’s LCMS should provide even greater accuracy in identifying students at risk of academic failure; however, this is not necessary in order to efficiently focus on those students most at need. As a stand-alone measure, the CourseSmart Engagement Index™ can help faculty plan their courses and conduct more efficient formative assessments of their students as compared to more traditional forms of observable data such as class participation, performance on quizzes, or using analytics data from the LCMS.
FIGURE 1

Relationship between, CourseSmart Engagement Index™, transfer GPA, and course grade. Note. Coefficients listed are the Beta (β) standardized regression coefficients.

![Diagram of relationships between Course Smart Engagement Index, transfer GPA, and course grade with coefficients and p-values]
REFERENCES


