

Identifying Challenging Spreadsheet Skills Using Reading and Homework Analytics from an Interactive Textbook

Prof. Matthew W. Liberatore, University of Toledo

Matthew W. Liberatore is a Professor of Chemical Engineering at the University of Toledo. He earned a B.S. degree from the University of Illinois at Chicago and M.S. and Ph.D. degrees from the University of Illinois at Urbana-Champaign, all in chemical engineering. His current research involves the rheology of complex fluids as well as active learning, reverse engineering online videos, and interactive textbooks. His website is: <http://www.utoledo.edu/engineering/chemical-engineering/liberatore/>

Kayla Chapman

Identifying challenging spreadsheet skills using reading and homework analytics within an interactive textbook

Abstract

Spreadsheets provide a user-friendly environment to complete many calculations for engineering students as well as practicing engineers. Spreadsheet programs have been available for decades, including Microsoft Excel, Google Sheets, and many others; numerous formulas, functions, and other tasks are common across versions and platforms (e.g., Windows, Mac, iOS, Android). An interactive textbook from zyBooks titled *Spreadsheets Essentials* provides students opportunities to step through animations, complete question sets, and perform spreadsheet functions on auto-graded, randomized problems. Advanced topics include solver, error, statistics, interpolation, numerical integration, matrix operations, and more. The interactive format creates ‘big data’ on student behaviors, such as reading completion, auto-graded problem success, and auto-graded problem attempts. Providing nominal course credit led to very high reading rates in 2018. Specifically, at least 75% of students (n=98) completed 100% of the reading participation encompassing over 250 clicks across 12 sections. In addition, animation view rates quantified repetition with rates as high as 118%. Finally, success on over 100 auto-graded questions was higher for basic skills and formulas than more advanced topics, such as integration and interpolation.

Introduction

Spreadsheets are common in home and professional life from budget spreadsheets to engineering calculations. Leveraging the ability to efficiently organize, calculate, and plot, spreadsheets have become commodities after decades of use. While the history of spreadsheets centers on Microsoft Excel as computer software, e.g., [1], the growth of open source and online spreadsheets, such as Google Sheets, fuels the current growth and ubiquity of spreadsheets.

Spreadsheet education and training resources are abundant on the Internet. Any web search using terms such as spreadsheet, excel, or MS Excel yield hundreds or thousands of online resources, many free of charge. While many free websites in this genre generate revenue from embedded ads, other businesses focus on synchronous online courses, in-person workshops, short courses, or boot camps. However, almost all spreadsheet education or training adopt a master-student learning framework, i.e., the spreadsheet instructor demonstrates with screenshots, movies, or live demonstration how the formulas and functions of a spreadsheet can be implemented. Vision is the key sense to stimulate learning in most contexts [2], and demonstrative training practices involve visuals. However, most demonstrative online courses lack individualized feedback without significant additional costs. Here, an interactive textbook combines visuals with feedback through auto-graded problems, which may be considered active learning.

Both single studies and meta-analyses of active learning show that students learn more through doing [3-6]. Active learning encompasses many different techniques, which will not be reviewed here. However, examples of active learning in spreadsheet education or training are scarce. To move away from static text in books and demonstrative websites and videos, interactivity in an

online platform can create learning by doing opportunities [7, 8]. Interactive platforms may leverage the strengths of many college students that are digital natives [9, 10] as well as anyone entrenched in digital technologies.

For about 100 years, textbooks have been a core technology in higher education. However, personalized electronics, such as smart phones, tablets, and laptops, have become an indispensable resource, so web searches are commonly seen as more convenient than reading static textbooks [11]. Additionally, web searches are generally free while textbook prices have risen many times faster than inflation [12]. Despite the multi-billion dollar per year industry size [12], textbook usage is poorly documented. As discussed in our previous publications, most research shows a majority of college students do not read textbooks [13-18]. However, technology has empowered reading quizzes, interactive textbooks, and other student-centered activities that lead to positive results related to out of class reading and participation [6-8, 19-24]. Specifically, fully interactive textbooks from zyBooks have shown excellent results. Course median reading rates up to 99% were observed. The interactive textbook format has also been demonstrated to help students in the lower third of the class engage in the course [7, 8, 19, 20, 25-27]. Reading participation using a zyBook is also discussed in another 2019 ASEE contribution.

In this contribution, an interactive textbook for teaching spreadsheets will be reviewed, reading participation and repetition analyzed, and successes and challenges of auto-graded problems summarized.

Materials: An interactive textbook with spreadsheets

zyBooks creates interactive textbook replacements using the philosophy: Less text, more action™. These interactive textbooks are viewed, read, and interacted with in any HTML5 compliant web browser without additional plug-ins. The lead author created a *Material and Energy Balances* zyBook, available since January 2016. In 2018, a chapter on spreadsheets was added along with additional spreadsheet content to help students solve material and energy balance problems. The chapter on spreadsheet skills is the focus of this work, and this content is also available as a stand-alone title, *Spreadsheets Essentials*, for a price of \$10 per student per semester. The *Material and Energy Balances* zyBook is required for the course; reading participation and auto-graded problems account for 10% of the final course grade, which is detailed in another paper for this conference. The necessity of interactivity to earn a grade eliminates the common issue with traditional textbooks where a static file (e.g., a pdf) is produced and distributed illegally without cost to the student.

A total of 15 sections are available related to spreadsheets, 14 containing content (Table 1) and one aggregating other resources to continue exploring the topic. Spreadsheet concepts are divided into three general categories ranging from very basic to advance skills useful for science and engineering students as well as scientists and engineers.

Table 1. Groupings of section titles in *Spreadsheets Essentials* zyBook as of February 2019.

General skills	Functions	Advanced skills
Spreadsheet basics	Spreadsheet functions	Calculating error and basic statistics
Spreadsheet formulas	Functions: Math and trigonometry	Interpolation
Sort and filter	Functions: Logical	Integration and numerical integration
Creating a chart	Matrix functions	Systems of linear equations
Trendlines		
Solvers		

Data generated by two cohorts of students at a public research institution during the Spring 2018 and 2019 semesters will be discussed. Approximately 100 students in each cohort were assigned reading and auto-graded exercises. Students were primarily in their first year of college (freshman) majoring in chemical engineering and environmental engineering, approximately 60-65% male and 35-40% female. Many students received spreadsheet training during the previous semester.

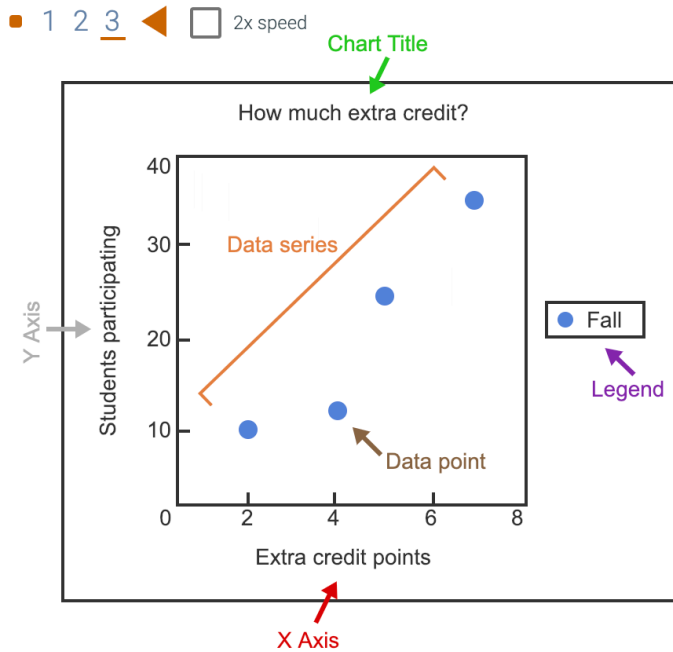
Reading participation and success on challenge activities are fully transparent, so both student and instructor can monitor progress in real time. Several types of data are collected and relevant to this study: 1. Clicks record reading participation, such as steps in an animation or completing a matching exercise, 2. Success on auto-graded challenge activities, and 3. Attempts before correct for challenge activities. New content is created with most clicks. For example, clicking an answer to a multiple-choice question reveals new, unique explanations for correct and incorrect responses. Neither clicking incorrect answers when reading or incorrect attempts on challenge activities lower students' grades for using the interactive textbook. Some salient features of *Spreadsheets Essentials* are discussed next and summarized in Table 2.

Table 2. Features of *Spreadsheets Essentials* zyBook as of February 2019.

Feature	Number
Sections	14
Animations	45+
Clicks to 'read'	290+
Auto-graded challenge activities	120+

As discussed in previous work, sections or subsections are divided into manageable chunks, which is in sync with cognitive load theory [3, 28-30]. Each unit of content progresses through Define, Demonstrate, Practice, and Challenge. While more detailed explanations and examples of the interactive textbooks format were included in previous publications [31], an example animation, question set, and challenge activity question are included below. In brief, Define uses text and tables to define key terms, which are indexed and searchable as most electronic books. Demonstrate conveys figures and derivations as interactive animations. By creating a figure in a small number of steps, about 3 to 6, students are prescribed to spend more time to contextualize new, visual information. For example, a chart, specifically a scatter plot (Figure 1), contains three steps with the final step describing the function of a legend. Taking less than 2 minutes to watch in most cases, animations are in sync with human's attention span – based upon billions of video views [32].

Previous work included surveys of students, which found favorable reviews of the interactive format. However, additional student surveys are outside the scope of this work.



A legend matches the symbols in the chart with a brief description. Legends can be located inside or outside the main chart area.

Figure 1. Final frame of an animation about the different parts of a chart.

After Define and Demonstrate, Practice involves learning questions, which vary in format from multiple choice to true/false to matching. Learning questions replace static text from a traditional textbook, e.g., explanations or limitations, with interactivity. A matching exercise related to interpolation (Figure 2) mixes stored content (formulas) with displayed content (results from formulas). Finally, auto-graded problems with rolling numbers and content make up the Challenge component of each content unit. Rolling numbers are usually manifested with changing the number in a cell or cells of a spreadsheet of a question level. For example, numerous images of charts (Figure 3) are included as a multiple-choice question to begin a challenge activity on charts. Both learning questions and challenge activities are designed using scaffolding [6], so easier questions precede more difficult ones.

Match the black numerals with the correct value or spreadsheet formula. Properties are for 1 kg of water.

	A	B	C	D	E
1	P (MPa)	5	5.39	6	
2	700	88.5	II	73.5	Volume (L)
3	771	I	III	V	
4	800	98.2	IV	81.6	
5	T (°C)				

III I V IV II

	95.4
	=B2+(D2-B2)/(D1-B1)*(C1-B1)
	89.1
	91.7
	=D2+(D4-D2)/(A4-A2)*(A3-A2)

Figure 2. Matching participation activity for interpolation, which mixes both stored and displayed content.

Select the type of chart shown below.

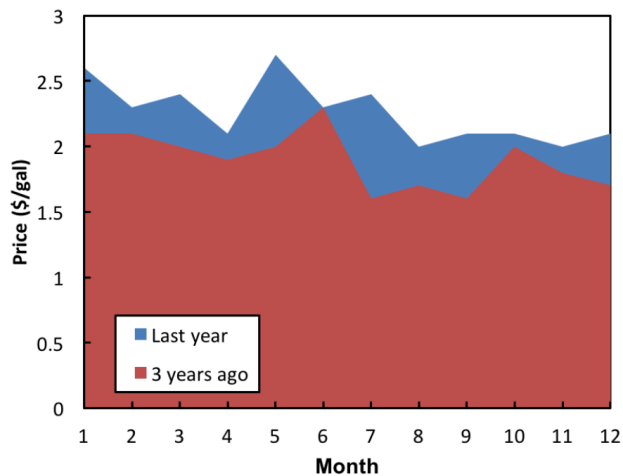


Chart type. ▾

1 2 3 4 5 6 7

Figure 3. First question of a challenge activity about charts. A multiple-choice, pull down menu displaying “Chart type.” allows students to select answers. Multiple versions of this question cover 6 different chart types.

Results and discussion

Reading data

Reading participation is quantified through students' clicks. Specifically, only reading participation completed before the due date are reported here, while reading before and after the due date is discussed in another contribution for this conference. Reading participation translates to 5% of the total course grade, and the spreadsheet content accounted for about 15% of the semester's reading assignments. Small incentives have led to high reading rates with interactive textbooks [20, 25, 26].

Twelve sections of the interactive spreadsheet textbook were completed by the 2018 cohort, and fourteen sections are being completed by the 2019 cohort. The quantity of reading clicks allows for further data analysis. For example, over 24,000 student interactions were collected for the 2018 cohort. Average reading rates with standard deviations as well as median, 1st quartile, and 3rd quartile scores provided ways to monitor students' effort in real time during a semester. Additional analysis will compare reading between cohorts for the talk, since the 2019 data are being collected after the conference deadlines.

Average and 1st quartile reading participation demonstrate the high level of student engagement (Figure 4). With average reading rates near 90% or higher as well as 1st quartile values of 100% for all sections, the fraction of students not fully participating is low. In other words, at least 75% of students earned 100% reading participation. In most cases, students complete 100% of the reading or do not start the reading and earn 0% at the due date, which agrees with our previous work [25, 26].

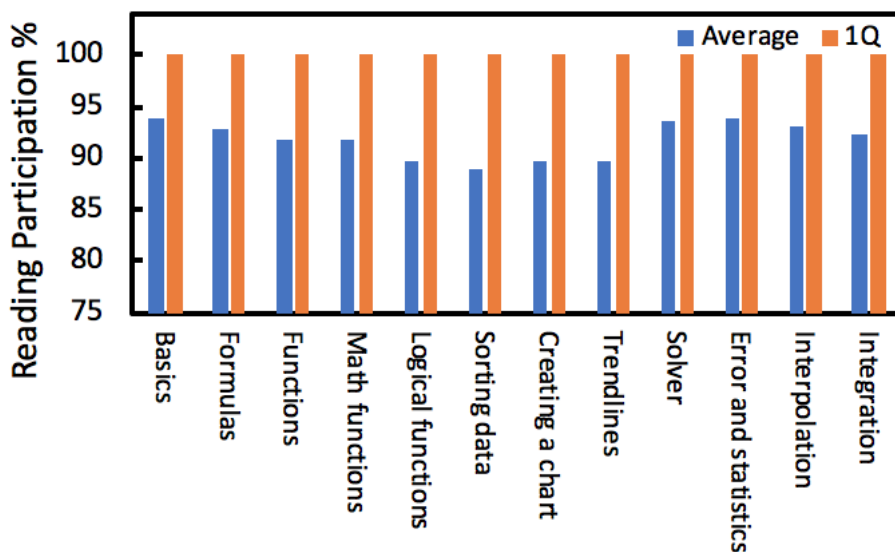


Figure 4. Reading participation (average and 1st quartile) for 12 sections (2018 cohort, n = 98 students).

Animation usage

Determining which animations receive the most views is analogous to counts of views on YouTube or other video sites. Examining the view rates may help elucidate the concepts that were more

difficult for students. The connection between repetition and learning is well documented [29], but discussion of repetition with textbooks is scarce. Unlike reading participation, animation views were taken after the semester and not at each assignments' due date. For the 2018 cohort, the three most watched animations span three different sections of the interactive textbook (Table 3). All three of the most watched animations gathered at least 11% extra views, so at least 10 students re-watched each animation.

Table 3. Most watched animations for 2018 cohort (n = 98 students).

Animation title	Section title	Views (%)
Formulas using \$	Spreadsheet formulas	118
Worksheet basics	Spreadsheet basics	114
If statements in a spreadsheet	Spreadsheet functions	111

Using the dollar sign (\$) in formulas was the most watched animation, which quantifies the ubiquity of the dollar sign in formulas as well as the different uses of the dollar sign. While using two dollar signs to anchor a cell is common (e.g., \$B\$3), using only a single dollar sign is useful but not as common in the authors' experiences (e.g., \$C2 or C\$2). Worksheet basics is the first animation of the chapter and covers renaming and relocating sheets within a workbook. On one hand, the novelty of a spreadsheet animation may contribute to the high view rate. On the other hand, students had seen many material and energy balances animations at this point in the semester. Alternatively, the content of the Worksheet basics animation may be simpler than and not covered by many online spreadsheet training videos, which may have led to the high view rate. Finally, using the IF function is a more advanced topic that students may find useful but have little experience with until the section's challenge activities.

Challenge activity data

Students' skills assessment is completed using challenge activities. For spreadsheets, challenge activities allow students to enter formulas, functions, and calculated values to test their strengths using spreadsheets. With hundreds of numeric combinations on many problems, students can repeat a new version of any question until they compute correct answer(s). With over 100 different questions, the most difficult spreadsheet skills can be identified from students' success. The number of attempts before correct and total attempts will complement the percent correct to give multiple metrics. Over 9,000 questions were attempted by the 2018 cohort and will be analyzed here. Responses from the 2019 cohort will be compared in the conference presentation.

Challenge activity scores varied between 65 to 90% across 12 sections for the 2018 cohort (Figure 5). Each section was comprised of 3 to 17 questions (white numbers in Figure 5). For example, 7 question levels in the Creating a chart section were attempted by 100 students. The first column quantifies 93% correct, which indicates over 650 of the 700 questions were answered correctly – independent of the number of attempts before answering correctly. Parsing the sections into the three categories discussed earlier in Table 1, challenge activity success follows a trend. First, general skills in six sections corresponded with the highest success rates between 87 and 93%. Next, the three sections related to functions saw slightly lower scores (84 to 87%), and the advance skills sections returned even lower scores (65 to 80%). Further analysis of individual question levels is included next.

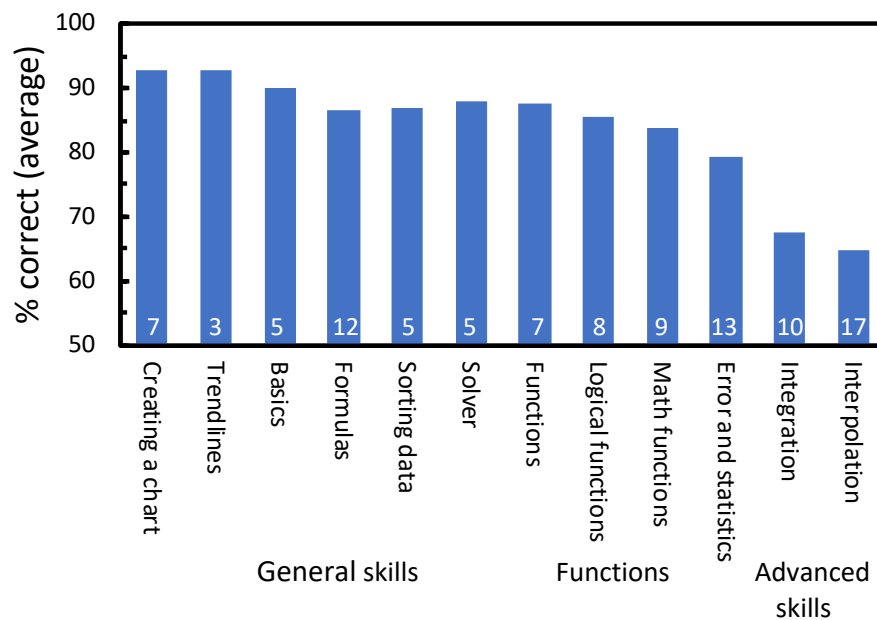


Figure 5. Average success on challenge activities across different sections about spreadsheets. White numbers within bars indicate the number of questions in each section.

The interpolation section serves as a case study of individual questions (Table 4). Ten question levels make up two challenge activities, one covering linear interpolation and a second on double interpolation. The first question is multiple choice with only two choices to reinforce definitions of interpolation and explanation. Only 82% of students answered correctly, compared to much higher success rates in earlier sections on spreadsheets (discussed above). The next four questions have about the same success rate and median number of attempts before correct. With only one or two attempts before correct (median), students who persisted answered these questions without significant struggle. Therefore, students not answering correctly likely gave up quickly and did not ask questions of the professor or teaching assistants. As mentioned earlier, challenge activity scores are independent of the number of attempts, so students finding the correct answer on the 10th attempt earn the same credit as students calculated a correct answer on the 1st attempt.

The second activity shows very low success rates (44 to 56%) while the difficulty is similar to the first challenge activity about interpolation. Since double interpolation is repeating single linear interpolation multiple times, students are likely wary of the number of calculations. Both single and double interpolation using steam tables are covered during the material and energy balances course; one interpretation of the success rates on these problems is the lack of units to relate to the physical world, such as specific volume or enthalpy from steam tables. Additional data will need to be analyzed to better quantify students' difficulty with this advanced spreadsheet topic.

Table 4. Types of challenge activities and success metrics for interpolation section.

Number (activity level)	Response Type	Question Statement Summary	% correct	Median attempts before correct
1.1	Multiple choice	Interpolation or extrapolation	82	1
1.2	Number	Linear interpolation calculation	76	2
1.3	Number	Linear interpolation calculation	75	2
1.4	Number and formula	Linear interpolation calculation and formula	75	2
1.5	Number and formula	Linear interpolation calculation and formula	73	1
2.1	Number and formula	Linear interpolation calculation and formula	58	2
2.2	Number and formula	Linear interpolation calculation and formula	56	2
2.3	Number	Third linear interpolation calculation	55	1
2.4	Formula	Double interpolation calculation and formula	55	2
2.5	Number	Double interpolation calculation	44	1

Conclusion

Spreadsheets are a common tool for both engineering students and engineers. The popularity of spreadsheets has spurred the development of many different tools; however, many features, formulas, and functions are independent of the program, platform, or device. Therefore, an educational tool using interactivity has been developed and deployed to leverage tenets of cognitive load theory and active learning. Specifically, *Spreadsheets Essentials* from zyBooks enabled us to measure reading participation, animation views, and success on over 100 auto-graded problems with randomized numbers and content. Reading rates of 100% were observed for at least 75% of students across all 12 sections in 2018 as well as animation view rates as high as 118%. Success on challenge activities was higher for basic skills and formulas than more advanced topics, such as integration and interpolation. With new sections being added for the 2019 cohort, this interactive tool will continue to yield new information about spreadsheet education for students and faculty in the future.

Acknowledgments

The author thanks recent contributions from Katherine Roach, zyBooks, and teaching assistants. This work was completed within the framework of University of Toledo IRB protocol 201808.

Disclaimer

One of the authors may receive royalties from sales of the zyBook detailed in this paper.

Bibliography

- [1] D. E. Clough, "Use Spreadsheets for ChE Problem-Solving," (in English), *Chemical Engineering Progress*, vol. 112, no. 8, pp. 25-34, Aug 2016.
- [2] J. Hawkins and S. Blakeslee, *On intelligence*. St. Martin's Griffin, 2005.
- [3] M. T. Chi, "Active-constructive-interactive: a conceptual framework for differentiating learning activities," *Topics in Cognitive Science*, vol. 1, no. 1, pp. 73-105, Jan 2009.
- [4] A. W. Chickering and Z. F. Gamson, "Seven Principles for Good Practice in Undergraduate Education," *AAHE Bulletin*, pp. 1-7, 1987.
- [5] S. Freeman *et al.*, "Active learning increases student performance in science, engineering, and mathematics," *Proceedings of the National Academy of Sciences*, vol. 111, no. 23, pp. 8410-8415, 2014.
- [6] R. M. Felder and R. Brent, *Teaching and Learning STEM: A Practical Guide*. San Francisco, CA: Jossey-Bass, 2016.
- [7] A. Edgcomb and F. Vahid, "Effectiveness of Online Textbooks vs. Interactive Web-Native Content," in *ASEE Annual Conference*, Indianapolis, IN, 2014.
- [8] A. Edgcomb, F. Vahid, R. Lysecky, A. Knoesen, R. Amirtharajah, and M. L. Dorf, "Student Performance Improvement using Interactive Textbooks: A Three- University Cross-Semester Analysis," in *ASEE Annual Meeting*, Seattle, WA, 2015.
- [9] D. Zax, "Learning in 140-Character Bites," *ASEE PRISM*, pp. 1-3, 2009.
- [10] D. Tapscott, *Grown Up Digital: How the Net Generation is Changing Your World*. New York: McGraw-Hill, 2009.
- [11] C. S. Lee, N. J. McNeill, E. P. Douglas, M. E. Koro-Ljungberg, and D. J. Therriault, "Indispensable Resource - A Phenomenological Study of Textbook Use in Engineering Problem Solving," *Journal of Engineering Education*, vol. 102, no. 2, pp. 269-288, 2013.
- [12] American Enterprise Institute. (2016, March). *Price changes selected consumer goods and services 1996-2016* [Online]. Available: <https://www.facebook.com/aei/photos/a.290944073957.147277.44951363957/10155200085953958/?type=3&theater>.
- [13] E. H. Hobson. (2004, December). *Getting Students to Read: Fourteen Tips* [Online]. Available: http://www.ideaedu.org/Portals/0/Uploads/Documents/IDEA%20Papers/IDEA%20Papers/Idea_Paper_40.pdf.
- [14] University of Indiana College of Education. (2018, January). *National Survey of Student Engagement - Question 1c. During the current school year, about how often have you done the following? Come to class without completing readings or assignments* [Online]. Available: http://nsse.indiana.edu/html/summary_tables.cfm.
- [15] B. D. Brost and K. A. Bradley, "Student Compliance with Assigned Reading: A Case Study," *Journal of Scholarship of Teaching and Learning*, vol. 6, no. 2, pp. 101-111, 2006.
- [16] P. Marshall, "How much, how often?," *College and Research Libraries*, vol. 35, no. 6, pp. 453-456, 1974.
- [17] C. M. Burchfield and T. Sappington, "Compliance with required reading assignments," (in English), *Teaching of Psychology*, vol. 27, no. 1, pp. 58-60, Win 2000.

- [18] T. Berry, L. Cook, N. Hill, and K. Stevens, "An Exploratory Analysis of Textbook Usage and Study Habits: Misperceptions and Barriers to Success," *College Teaching*, vol. 59, no. 1, pp. 31-39, 2010.
- [19] J. S. Yuen, A. Edgcomb, and F. Vahid, "Will Students Earnestly Attempt Learning Questions if Answers are Viewable?," in *ASEE Annual Meeting*, New Orleans, LA, 2016, p. 16595.
- [20] A. Edgcomb and F. Vahid, "How Many Points Should Be Awarded for Interactive Textbook Reading Assignments?," in *45th Annual Frontiers in Education Conference (FIE)*, El Paso, TX, 2015, pp. 1-4.
- [21] A. G. Carney, S. Winstead Fry, R. V. Gabriele, and M. Ballard, "Reeling in the big fish: Changing pedagogy to encourage the completion of reading assignments," *College Teaching*, vol. 56, no. 4, pp. 195-201, 2008.
- [22] R. E. Landrum, R. A. R. Gurung, and N. Spann, "Assessments of Textbook Usage and the Relationship to Student Course Performance," *College Teaching*, vol. 60, no. 1, pp. 17-24, 2012.
- [23] D. Lepek and M.-O. Coppens, "Nature-Inspired Chemical Engineering: Course Development in an Emerging Research Area," in *ASEE Annual Meeting*, New Orleans, LA, 2016, p. 15218.
- [24] E. Mazur, G. King, B. Lukoff, and K. Miller. (2018, August). *Perusall* [Online]. Available: <http://perusall.com/>.
- [25] M. W. Liberatore, "High textbook reading rates when using an interactive textbook for a Material and Energy Balances course," *Chemical Engineering Education*, vol. 51, no. 3, pp. 109-118, 2017.
- [26] M. W. Liberatore, "Reading analytics and student performance when using an interactive textbook for a material and energy balances course," in *ASEE Annual Conference & Exposition*, Columbus, OH, 2017, pp. 1-13.
- [27] M. W. Liberatore, "Quantifying reading and online homework completion using an interactive material and energy balances textbook," presented at the AIChE Annual Meeting Minneapolis, MN, October 2017, 2017.
- [28] A. S. Bowen, D. R. Reid, and M. D. Koretsky, "Development of interactive virtual laboratories to help students learn difficult concepts in thermodynamics," *Chemical Engineering Education*, vol. 49, no. 4, pp. 229-238, 2015.
- [29] E. D. Sloan and C. Norrgran, "A neuroscience perspective on learning," *Chemical Engineering Education*, vol. 50, no. 1, pp. 29-37, 2016.
- [30] F. Paas, A. Renkl, and J. Sweller, "Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture," *Instructional science*, vol. 32, no. 1, pp. 1-8, 2004.
- [31] M. W. Liberatore and K. Roach, "Building Spreadsheet Skills Using an Interactive Textbook," in *ASEE Annual Meeting*, Salt Lake City, UT, 2018, pp. 1-12.
- [32] Wistia. (2016, January). *How long should a video be?* [Online]. Available: <https://wistia.com/blog/optimal-video-length>.