Internet and Higher Education 27 (2015) 54–63

Contents lists available at ScienceDirect

Internet and Higher Education

Predicting course outcomes with digital textbook usage data

Reynol Junco [a,b,*], Candrianna Clem [c]

[a] School of Education, Human Computer Interaction, Iowa State University, 2628 Lagomarcino Hall, Ames, IA 50011, United States
[b] Berkman Center for Internet and Society, Harvard University, 23 Everett Street, 2nd Floor, Cambridge, MA 02138, United States
[c] Department of Sociology, The University of Texas at Austin, Austin, TX 78712, United States

ARTICLE INFO

Article history:
Accepted 5 June 2015
Available online 11 June 2015

ABSTRACT

Digital textbook analytics are a new method of collecting student-generated data in order to build predictive models of student success. Previous research using self-report or laboratory measures of reading show that engagement with the textbook was related to student learning outcomes. We hypothesized that an engagement index based on digital textbook usage data would predict student course grades. Linear regression analyses were conducted using data from 233 students to determine whether digital textbook usage metrics predicted final course grades. A calculated linear index of textbook usage metrics was significantly predictive of final course grades and was a stronger predictor of course outcomes than previous academic achievement. However, time spent reading, one of the variables that make up the index was more strongly predictive of course outcomes. Additionally, students who were in the top 10th percentile in number of highlights had significantly higher course grades than those in the lower 90th percentile. These findings suggest that digital textbook analytics are an effective early warning system to identify students at risk of academic failure. These data can be collected unobtrusively and automatically and provide stronger prediction of outcomes than prior academic achievement (which to this point has been the single strongest predictor of student success).

1. Introduction

Learning analytics is 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" (Siemens, 2010). Learning analytics projects collect and assess student-produced data in order to predict educational outcomes with the goal of tailoring education. In other words, learning analytics is an application of big data and predictive analytics in educational settings (Ferguson, 2012; Manyika et al., 2011). These methods do not involve direct input from students but instead use data collected unobtrusively from their use of educational technologies such as a university's learning and course management systems (LCMS) (Mattingly, Rice & Berge, 2012; Verbert, Manouselis, Drachsler & Duval, 2012). Data are collected and analyzed in real-time giving educators the ability to identify students at risk of academic failure (Picciano, 2012). Such predictive modeling is the ultimate form of student formative assessment—educators can have information about how a student might fare in their courses even before the student submits gradable work.

Campbell and Oblinger (2007) identified five steps of the learning analytics process:

1. *Capturing*: Data are captured from real-time learning analytics systems and combined with student information, such as demographic or background information (Ferguson, 2012). Specifically, data are collected in real-time from virtual learning environments (VLEs), learning management systems (LMSs), virtual machines, personal learning environment (PLEs), web portals, intelligent tutoring systems (ITS), forums, chat rooms and email (Mattingly, Rice & Berge, 2012).
2. *Reporting*: Data collected in the first step is reported to key individuals, identifying trends and patterns in the data to generate accurate models for measuring student progress and success. In order for the data to be reported in a coherent manner, it must be translated from raw data to comprehensive information through theoretical constructs, algorithms, and weightings (Greller & Drachsler, 2012). Oftentimes, the transformed data are visualized using learning analytics dashboards so that they can be more easily understood.
3. *Predicting*: A key affordance of learning analytics is using data to identify predictors of student success and to create models to predict student outcomes and identify at-risk students. In addition, predictive modeling is used to examine real-time learning in the classroom, to predict and make decisions about course and resource allocation, and to inform institutional decision-making (Ferguson, 2012).
4. *Acting*: Instructors and students act on the information discovered through learning analytics, for instance by intervening when a student is not doing well.
5. *Refining*: Ideally, as information is gathered, processed, and evaluated, learning analytics are refined through a cyclical process and continuous improvement model (Ferguson, 2012).

* Corresponding author at: 2628 Lagomarcino Hall, Ames, IA 50011, United States.
E-mail address: rey.junco@gmail.com (R. Junco).

http://dx.doi.org/10.1016/j.iheduc.2015.06.001
1096-7516/

While learning analytics show great promise for education, presently these tools and projects collect limited data on student academic work. For instance, most learning analytics tools collect information solely from the LCMSs and restrict the data collected to number of times students log on to the system, number of posts in online discussions, and assignment grades (Dawson, McWilliam & Tan, 2008; MacFayden & Dawson, 2012). The Purdue Signals learning analytics project 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 (Arnold & Pistilli, 2012). Based on the results of the system's predictive student success algorithm, an intervention schedule is created consisting of automated email message or reminders, text messages, referral to an academic advisor, or a referral for face to face meetings with the instructor. Such a system helps instructors become aware of students who might be at risk of failure and helps focus their limited time on the students who need help.

Predictive models based on LCMS data show that activity on these sites is related to course outcomes. For instance, Arnold and Pistilli (2012) found that students in courses that used the Purdue Signals system earned better course grades and were retained at higher rates. Macfayden and Dawson (2012) found that number of discussion messages posted, read, and replied to positively predicted final course grades. Student use of LCMS-based course content and visits to the online grade book were also positively correlated with final course grade (Dawson, McWilliam & Tan, 2008). Smith, Lange, and Huston (2012) found that LCMS log-in frequency, site engagement (activities such as viewing the course syllabus, viewing an assessment, and completing an assessment), and points submitted were all correlated with successful course outcomes. Also, the more a student performs a certain course activity the better they will score in that area as course outcomes are predicted with increased accuracy as more activity and grade information accumulate over the duration of a course (Smith, Lange, & Huston, 2012).

Learning analytics applications to this point have focused on collecting data from LCMSs, although newer methods allow for collection of broader types of student-generated data (Junco, 2013a; Junco, 2013b). For instance, using monitoring software to collect data on student use of social media might predict how well they will perform in a course (Junco, 2013a; Junco 2013b). Furthermore, the predictive ability of models based solely on LCMS data is overestimated because they relate graded activities (like number of discussion board posts) to course grades. The relationship between discussion board activity and course grade should be significant because students are being graded on such activity.

1.1. Digital textbooks

Digital books continue to be a fast-growing sector of the publishing market. The Association of American Publisher's (2014) data on trade books found that there was an increase of 43% in ebook sales between 2011 and 2013. Youth digital book usage has grown substantially since 2011. Nielsen (2014) reports that while only 13% of children less than 12 years old read ebooks on a tablet in 2011, 34% did so in 2014. These numbers have increased since 2011 with Pew Internet Project data showing that more adults and youth own ereading-capable devices (such as tablets and dedicated ereaders) and read ebooks than ever before (Rainie, 2012; Zickuhr & Rainie, 2014).

While adoption of digital textbooks in higher education is growing, there are mixed findings when looking at student preferences. Longhurst (2003) found that 64% of students preferred printed materials to online texts, and that almost all students printed out materials when they had the option. Among both graduate and undergraduate students, 92% of students printed out materials when working with someone else, over 80% of students printed out materials if they were long or complicated or they wanted to study from them and 75% of students printed out materials if they wanted to take notes (Sandberg, 2011; Spencer, 2006). The reason students give for choosing print material over online text is because of difficulty reading from a screen (Rockinson-Szapkiw, Courduff, Carter & Bennett, 2013), that it is easier to concentrate when using paper than on a screen, and that it is easier to remember and understand information with paper than with online text (Le Bigot & Rouet, 2007). Dennis (2011), on the other hand, found that 36% to 84% of students preferred a digital textbook depending on the course; however, the strongest factor predicting preference was whether the instructor made substantial use of the digital text (assigning readings, making annotations, and referring to the book in class), instead of solely using it as a reference. Previous experience using a digital textbook for class also influenced student preference. In a related study, Weisberg (2011) provided students with tablets, e-readers, and print textbooks. If given a choice, 87% of students said they would use a digital textbook. Students reported that lower cost, convenience, and the ability to easily search book content were positive factors for adoption; while reporting that it was easier to concentrate on and comprehend paper textbooks as negative factors for adoption (Weisberg, 2011).

Capitalizing on the growth of ereading-capable devices and the ubiquitousness of learning and course management systems, textbook publishers have invested a great deal of resources into developing and promoting digital textbooks (Young, 2013). Newer versions of digital textbooks are intended to serve not just as texts for the course, but also as primary sources of course content. Indeed, some have argued that the publishers may soon control not just the textbook material but the course content as well (Young, 2013). In their push to increase the interactivity and usefulness of digital textbooks, publishers have included interactive content such as dynamic quizzes that feed results back into LCMS grade books (Young, 2013). Such integration will become more commonplace as textbook companies have moved to acquire educational technology companies, like those that develop LCMSs. In addition to acquiring companies that develop platforms like LCMSs, textbook companies are acquiring adaptive learning and learning analytics startups. For example, in 2013 *Pearson* acquired *Learning Catalytics*, a company that uses predictive analytics to help provide feedback to faculty and improve student engagement (New, 2013).

The merger of textbook companies with LCMS, adaptive learning, and learning analytics products hints at the future of digital textbooks. Indeed, companies have developed textbooks that are not only intended to help students become more engaged, but that can track that engagement much in the same way that LCMSs use learning analytics to track and predict student success. The advent of digital textbooks, then, affords educators the opportunity to unobtrusively collect learning analytics data from student use of reading materials. While digital textbooks have had the ability to collect usage data, only recently have technology companies started to develop methods to use these data to predict student outcomes (Bossaller & Kammer, 2014).

Textbook analytics is an emerging sub-category of learning analytics and such applications can not only provide additional data to existing learning analytics models but can serve as stand-alone predictors of student success, for engagement with digital textbooks should be predictive of course outcomes. For instance, computer science students who completed more exercises in a digital text scored better on written exams in the course (Fouh, Breakiron, Hamouda, Farghally, & Shaffer, 2014). Fouh et al.'s (2014) data show that while students did not read the static text, they did engage with the digital text's interactive elements, even when such engagement was ungraded.

1.2. Reading and student performance

Unsurprisingly, reading a textbook is directly related to course outcomes (Daniel & Woody, 2013). For instance, Landrum, Gurung, and Spann (2012) found that the self-reported percentage of reading completed in a textbook was positively correlated with quiz scores, total course points, and final course grades. On average, students spend less than 3 h a week reading their print textbook, and reported using surface study strategies like study guides more often than textbooks (Berry, Cook, Hill & Stevens, 2010).

Given the rise of digital textbooks, researchers have been curious about differences between digital and print reading. Woody et al. (2010) found that there were no differences in the rates that students read sections or answered study questions and were more likely to use embedded links in e-textbooks than online resources in print books while Schugar, Schugar, and Penny (2011) saw no differences in reading comprehension between students who chose to read course readings using an e-reader versus those who didn't. Rockinson-Szapkiw et al. (2013) found there were no differences in perceived learning and grades between students who chose to use a digital version of their course textbook. Weisberg (2011) and Taylor (2011) both found no differences in quiz scores between students who read digital versus paper textbooks. Taylor (2011) uncovered no differences in learning between students who annotated the text and those who didn't. However, Dennis (2011) found that while number of pages students read in a digital textbook was not related to course grades, number of annotations was positively related. Daniel and Woody (2013) discovered that irrespective of format, students who read a chapter had significantly better quiz scores than those who did not.

Observed and self-reported reading times were not related to self-reported GPA, although students spent significantly more time reading digital material than print in one study (Daniel & Woody, 2013). Reading time varies by location as students spend more time reading at home than in controlled lab conditions (Daniel & Woody, 2013). Engagement with the textbook varies by discipline with average reading time being higher for more advanced courses (Fitzpatrick & McConnell, 2008). Possible explanations for these differences could be how textbooks are used differently across disciplines and the role, expectations, and involvement of the instructor.

Variations in highlighting and other reading strategies have been found between digital and print reading although no differences have been found in GPA between electronic and print textbooks (Daniel & Woody, 2013). Research has found differences in reading skills based on how much students highlight with low-skilled readers highlighting more than high-skill readers (Bell & Limber, 2009). Highlighting should be positively related to reading skills up to a point—when students are unable to discriminate important from trivial textbook information, they may highlight excessively. Those who read in print used fewer reading strategies than those who read digitally (O'Brien & Voss, 2011; Uso-Juan & Ruiz-Madrid, 2009). The reading strategies that most successfully influenced comprehension were annotating (taking notes) and highlighting (Yang, 2010), rereading (Hsieh & Dwyer, 2009), matrix style note-taking and outlining; while the least successful was listing (Kauffman, Zhao & Yang, 2011).

1.3. Student engagement

According to Astin (1984), engagement is “the amount of physical and psychological energy that the student devotes to the academic experience” (p. 518). His theory of student engagement includes five tenets: (1) engagement involves investment of physical and psychological energy; (2) engagement occurs along a continuum (some students are more engaged than others, and individual students are engaged in different activities at differing levels); (3) engagement has both quantitative and qualitative features; (4) the amount of student learning and development associated with an educational program is directly related to the quality and quantity of student engagement in that program; and (5) the effectiveness of any educational practice is directly related to the ability of that practice to increase student engagement. Today, engagement is conceptualized as the time and effort students devote to educational activities that are empirically linked to desired college outcomes (Kuh, 2009).

Chickering and Gamson (1987) proposed seven principles for good practice in undergraduate education, all of which are related to student engagement. The seven principles are: (1) student/faculty contact; (2) cooperation among students; (3) active learning; (4) prompt feedback; (5) emphasizing time on task; (6) communicating high expectations; and (7) respecting diversity. Kuh (2009) reported that institutions of higher education could directly influence engagement by implementing these seven principles. Chickering and Ehrmann (1996) showed how technology might be used to apply the seven principles. Student engagement has been researched extensively since Astin's (1984) original work. Researchers have found that engagement is directly related to student success (Pascarella & Terenzini, 2005) leading Kuh (2009) to report that “student engagement and its historical antecedents… are supported by decades of research showing positive associations with a range of desired outcomes of college” (p. 698).

Engagement is a necessary prerequisite for student learning. Student learning, however, is related to a host of factors including students' motivation, regulation of learning, and views of learning (Lonka, Olkinuora, & Mäkinen, 2004). The relationship between student engagement with their textbook and academic performance can be understood in the context of Astin's (1984) tenets that engagement requires the investment of psychological energy and that the amount of learning is related to the amount of engagement in an educational activity. Students who invest more psychological energy in and are engaged in reading are more likely to perform better in a course. Furthermore, increasing time on task and promoting active learning should lead to improved learning outcomes (Chickering & Gamson, 1987). Studies have shown that while most students don't read text elements, most do engage with interactive features of digital texts; engaging with these interactive features leads to better course outcomes (Berry et al., 2010; Dennis, 2011; Fouh et al., 2014). In general, engagement with the textbook is related to better course outcomes (Daniel & Woody, 2013; Landrum et al., 2012); however, format is not important leading Taylor (2011) to summarize that getting students to read in the first place is of utmost importance.

Astin's (1984) theory of student engagement and Chickering and Gamson's (1987) seven principles can be used to predict outcomes from the use of digital textbooks. Students who are more engaged with digital textbooks, should perform better in their courses. Moreover, additional engagement with digital textbooks such as highlighting or taking notes should lead to improved course outcomes. Therefore, it should be possible to use textbook analytics to directly measure engagement with a digital textbook and to use this information to predict student course outcomes (Bossaller & Kammer, 2014).

1.4. Research questions

Digital textbooks afford instructors and researchers the opportunity to unobtrusively collect data on how students are interacting with course materials. Previous research on reading and student success and on digital textbook use and student success has relied primarily on self-reported measures of reading and interactions with digital textbooks. Knowing that student reading is directly related to academic success, digital textbook analytics could suggest how a student will perform in a course (Berry et al., 2010; Fouh et al., 2014). 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 (Daniel & Woody, 2013; O'Brien & Voss, 2011).

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” (Garrison &

Ehringhaus, 2009). Furthermore they state, “Formative assessment helps [educators] determine next steps during the learning process as the instruction approaches the summative assessment of student learning” (Williford, 2009). Therefore, faculty can use digital textbook analytics to have their “finger on the pulse” of how students are performing, or are about to perform, in their course.

CourseSmart provides a large percentage of the most popular textbooks in digital form and they are the first digital textbook provider to offer textbook analytics. The CourseSmart analytics platform was developed to address Campbell and Oblinger's (2007) five steps of the learning analytics process. First, the analytics platform *captures* data on interactions with the digital textbook in real-time. Second, the platform translates the raw data into a calculated Engagement Index and *reports* this information to faculty. However, the *predictive* ability of the Engagement Index has not yet been tested. Because the *predictive* ability of the Index has not yet been tested, it is unclear as to how instructors and students can *act* on the data and what needs to be *refined* in order to improve the Engagement Index.

The Engagement Index was developed as a real-time indicator of how students are interacting with their textbook. Through initial beta testing and with guidance from Astin's (1984) theory of student engagement, the Engagement Index includes multiple measures of reading frequency (time spent reading) and engagement (like highlighting and annotating). The Engagement Index score 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. Since the Engagement Index is a corporate for-profit product developed by CourseSmart, the exact formula for calculation of the Index is not shared with researchers.

Because 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 digital textbook reading patterns and evaluates how well the Engagement Index and its composite elements can predict student course outcomes. The study is also intended to inform ways that instructors can act on data collected by digital textbooks and to provide recommendations to refine the Engagement Index and/or other textbook analytics predictive calculations (Campbell & Oblinger, 2007). 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 positive relationship between the Engagement Index and course outcomes (DeBerard, Speilmans & Julka, 2004; Dawson, McWilliam & Tan, 2008; Woody et al., 2010). Furthermore, based on previous research it is hypothesized that there will be differences in course outcomes by the number of highlights (Daniel & Woody, 2013), but the number of notes, bookmarks, and time spent reading may not necessarily reflect a difference in academic performance (Fitzpatrick & McConnell, 2008, Hsieh & Dwyer, 2009).

The current study uses CourseSmart digital textbook analytics to answer the following questions:

1. *What are the digital reading patterns of students?*
2. *How is the CourseSmart Engagement Index related to course performance?*
3. *How are the individual components of the CourseSmart Engagement Index related to course performance?*

2. Methods

2.1. Ethics statement

The study was categorized as exempt by the Institutional Review Board policies of Texas A&M University—San Antonio because it involved “research conducted in established or commonly accepted educational settings, involving normal educational practices.” All participants gave written informed consent to participate in the study through an online system.

2.2. Participants

Texas A&M University—San Antonio is a medium-sized two year upper-level only Hispanic Serving Institution. The university was asked to identify courses with willing instructors to take part in the study that used a book that was offered by CourseSmart, used the campus LCMS, and where at least 50% of the students used the CourseSmart eTextbook. The institution identified 11 courses and 8 faculty who participated. Table 1 lists the course titles, the distribution of instructors, and the distribution of students in each course after deletion of outliers (see *Statistical Analyses* section below for an explanation of the handling of outliers). The digital textbooks were provided to the 307 students enrolled in these courses; however, students were free to print the book instead of using the digital textbook. While all students consented to take part in the study, only 236 students used the digital textbook for an overall participation rate of 77%. Examination of the textbook usage data suggests that these 236 students used the digital textbook exclusively.

2.3. Measures

Data were collected using the CourseSmart eTextbook analytics platform, which unobtrusively tracked student interaction with the digital textbook throughout the Spring 2013 semester. The analytics platform yielded an Engagement Index score that ranged from 20 to 80, with 20 representing no engagement and 80 representing the highest level of engagement. The platform also yielded: number of pages read, the number of times a student opened/interacted with the digital textbook, number of days the student used their textbook, time spent reading, number of highlights, number of bookmarks, and number of notes. Additionally, transfer GPA, course grades, and student demographic information were obtained directly from university records. Textbook use data were downloaded directly from the CourseSmart servers, combined with the data from university records, checked for anomalies, coded, and analyzed using SPSS Statistics 21.

2.4. Statistical analyses

To answer the first research question, descriptive statistics were used to illustrate how students interacted with their digital textbooks over the course of an entire 16-week semester. To answer the second and third research questions, blocked linear regression analyses were conducted to determine whether the Engagement Index and its components could predict student course grades. Because of the significant and strong correlations (Pearson's $r > .87$) between the time components (number of pages read, the number of times a student opened/interacted with the digital textbook, number of days the student used

Table 1
Distribution of students by course and instructor.

Instructor	Course	Students
1	Introduction to Accounting S1	31
	Introduction to Accounting S2	26
2	Evolutionary Theory	36
3	American Judicial Process S1	17
	American Judicial Process S2	9
4	Emotional Disorders and Related Issues	12
5	Educational & Psychological Measurement & Evaluation	9
6	Assessment of Exceptional Students	25
7	Human Resource Management	33
8	Production/Operations Management S1	28
	Production/Operations Management S1	7
	Total	233

their textbook, time spent reading), the *number of days* variable was used as the overall proxy for time in the Engagement Index components regression. Using a blocked linear regression allows the researcher to insert variables into the model in groups also called "blocks." The blocks were gender, race/ethnicity, transfer GPA, course, and either Engagement Index score or component values. Race and ethnicity categories were those collected by the institution and match the US Census/Office of Management and Budget (OMB) standards. Gender and race/ethnicity are demographic variables that strongly correlate to student academic achievement and were used as control variables in these analyses (Pascarella & Terenzini, 2005). Transfer GPA was included as both a control variable and in order to compare other variables' relative impact on student course grades.

Research has consistently found that previous GPA is the strongest predictor of future academic achievement (Burton & Dowling, 2005; DeBerard, Speilmans & Julka, 2004; Dennis, 2011; Geiser & Santelices, 2007; Williford, 2009). Lastly, course was included to control for the effects of the difficulty level of each course and for instructor effects. Overall, research has shown the effect of instructor and academic discipline on student academic performance (Cohen, 1981; Vanderstoep, Pintrich, & Fagerlin, 1996; Vermunt, 2005). Typically students taking courses from instructors who are rated more highly have higher levels of academic achievement (Cohen, 1981). Additionally, engagement with the textbook varies by discipline and course level (Fitzpatrick & McConnell, 2008) and students from different academic disciplines have different levels of reading comprehension (Alderson & Urquhart, 1985). Because of the strong correlation between instructor and course variables, only course variable was used in the analyses as it captures variance from instructor, discipline, and course section.

The data were analyzed to evaluate whether they met the assumptions necessary for regression analyses. Collinearity and evaluation of outliers was examined through collinearity diagnostics and by examination of residual plots. All tolerances were above .20 and all VIF's were below 5. The curve estimation procedure of SPSS was used to examine the linearity of the relationships in the data. Functions were plotted for linear, quadratic, cubic, compound, growth, exponential, and logistic functions and examination of these functions revealed that linear functions were the most 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 4900. 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, White students, and the Emotional Disorders and Related Issues course.

3. Results

3.1. Descriptive statistics

Sixty four percent of students in the sample were female and 36% were male. The mean age of the participants was 31 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 gender, age, race, and ethnicity breakdown of the sample was congruent with that of the overall university population excepting a slight underrepresentation of Latino students in this sample. 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−.

3.1.1. Research question 1: what are the digital reading patterns of students?

Table 2 shows the descriptive statistics of reading behaviors. These results show that students read an average of slightly over 7 h and 20 min over 11 days throughout the entire 16-week semester.

3.1.2. Research question 2: How is the CourseSmart Engagement Index related to course performance?

The hierarchical linear regression predicting course GPA using the CourseSmart Engagement Index was significant ($F_{(18,180)} = 4.442$, $p < .001$, adjusted $R^2 = .238$). Table 3 shows that the Engagement Index Score was a significant positive predictor of course GPA when taking into account student demographics, course enrollment, and prior academic achievement. Congruent with previous research on the effect of instructor and discipline, students enrolled in the second section of *Introduction to Accounting* and in *Evolutionary Theory* had significantly lower course grades (Cohen, 1981; Vanderstoep, Pintrich, & Fagerlin, 1996; Vermunt, 2005). Additionally, congruent with previous research that has shown that student past academic achievement is predictive of future achievement, transfer GPA was positively predictive of course grade (Burton & Dowling, 2005; DeBerard, Speilmans & Julka, 2004; Dennis, 2011; Geiser & Santelices, 2007; Williford, 2009). An examination of the β coefficients shows that the Engagement Index score was more than twice as strong a predictor of student success than prior academic achievement. Furthermore, being enrolled in the second section of *Introduction to Accounting* was almost as strong a predictor as the Engagement Index score while being enrolled in *Evolutionary Theory* was almost twice as strong a predictor as the Engagement Index score.

3.1.3. Research question 3: how are the individual components of the CourseSmart Engagement Index related to course performance?

The hierarchical linear regression predicting course GPA using the components of the CourseSmart Engagement Index was significant ($F_{(21,177)} = 4.205$, $p < .001$, adjusted $R^2 = .254$). As in the previous analyses, Table 4 shows that transfer GPA was positively predictive of course grades while being enrolled in Section 2 of *Introduction to Accounting* or in *Evolutionary Theory* were negatively predictive of course grades. The only component of the Engagement Index predictive of course grades was number of days students read during the semester. As with the analysis of the CourseSmart Engagement Index, number of days students read was a stronger predictor of course outcomes than transfer GPA. Number of days reading was more than twice as strong of a predictor as transfer GPA; it was also as strong of a predictor of course grades as being enrolled in Section 2 of *Introduction to Accounting* and a slightly weaker predictor than being enrolled in *Evolutionary Theory*.

Because of previous research suggesting relationships between highlighting, notes, bookmarks, and outcomes and because the distributions for these variables were skewed, t-tests were conducted to evaluate differences in course grades between students in the upper 10th percentile and those in the lower 90th percentile. Independent samples t-tests were not significant for number of bookmarks ($t(231) = -.175$, $p > .5$) and for number of notes($t(231) = -1.709, > .08$). However, students who were in the top 10th percentile of number of highlights (those highlighting two or more times) had significantly higher course

Table 2
Descriptive statistics of reading behaviors (N = 233).

	Mean (SD)	Median
Pages read	551 (782)	224
Number of days	11 (10)	7
Reading sessions	17 (20)	10
Time reading (min)	442 (664)	169
Highlights	4 (20)	0
Bookmarks	.42 (2)	0
Notes	.16 (1.2)	0

Table 3
Hierarchical regression model exploring how demographics, course/section, transfer GPA, and CourseSmart Engagement Index score predict course grades (N = 233).

	Block 1 gender	Block 2 Race/ethnicity	Block 3 Course/section	Block 4 Transfer GPA	Block 5 Engagement
Independent variables	β	β	β	β	β
Male	−.030	−.038	−.019	−.023	−.016
Latino/Hispanic		−.057	−.026	−.007	−.023
African American		.027	.124	.139	.142 *
Asian American		−.088	−.027	−.034	−.044
Native American		−.034	.013	.026	.035
Other		−.038	−.024	−.013	−.033
Intro to Accounting S1			−.164	−.111	−.241
Intro to Accounting S2			−.346**	−.287*	−.366**
Evolutionary Theory			−.600***	−.524***	−.528***
American Judicial Process S1			−.223	−.162	−.126
American Judicial Process S2			−.175	−.118	−.097
EdPsych Measurement			−.098	−.108	−.093
Assessment of Except Students			−.104	−.049	−.075
Human Resource Management			−.272	−.204	−.218
Production/Operations Management S1			−.099	−.045	−.013
Production/Operations Management S1			−.021	.029	.051
Transfer GPA				.166*	.134*
Engagement score					.288***
Adjusted R^2	−.004	−.015	.166	.186	.238
R^2 change	.001	.015	.217***	.023*	.051***

Note. β = Beta, the standardized regression coefficient.
* $p \le .05$.
** $p \le .01$.
*** $p \le .001$.

grades ($M = 3.52$, SD = 1.23) than students in the lower 90th percentile ($M = 2.69$, SD = 1.55, $t(231) = -2.594$, $p < .01$).

4. Discussion

4.1. Research questions

4.1.1. Question 1: what are the digital reading patterns of students?

Results from this study showed that students did not engage much with their digital textbooks. The mean amount of reading time was just slightly over 7 h and 20 min over the entire semester. Given the skew of the reading time distribution, the median was a better measure of central tendency, in which case students read less than 3 h over an entire 16-week semester. These data are supported by the number of days that students read (mean of 11 and median of 7) and the number of times that students "opened" their digital textbooks (mean of 17 and median of 10). These results are much lower than previous research on student reading time showing that students reported reading 3 h or less per week (Berry et al., 2010). However, 18% of students in Berry et al.'s (2010) study did not read the textbook, which is congruent with Fouh et al.'s (2014) data showing that 50% of students didn't read the text-based components of the course ebook. Future research should examine broader samples (in terms of student characteristics, courses, and institutional type) in order to evaluate normative levels of student digital

Table 4
Hierarchical regression model exploring how demographics, course/section, transfer GPA, and reading behaviors predict course grades (N = 233).

	Block 1 Gender	Block 2 Race/ethnicity	Block 3 Course/section	Block 4 Transfer GPA	Block 5 Engagement
Independent variables	β	β	β	β	β
Male	−.030	−.038	−.019	−.023	−.030
Latino/Hispanic		−.057	−.026	−.007	−.028
African American		.027	.124	.139	.123
Asian American		−.088	−.027	−.034	−.037
Native American		−.034	.013	.026	.010
Other		−.038	−.024	−.013	−.045
Intro to Accounting S1			−.164	−.111	−.249
Intro to Accounting S2			−.346**	−.287*	−.338**
Evolutionary Theory			−.600***	−.524***	−.491***
American Judicial Process S1			−.223	−.162	−.092
American Judicial Process S2			−.175	−.118	−.067
EdPsych Measurement			−.098	−.108	−.072
Assessment of Except Students			−.104	−.049	−.089
Human Resource Management			−.272	−.204	−.229
Production/Operations Management S1			−.099	−.045	.003
Production/Operations Management S1			−.021	.029	.078
Transfer GPA				.166*	.137*
Bookmarks					−.031
Highlights					.030
Number of days					.338***
Notes					.073
Adjusted R^2	−.004	−.015	.166	.186	.254
R^2 Change	.001	.015	.217***	.023*	.077***

Note. β = Beta, the standardized regression coefficient.
* $p \le .05$.
** $p \le .01$.
*** $p \le .001$.

textbook usage. For if students generally read as little as the current study suggests, instructors and institutions need to critically evaluate their approach to reading requirements. Of the three annotation strategies measured, highlighting was the most popular; however, all three were not conducted with much frequency in the sample.

4.1.2. Question 2: how is the CourseSmart Engagement Index related to course performance?

The CourseSmart Engagement Index significantly predicted course performance. In fact, the CourseSmart Engagement Index was a stronger predictor of course grades than previous grades, the single strongest predictor of future student success (Burton & Dowling, 2005; DeBerard, Speilmans & Julka, 2004; Dennis, 2011; Geiser & Santelices, 2007; Williford, 2009). These results are congruent with the idea that students who invest more time and energy in their readings will perform better (Astin, 1984; Chickering & Gamson, 1987). The Engagement Index includes both elements of time on task (like reading time) and active reading (highlighting) and as such seems to be an omibus measure of engagement with the text. Because the Engagement Index can predict course outcomes, it can be used to identify students at-risk of academic failure with better predictive ability than prior achievement (Campbell & Oblinger, 2007). While prior student achievement is not readily available to course instructors, textbook analytics can be easily accessed through the institution's LCMS. Additionally, prior achievement is a *static* measure that may not reflect other important factors, like student engagement, that are related to student academic success. The CourseSmart Engagement Index is a *dynamic* and real-time measure that reflects students' current behaviors that impact course success.

A measure such as the CourseSmart Engagement Index is not dependent on graded course activities, like other learning analytics systems. The predictive ability of graded content is straightforward because poor assignment and quiz grades will predict poor course performance; however, such prediction requires input from the instructor while a measure like the Engagement Index requires no such input. Data that drive the Engagement Index are collected unobtrusively in the background. In the past, a student would have to submit a quiz, exam, essay, or other gradable assignment in order for instructors to estimate how they might do in their course. With the CourseSmart Engagement Index, an instructor can identify students at risk and act on this identification earlier in the semester than if they had to wait to calculate a student's grade (Campbell & Oblinger, 2007).

Of note is the fact that course and instructor variables were controlled for in these analyses. In other words, the Engagement Index was a significant predictor of course grades across disciplines, instructors, and course sections. This suggests that the predictive ability of the Index does not change much, if at all, when using it for courses that require more reading, for textbooks requiring different levels of reading comprehension, for instructors who have different teaching styles, or for subject areas that vary in their technical nature (Campbell & Oblinger, 2007). This is especially important given previous research showing that students from different academic disciplines may have different reading comprehension skills (Alderson & Urquhart, 1985).

4.1.3. Question 3: how are the individual components of the CourseSmart Engagement Index related to course performance?

The model using the individual components of the Engagement Index yielded greater predictive power than the model using the Engagement Index. Just like the previous model, only four variables were significant predictors of course grades in the individual components model—two course variables (Section 2 of Introduction to Accounting and Evolutionary Theory), transfer GPA, and number of days spent reading. Number of days spent reading was a stronger predictor of course grade than the CourseSmart Engagement Index, predicting 2% more of the variance.

Astin (1984) and Chickering and Gamson's (1987) theories predict that time on task will promote engagement and lead to improved student learning. According to these models, more time spent reading will lead to better learning outcomes. Indeed, the current study found that time spent reading was significantly related to course grade. This was congruent with research by Landrum et al. (2012), who found that percentage of reading completed was related to final course grade. However, this was not congruent with Daniel and Woody's (2013) work that found that reading time differences did not indicate performance differences. The discrepancy in these findings is likely due to the fact that the Landrum et al. (2012) and the Daniel and Woody (2013) studies used self-report measures while the current study used actual digital textbook usage and course grades (the Daniel and Woody (2013) study observed reading times in the lab but had students self-report their home reading habits).

Digital textbook analytics might help elucidate a relationship between reading behaviors and learning outcomes not obtainable through research using self-reports and laboratory studies of reading. Self-reported measures are notoriously bad at relating to actual behaviors (Celis-Morales et al., 2012; Hald, Overgaard & Grau, 2003; Junco, 2013b; Otten, Littenberg & Harvey-Berino, 2010) and this is also the case with reading (Daniel & Woody, 2013). As such, the error introduced in these measurements makes it difficult, if not impossible to relate them to outcomes. Digital textbook analytics, however, bypass the measurement error inherent in self-reported reading behaviors and, in this case, help identify a reading frequency index helpful in predicting outcomes.

Engagement with the textbook is understood through Astin (1984) and Chickering and Gamson's (1987) engagement frameworks. Students who spent more time on task engaged with their text did better in their courses. Indeed, time on task as measured by the number of days a student read is a better predictor of course outcomes than the calculated Engagement Index measure. The positive effects of such engagement with the textbook over time might also be conceptualized as increased retention of textbook material through spaced practice. In other words, course performance will be better for students who study material over multiple days versus those who study material all at once (Carpenter, Cepeda, Rohrer, Kang & Pashler, 2012). Engagement is a necessary condition for the benefits of such spaced practice (Astin, 1984; Chickering & Gamson, 1987).

Highlighting showed an ability to predict student grades, but not as part of a linear model. Previous research found that low-skilled readers highlighted more than high-skilled readers and suggested that highlighting may be used as a proxy for reading skill (Bell & Limber, 2009); however, that was not the case in the current study. Students who highlighted more performed better in their courses perhaps not indicating that they had better reading skills but that they were much more engaged with their textbooks in general. Chickering and Gamson (1987) emphasize the importance of active learning—that is, learning that engages the student and makes them active participants in the learning process.

Highlighting and annotating are active reading strategies that help students engage with the text in order to remember and understand key concepts (Bell & Limber, 2009). Indeed, a post hoc analysis showed that students who highlighted two or more times had significantly higher Engagement Index Scores ($M = 1663.64$, $SD = 1219.08$) than students who highlighted less than two times ($M = 588.59$, $SD = 755.56$, $t(231) = -6.223$, $p < .001$). Although the curve estimation procedure found that the best relationship between highlighting and grades was linear, the post hoc t-tests suggested the possibility of using a different function for a better fit. A number of transformations were attempted on the highlighting data (including logarithmic and quadratic) in order to see if predictive power could be improved; however, none of these functions yielded a better solution than the linear model as suggested by the curve estimation analyses.

4.2. Summary

Digital textbook analytics can serve as an effective and unobtrusive method of formative assessment. The results of this study show that both the CourseSmart Engagement Index and the number of days students read were both strong predictors of student course performance even when controlling for prior academic achievement, discipline, course, and instructor. However, the number of days students read was a much stronger predictor of course performance than the CourseSmart Engagement Index. Further research should examine how best to combine the individual factors of the Engagement Index to strengthen its predictive ability. While transforming the individual components of the index might seem like a place to start, the results of the curve estimation suggest that nonlinear transformations would not yield better predictive ability. Another possibility is to weigh the different variables in the Index to reflect their predictive ability. For instance, number of days spent reading might be weighed more heavily while also weighing number of highlights above the 90th percentile more heavily.

Typically, it is difficult to measure the level of student engagement in college courses (Pascarella & Terenzini, 2005). Traditionally, student engagement has been measured by its outcomes—grades on academic work, attendance, semester grades, and student persistence (Pascarella & Terenzini, 2005). However, digital textbook analytics allow instructors and institutions to evaluate time on task for the academic task of reading as well as how much students are engaging with active learning using their textbooks, both of which are important indicators of student engagement (Astin, 1984; Chickering & Gamson, 1987). Up until now, student performance information has only been available through traditional assessment measures such as quizzes, exams, essays or other gradable assignments. However, textbook engagement data provided by digital textbook analytics can help instructors track how engaged students are at any given time in the course and help them predict course outcomes. According to Campbell and Oblinger (2007), since instructors can access these predictive data, they can act on this information to help support students. Therefore, digital textbook analytics show promise to be used as an effective early warning system to identify students at risk of academic failure. Furthermore, digital textbook analytics can help faculty plan, adjust, and assess their teaching strategies in order to maximize student engagement with the text. For instance, faculty can have a readily available way to evaluate how well students are (or are not) completing their course readings and intervene appropriately.

Even though previous research showed that the instructor and the academic discipline of the course affect student academic performance (Cohen, 1981; Vanderstoep, Pintrich, & Fagerlin, 1996; Vermunt, 2005), that engagement with the textbook varies by discipline and course level (Fitzpatrick & McConnell, 2008), and that levels of reading comprehension vary by discipline (Alderson & Urquhart, 1985), these factors did not affect the ability of the CourseSmart Engagement Index or the number of days reading to predict course outcomes. It is important to note that other factors such as how much of the course grade relied upon textbook material, how much reading was required, and whether students were instructed how to use the digital text were not standardized across courses in this study—this shows the ability for digital textbook analytics to predict learning outcomes with different course reading requirements.

This study found that students don't spend much time reading, as has been found in other studies (Berry et al., 2010; Taylor, 2011). Unlike previous research (Bell & Limber, 2009; Daniel & Woody, 2013; Fitzpatrick & McConnell, 2008; Uso-Juan & Ruiz-Madrid, 2009), 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 although those who did highlight a great deal had better academic outcomes. This finding was incongruent with Taylor's (2011) discovery that highlighting and annotating were not related to learning; however, Taylor (2011) did not evaluate how number of highlights related to learning, only whether students were asked to highlight or to refrain from highlighting in an experimental protocol. This could hint at an active learning process issue for further study—students who choose to highlight (perhaps because they are more engaged in general) do better in their courses. It is possible that the act of highlighting in digital textbooks is different than in print textbooks: in print, excessive highlighting signals an inability to select relevant information while frequent digital highlighting signals greater interaction and engagement with the reading materials, an issue for consideration in future research (Bell & Limber, 2009).

4.3. Limitations

This study has a number of limitations for consideration. First, the sample used for this study was a convenience sample of 236 students—the results, while significant, may not be able to be generalized to students at other institutions of higher education because of the unique and diverse nature of the student population at this particular institution. Further digital textbook analytics research will want to recruit larger samples that reflect the makeup of higher education institutions across the United States and internationally. Another limitation is that although the CourseSmart Engagement Index and number of days spent reading were significant and strong predictors of course grade relative to other variables like course and previous GPA, the percentages of variance explained by the addition of the textbook analytics elements are low (5% and 8%, respectively). Further research should work to elucidate other digital textbook engagement variables that would increase the overall predictive ability of these models. A final limitation is that students were not engaged with the digital textbooks much in this study. Again, further research should replicate such work with larger samples to be able to generalize how students engage with course reading materials.

5. Conclusion

This study showed the utility of digital textbook analytics in predicting student course grades. While the CourseSmart Engagement Index was a good predictor, the number of days students spent reading was a more powerful predictor of course outcomes. This suggests that the calculated Engagement Index does not yet capture the important factors related to engagement with the textbook. Time on task is an important factor in student engagement (Astin, 1984; Chickering & Gamson, 1987); however, the Engagement Index does not weigh this factor heavily enough. Also while highlighting was related to final course grades, it was not significant in the predictive models. Additional transformations of the highlighting data did not yield a more fitting solution. Future work on digital textbook analytics should consider how to best account for highlighting in the predictive model. This work suggests that transforming highlighting using a probability-based algorithm and adding that to time spent reading would yield an even more powerful predictor of student course grades. Perhaps another variable to consider in a revised digital textbook engagement index would be a measure of re-reading—re-reading the same pages increases time on tasks and improves retention of textbook material through spaced practice (Carpenter et al., 2012).

The CourseSmart Engagement Index is limited by the features available in CourseSmart books—the promise of digital textbooks is the ability for them to be more engaging than regular texts (Young, 2013). Digital textbooks with interactive exercises give students the opportunity to engage with the text in ways that are more congruent with theories of student engagement—such engagement leads to better course performance (Astin, 1984; Fouh et al., 2014). Having exercises that employ active learning more fully capitalizes on the unique affordances of digital textbooks. A refined engagement index could take advantage of these engaging elements to improve predictive ability.

Data from this study show that the CourseSmart Engagement Index needs to be refined, the last step in Campbell and Oblinger's (2007) learning analytics process. Even though the CourseSmart Engagement Index should be refined to include better measures of engagement, the current study's results show that digital textbook analytics are a promising new form of learning analytics. Adding digital textbook engagement data to existing learning analytics tools that base predictions on activities conducted on the 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, digital textbook analytics can help faculty plan their courses and conduct more efficient formative assessment of their students than ever before possible. In either case, digital textbook analytics show that student-generated data outside of a LCMS can be used to effectively predict learning and to help plan meaningful student-centered interventions.

References

Alderson, J.C., & Urquhart, A.H. (1985). The effect of students' academic discipline on their performance on ESP reading tests. *Language Testing*, *2*(2), 192–204.

Junco, R. (2013a). iSpy: Seeing what students really do online. *Learning, Media and Technology*, *39*(1), 75–89.

Junco, R. (2013b). Comparing actual and self-reported measures of Facebook use. *Computers in Human Behavior*, *29*(3), 626–631.

Arnold, K.E., & Pistilli, M.D. (2012). *Course signals at Purdue: Using learning analytics to increase student success. Proceedings of the 2nd International Conference on Learning Analytics & Knowledge.* New York, NY: ACM Retrieved May 7, 2015, from: https://www.itap.purdue.edu/learning/docs/research/Arnold_Pistilli-Purdue_University_Course_Signals-2012.pdf.

Association of American Publishers (2014, June). *US Publishing Industry annual survey reports $27 Billion in net revenue, 2.6 billion units For 2013. Press Release* Retrieved May 7, 2015, from: http://www.publishers.org/press/138/.

Astin, A.W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel*, *25*(4), 297–308.

Bell, K.E., & Limber, J.E. (2009). Reading skill, textbook marking, and course performance. *Literacy Research and Instruction*, *49*(1), 56–67.

Berry, T., Cook, L., Hill, N., & Stevens, K. (2010). An exploratory analysis of textbook usage and study habits: Misperceptions and barriers to success. *College Teaching*, *59*(1), 31–39.

Bossaller, J., & Kammer, J. (2014). Faculty views on eTextbooks: A narrative study. *College Teaching*, *62*(2), 68–75.

Burton, L.J., & Dowling, D.G. (2005). In search of the key factors that influence student success at university. *Proceedings of the 28th HERDSA Annual Conference* (pp. 68–78).

Campbell, J.P., & Oblinger, D.G. (2007). *Academic analytics. EDUCAUSE paper* Retrieved May 8, 2015, from: https://net.educause.edu/ir/library/pdf/pub6101.pdf.

Carpenter, S.K., Cepeda, N.J., Rohrer, D., Kang, S.H.K., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review*, *24*(3), 369–378.

Celis-Morales, C.A., Perez-Bravo, F., Ibañez, L., Salas, C., Bailey, M.E.S., & Gill, J.M.R. (2012). Objective vs. self-reported physical activity and sedentary time: Effects of measurement method on relationships with risk biomarkers. *PLOS One*, *7*(5) Retrieved May 8, 2015, from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0036345.

Chickering, A.W., & Ehrmann, S.C. (1996, October). Implementing the seven principles: Technology as a lever. *AAHE Bulletin*, *3–6*.

Chickering, A.W., & Gamson, Z.F. (1987, March). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, *3–7*.

Cohen, P.A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of Educational Research*, *51*(3), 281–309.

Daniel, D.B., & Woody, W.D. (2013). E-textbooks at what cost? *Performance and use of electronic v. print texts.. Computers & Education*, *62*. (pp. 18–23).

Dawson, S.P., McWilliam, E., & Tan, J.P. -L. (2008). Teaching smarter: How mining ICT data can inform and improve learning and teaching practice. *Annual conference of the Australasian Society for Computers in Learning in Tertiary Education, 221-230*. Melbourne, Australia: Deakin University.

DeBerard, M.S., Speilmans, G.I., & Julka, D.L. (2004). Predictors of academic achievement and retention among college freshmen: A longitudinal study. *College Student Journal*, *38*(1), 66–80.

Dennis, A. (2011). *e-Textbooks at Indiana University: A summary of two years of research. Indiana University Kelly School of Business Research Report.* Retrieved June 6, 2014, from: http://www.indiana.edu/~etext/files/eText%20Pilot%20Data%202010-2011.pdf.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5/6), 304–317 Retrieved May 8, 2015, from: http://oro.open.ac.uk/36374/1/IJTEL40501_Ferguson%20Jan%202013.pdf.

Fouh, E., Breakiron, D.A., Hamouda, S., Farghally, M.F., & Shaffer, C.A. (2014). Exploring students learning behavior with an interactive etextbook in computer science courses. *Computers in Human Behavior*, *41*, 478–485.

Fitzpatrick, L., & McConnell, C. (2008). Student reading strategies and textbook use: An inquiry into economics and accounting courses. *Research in Higher Education Journal* Retrieved May 8, 2015, from: http://www.aabri.com/manuscripts/09150.pdf.

Garrison, C., & Ehringhaus, M. (2009). Formative and summative assessment in the classroom. *National Middle School Association*, *1–3* Retrieved May 8, 2015, from: http://schools.nyc.gov/NR/rdonlyres/33148188-6FB5-4593-A8DF-8EAB8CA002AA/0/2010_11_Formative_Summative_Assessment.pdf.

Geiser, S., & Santelices, M.V. (2007). Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes. *Berkeley, CA: Center for Studies in Higher Education Research and Occasional Paper Series, University of California, Berkeley* Retrieved May 8, 2015, from: http://www.cshe.berkeley.edu/sites/default/files/shared/publications/docs/ROPS.GEISER._SAT_6.13.07.pdf.

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society*, *15*(3), 42–57.

Hald, J., Overgaard, J., & Grau, C. (2003). Evaluation of objective measures of smoking status—A prospective clinical study in a group of head and neck cancer patients treated with radiotherapy. *Acta Oncologica*, *42*(2), 154–159.

Hsieh, P. -H., & Dwyer, F. (2009). The instructional effect of online reading strategies and learning styles on student academic achievement. *Educational Technology & Society*, *12*(2), 36–50.

Kauffman, D.F., Zhao, R., & Yang, Y. -S. (2011). Effects of online note taking formats and self-monitoring prompts on learning from online text: Using technology to enhance self- regulated learning. *Contemporary Educational Psychology*, *36*(4), 313–322.

Kuh, G.D. (2009). What student affairs professionals need to know about student engagement. *Journal of College Student Development*, *50*(6), 683–706.

Landrum, R.E., Gurung, R.A.R., & Spann, N. (2012). Assessments of textbook usage and the relationship to student course performance. *College Teaching*, *60*(1), 17–24.

Le Bigot, L., & Rouet, J. -F. (2007). The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *Journal of Literacy Research*, *39*(4), 445–470.

Longhurst, J. (2003). World history on the World Wide Web: A student satisfaction survey and a blinding flash of the obvious. *History Teacher*, *36*(3), 343–357.

Lonka, K., Olkinuora, E., & Mäkinen, J. (2004). Aspects and prospects of measuring studying and learning in higher education. *Educational Psychology Review*, *16*(4), 301–323.

Macfadyen, L.P., & Dawson, S. (2012). Numbers are not enough. Why e-Learning analytics failed to inform an institutional strategic plan. *Educational Technology & Society*, *15*(3), 149–163.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., et al. (2011). *Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute Big Data Report..* Retrieved May 8, 2015, from: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.

Mattingly, K., Rice, M., & Berge, Z. (2012). Learning analytics as a tool for closing the assessment loop in higher education. *Knowledge Management & E-Learning: An International Journal*, *4*(3), 236–247.

New, J. (2013, April). *Pearson acquires learning catalytics, a cloud-based assessment system.* The Chronicle of Higher Education Retrieved April 24, 2015 from: http://chronicle.com/blogs/wiredcampus/pearson-acquires-learning-catalytics-a-cloud-based-assessment-system/43543.

Nielsen (2014, August). *2014 Nielsen Children's Book Industry Report.* New York, NY: Report from the Nielsen Children's Book Summit.

O'Brien, D., & Voss, S. (2011). Reading multimodally: What is afforded? *Journal of Adolescent & Adult Literacy*, *55*(1), 75–78.

Otten, J.J., Littenberg, B., & Harvey-Berino, J.R. (2010). Relationship between self-report and an objective measure of television-viewing time in adults. *Obesity*, *18*(6), 1273–1275.

Pascarella, E.T., & Terenzini, P.T. (2005). *How college affects students: A third decade of research.* San Francisco, CA: Jossey-Bass.

Picciano, A.G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks*, *16*(3), 9–20.

Rainie, L. (2012). *Tablet and E-book reader ownership nearly double over the holiday gift-giving period.* Washington, DC: Pew Research Center Retrieved May 7, 2015, from: http://libraries.pewinternet.org/2012/01/23/tablet-and-e-book-reader-ownership-nearly-double-over-the-holiday-gift-giving-period/.

Rockinson-Szapkiw, A.J., Courduff, J., Carter, K., & Bennett, D. (2013). Electronic versus traditional print textbooks: A comparison study on the influence of university students' learning. *Computers & Education*, *63*, 259–266.

Sandberg, K. (2011). College Student Academic online reading: A review of the current literature. *Journal of College Reading and Learning*, *42*(1), 89–99.

Schugar, J.T., Schugar, H., & Penny, C. (2011). A Nook or a book? Comparing college students' reading comprehension levels, critical reading, and study skills. *International Journal of Technology in Teaching and Learning*, *7*(2), 174–192.

Siemens, G. (2010, July 22). *1st international conference on Learning Analytics and Knowledge 2011.* Retrieved May 7, 2015, from: https://tekri.athabascau.ca/analytics/.

Smith, V.C., Lange, A., & Huston, D.R. (2012). Predictive modeling to forecast student outcomes and drive effective interventions in online community college courses. *Journal of Asynchronous Learning Networks*, *16*(3), 51–61.

Spencer, C. (2006). Research on learners' preferences for reading from a printed text or from a computer screen. *Journal of Distance Education*, *21*(1), 33–50.

Taylor, A.K. (2011). Students learn equally well from digital as from paperbound texts. *Teaching of Psychology*, *38*(4), 278–281.

Uso-Juan, E.U., & Ruiz-Madrid, M.N. (2009). Reading printed versus online texts. A study of EFL learners' strategic reading behavior. *International Journal of English Studies*, *9*(2), 59–79.

Vanderstoep, S., Pintrich, P., & Fagerlin, A. (1996). Disciplinary differences in self-regulated learning in college students. *Contemporary Educational Psychology*, *21*(4), 345–362.

Verbert, K., Manouselis, N., Drachsler, H., & Duval, E. (2012). Dataset-driven research to support learning and knowledge analytics. *Educational Technology & Society*, *15*(3), 133–148.

Vermunt, J.D. (2005). Relations between student learning patterns and personal and contextual factors and academic performance. *Higher Education*, *49*(3), 205–234.

Weisberg, M. (2011). Student attitudes and behaviors towards digital textbooks. *Publishing Research Quarterly*, *27*(2), 188–196.

Williford, A.M. (2009). Secondary school course grades and success in college. *College and University*, *85*(1), 22–33.

Woody, W.D., Daniel, D.B., & Baker, C.A. (2010). E-books or textbooks: Students prefer textbooks. *Computers & Education*, *55*(3), 945–948.

Yang, Y. -F. (2010). Developing a reciprocal teaching/learning system for college remedial reading instruction. *Computers & Education*, *55*(3), 1193–1201.

Young, J. (2013, January). The object formerly known as the textbook. *The Chronicle of Higher Education* Retrieved April 24, 2015 from: http://chronicle.com/article/Dont-Call-Them-Textbooks/136835/.

Zickuhr, K., & Rainie, L. (2014). *A snapshot of reading in America in 2013.* Washington, DC: Pew Research Center Retrieved May 7, 2015, from: http://www.pewinternet.org/2014/01/16/a-snapshot-of-reading-in-america-in-2013/.