The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
ASEE
Paper ID #41684

Effects of problem type on completion and attempts on auto-graded homework problems for Material and Energy Balances

Samantha Yanosko, University of Toledo
Grant Valentine, University of Toledo

Grant Valentine is an undergraduate student studying chemical engineering at the University of Toledo expecting to graduate in 2025.

Prof. Matthew W Liberatore, University of Toledo

Matthew W. Liberatore is a Professor in the Department 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. From 2005 to 2015, he served on the faculty at the Colorado School of Mines. In 2018, he served as an Erskine Fellow at the University of Canterbury in New Zealand. His research involves the rheology of complex fluids, especially traditional and renewable energy fluids and materials, polymers, and colloids. His educational interests include developing problems from YouTube videos, active learning, learning analytics, and interactive textbooks. His interactive textbooks for Material and Energy Balances, Spreadsheets, and Thermodynamics are available from zyBooks.com. His website is: https://www.utoledo.edu/engineering/chemical-engineering/lib

Effects of problem type on completion and attempts on auto-graded homework problems for Material and Energy Balances

Abstract

Auto-graded online homework and interactive textbooks engage students and generate big data. Several new research questions investigate students' usage of and success on over 700 auto-graded questions within an interactive tool titled the Material and Energy Balances zyBook. Auto-grading occurs in real time, so students, teaching assistants, and faculty can see progress without waiting for assignments to be graded. Previous research examined reading participation and auto-graded problems at the course level; Findings included median reading participation over 93% for seven cohorts and median correct on auto-graded problems of 91% or higher for six cohorts. More specifically, auto-graded problems allowed unlimited attempts, so students received feedback and persisted until correct for these randomized problems. Here, two recent cohorts' responses on hundreds of auto-graded questions examined specific types of auto-graded problems. From one perspective, formative, single calculation problems with scaffolding appeared in most sections, while more summative, multi-concept problems appeared at the end of each chapter. From another perspective, many problems required numerical answers within a tolerance, and other problems were multiple choice. Our research questions examined these different types and locations of auto-graded problems. New findings showed that median percent correct was high (above 80%) for all problem types. Attempts before correct provided a valuable metric to distinguish between problem types with numeric problems taking more attempts than multiple choice. Finally, a metric combining both correct and attempts, called the deliberate practice score, provided another quantitative aggregate measure. Of note, end-of-chapter numeric response problems had a much larger fraction of problems at higher deliberate practice scores than in-chapter, numeric questions.

Introduction

Interactive technologies in engineering education are creating big data that can be used to measure student engagement and learning. Clicks are one way to capture interactivity from stepping through an animation or multi-step example to correctly answering multiple choice or true and false questions. However, online homework provides more advanced metrics to capture student's problem-solving skills. Numerous platforms have been in development for years with math and science courses having more tools and options than engineering courses [1-4].

Online homework is synonymous with terms such as auto-graded problems, and these tool have become common throughout science, technology, engineering, and mathematics (STEM) [5-9]. By providing instantaneous feedback to students, auto-graded problem capture some of the most important tenets of deliberate practice [10-13]. Other learning frameworks also describe the creation of these problems as well as their use by students, e.g., scaffolding, randomization, and chunking [14, 15]. Some limitations come with every technology, and online homework is not a perfect tool for learning engineering. Auto-grading usually is normally limited to multiple choice and numeric answers, which are generally algorithmic. Thus, other problem types common to engineering, including conceptual, drawing, and graphical problems, have seen more limited study. Here, auto-graded problems will be examined using multiple metrics, both by type and location, which will expand the knowledge in this understudied area.

Many contributions in chemical engineering consider interventions in a single course, which is true here also. However, subject matter expertise is outside the primary scope of the findings related to auto-graded problems. Specifically, the auto-graded problems are part of an interactive textbook for a Mass and Energy Balances (MEB) course. As the first chemical engineering course in most curricula, student develop engineering problem solving skills related to non-reacting and reacting processes as well as multi-unit operations. In some cases, interventions for this course were inspired by high attrition rates [16], so the literature contains many novel pedagogies and projects, e.g., [2, 17-19], but a thorough review is not relevant to this research.

This contribution's research questions explore types of auto-graded questions, multiple metrics related to correctness and attempts, and problem difficulty via the deliberate practice score. The types of problems vary in several distinct ways. First, the location of the problems are either in-chapter with primary content or at the end of each chapter. Second, numeric answers within tolerances or multiple choice are the two primary problem types being compared. Finally, two specific multiple-choice question types are investigated: vocabulary and concept questions. Examining these various problem types and locations leads to the following research questions.

How does correctness, attempts, and problem difficulty vary:

1. for numeric problems based on location - either comprehensive, end-of-chapter questions or more formative in-chapter questions?
2. for multiple choice problems based on location - either comprehensive, end-of-chapter questions or more formative, in-chapter questions?
3. between multiple choice question types – either concept questions or vocabulary?

Materials and Methods

The *Material and Energy Balances zyBook* is a fully interactive textbook intended for a first course in chemical engineering [20]. The student-centered reading experience and online homework have been presented in many previous settings, so only a brief synopsis is provided. As of January 2024, 157 animations, over 1400 clicks to complete reading participation, and 737 online homework problems are included across 9 chapters. Reading participation encourages weekly involvement by awarding 5% of the course grades for completing the interactive exercises before the due date. Reading participation has been discussed in previous publications and is outside the scope of this contribution [21-24].

The auto-graded problems are called challenge activities in the zyBook and also contribute 5% of the final course grade. While reading participation is an effort-based grade, challenge activities are machine graded as correct or incorrect. As these problems are considered formative assessment, so no limit on the number of attempts is given. The numbers and/or content changes after each incorrect attempt. Tolerances vary by question and complexity of the steps to find the final answer or answers; Tolerances are adjusted by the book's author, but rarely (<1% of problems) and usually in the first year of use. For challenge activities, 15 problems (2-3% of the total assigned problems) are forgiven when calculating students' grades. While this forgiveness factor can reduce anxiety when a student is stuck on auto-graded problems [25, 26], the raw, uncorrected fraction correct (%) are presented here.

Students in two cohorts taught at a medium-size, Midwest public university by one of the authors will be examined. Students completing the course, i.e., not withdrawing during the semester, are included in the two cohorts of interest. The 2021 cohort was taught online synchronously (n=66 students), and the 2022 cohort was taught in-person (n=57 students). Discussions of gender distribution and student success was examined in previous publications [21, 22], and further investigation of the diversity of the students is outside the scope of this work.

Auto-graded problems exist in many contexts in higher education, including multiple choice, true and false, and numeric entry. In addition to the type of student response, the problem style in engineering education, and specifically in the MEB zyBook, also varies (Table 1). Two problem locations are common within the MEB zyBook, i.e., in chapter and end of chapter. Two primary types of answers can be requested of the students. First, multiple choice uses drop-down menus with a limited number of choices (usually 3 to 6). Second, numeric problems provide an answer box for students to enter a number within a tolerance; All numeric answers, usually 1 to 4, need to be entered simultaneously to answer a problem correctly. Further details and example problems are provided next.

Table 1. Number and type of auto-graded problems in the Material and Energy Balances zyBook studied for each cohort.

Total Chapters 1-9	In chapter numeric	End of chapter numeric	In chapter multiple choice	End of chapter multiple choice	End of chapter vocabulary	End of chapter concept
712	224	157	56	126	43	100

Many in-chapter, auto-graded problems in the MEB zyBook are scaffolded exercises focusing on one primary course concept and a single numeric answer. For example, the 3rd of 5 levels on the topic of extent of reaction focused only on the extent of reaction without any additional systems or balances (Figure 1 Top). Another in-chapter, scaffolded problem type is multiple choice. Identifying the limiting reactant allows for changing the chemical reaction and/or the given flow rates to create dozens of versions of a single, auto-graded question (Figure 1 Bottom).

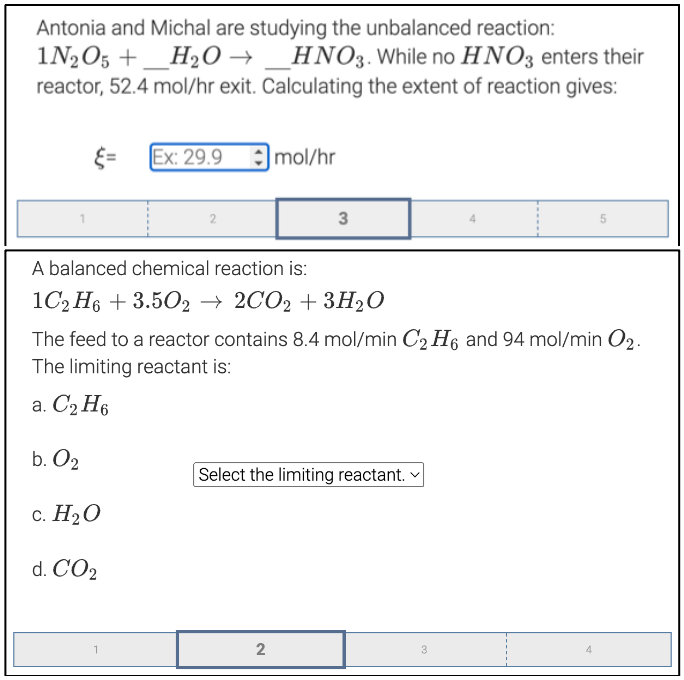

Figure 1. Examples of auto-graded problems. Top. Single numeric response for an in-chapter, scaffolded problem. Bottom. Multiple choice response for an in-chapter, scaffolded problem. Numbers below problems represent the current question level (bold) and total question levels within the challenge activity.

Alternatively, end-of-chapter questions involve more complicated systems and solutions, i.e., writing and solving multiple balances and extra equations. Applying the extent of reaction for multiple reactions is the way an end-of-chapter problem related to ethanol is solved (Figure 2 Top). Here, balancing chemical reactions, extent of reaction, yield, and other concepts are needed to solve for the three auto-graded, numeric answers. All three numeric responses are entered simultaneously, and all numbers have to be within the tolerance for the problem to be completed correctly. Accompanying many of the longer end-of-chapter problems are concept questions that are auto-graded with three choices in most cases, i.e., increase, decrease, or stay the same. These concept questions ask the student to re-examine their solution from the previous level and change one of the initial values (Figure 2 Bottom). The concept questions then qualitatively ask how or if the new initial value alters one of the just-calculated answers.

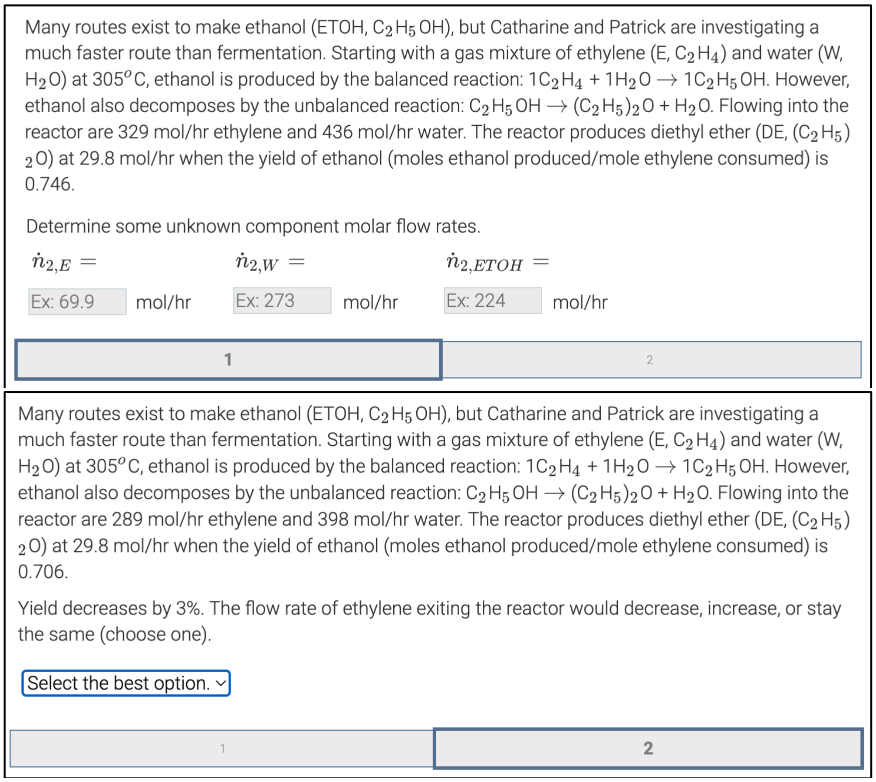

Figure 2. Examples of auto-graded problems. Top. Multiple numeric response for an end-of-chapter problem. Bottom. Multiple choice response for concept question as part of an end-of-chapter problem. Numbers below problems represent the current question level (bold) and total question levels within the challenge activity.

Finally, vocabulary problems are multiple-choice questions that quiz students' knowledge of new terms introduced within a section or chapter. These questions would map to the lowest level of Bloom's taxonomy (remembering).

Since the interactive textbook is different than most engineering textbooks, some context on use of the interactive textbook is provided here. The course meets three times per week (Monday, Wednesday, Friday) with a normal assignment schedule involving reading participation due on Mondays, in-chapter challenge activities due on Wednesdays, and static problems due on Fridays. Static problems are called zyExercises; static versions of challenge activities as well as YouTube problems are sometimes included for the Friday problem sets. Previous work discussed end-of-chapter and YouTube problems in detail [27, 28]. Students electronically scan and submit hand-written work for the Friday problems. In addition, some end-of-chapter auto-graded problems are assigned before each of three midterm exams [17]. Since many of the end-of-chapter problems are from the author's previous exams, these problems are very representative of exam problems.

The term learning analytics captures the research presented here with several data types quantifying student usage and behavior. The total number of available auto-graded problems was equivalent for these two cohorts. While the specific end-of chapter problems that were assigned for a grade varied, the number of problems was similar for both cohorts. The types and styles of these end-of-chapter problems were presented previously [27, 28] and will be detailed as needed in the results below.

Problem difficulty was examined using a metric called the deliberate practice (DP) score, which was detailed in a previous publication [10]. In brief, four metric are combined into the DP score. First, the fraction correct at the due date translates to 0 points for ≥90%, 1 point for 80 - 90%, and 2 points for < 80%. Next, modified correct accounts for the possibly diminishing number of students attempting subsequent levels due to scaffolding, i.e., for multi-level problems students cannot attempt a subsequent level without correctly completing the previous level. Modified correct at the due date is scored 0 points for ≥95%, 1 point for 85 - 95%, and 2 points for < 85%. The third and fourth metrics examine attempts before correctly answering a question the first time. Here, the 1st quartile and median attempts before correct capture struggle for 25% and 50% of a cohort. Specifically, attempts before correct translates to DP score with the following thresholds: 0 points for ≤ 2, 1 point for 2 – 3 attempts, and 2 points for ≥ 3. Thus, DP score is the sum of the four metrics leading to scores ranging from 0 to 8.

One or a small number of outliers can alter mean values, so box plots capture a broader view of the cohorts. By documenting the middle 50%, including the 1st quartile, median, and 3rd quartile, comparisons can be made that capture the behaviors of a large fraction of a cohort. In addition, skewness can be observed by also presenting the means.

Results and Discussion

Answering the research questions using tens of thousands of attempts on auto-graded homework problems will explore problem type and location. Specifically, multiple choice and numeric problems are placed in most sections as well as at the end of each chapter. Also, multiple-choice problems comparing concept and vocabulary questions will be investigated to complete this contribution.

> Research Question1: How does correctness, attempts, and problem difficulty vary for numeric problems based on location - either comprehensive, end-of-chapter questions or more formative in-chapter questions?

With unlimited attempts without penalty, students can continue to re-try problems. However, for problems with numeric solutions, the chance of randomly entering one or more correct answers simultaneously is small since, at a minimum, numbers change in the problem statement with each attempt. Here, the focus is on numeric entry problems in various contexts in the book, either in-chapter or end-of-chapter,. While both in-chapter, scaffolded and end-of-chapter problems were discussed in previous contributions [27-30], a shift from examining all problems in aggregate to the influence of problem type is new here. Additionally, different cohorts of students provide new data sets to add to or reproduce established research.

First, students answered auto-graded, numeric problems correctly at a high rate, which is likely related to the unlimited attempts mentioned earlier. When combining two cohorts, the median percent correct for numeric problems was similar between in-chapter (80%) and end-of-chapter

(82%) (Figure 3). Thus, the top half of the class performed similarly on the single concept in-chapter questions and end-of-chapter, multi-concept questions. Next, the 1st quartile was another useful metric when examining percent correct, since three quarters of the cohort was captured. For 1st quartile percent correct, the in-chapter (71%) was measurably larger than the end-chapter questions (50%). Here, the differences in the quartile below the median were noteworthy and served as an indicator of students who may scores lower on quizzes and exams. This type of observation was not possible without taking the broader, more inclusive lens available when using box plots (instead of the common mean and standard deviation). While t-test can be performed and very low p-values were measured, the differences in the distributions as noted by the quartiles' discussion limit the utility of t-tests.

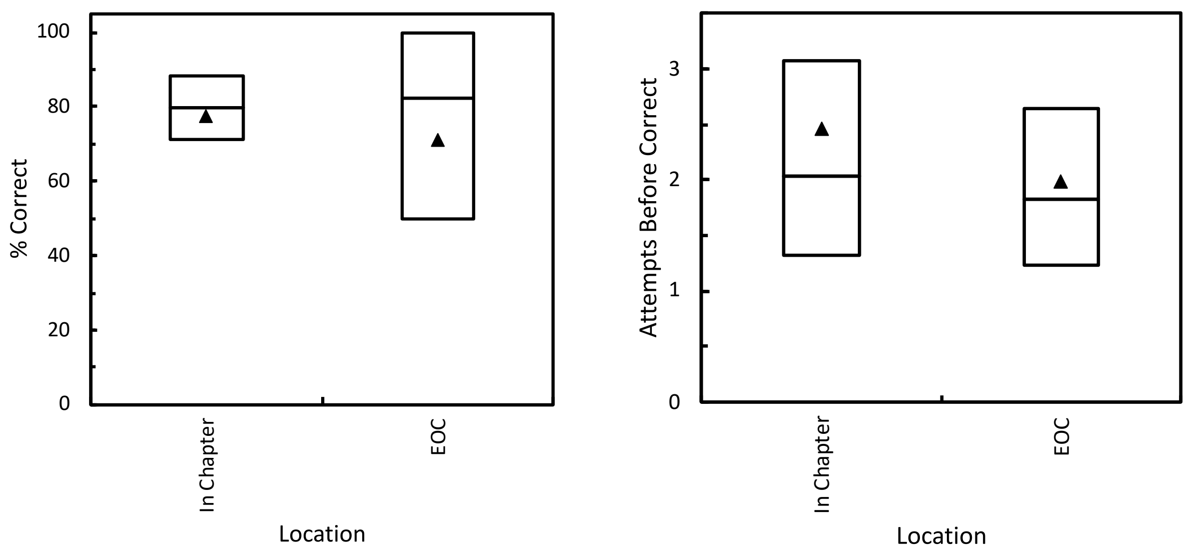

Figure 3. Percent Correct (left) and Attempts Before Correct (right) for Numeric questions located within chapters (n=448) or at the end of a chapter (n=313).

Quantifying students' persistence with auto-graded problems having unlimited attempts provided a metric beyond right and wrong. Similar to percent correct, the median attempts before correct were similar for in-chapter (2.0) and end-of-chapter problems (1.8) (Figure 3). While the quartile below the median was lower for percent correct on end-of-chapter problems, the opposite effect was observed with attempts. A higher 3rd quartile attempts before correct was observed for in-chapter, numeric problems (3.1) compared to end-of-chapter, numeric problems (2.7). This finding implied that students are less likely to continue on end-of-chapter problems requiring more effort, i.e., involving multiple balances, extra equations, or systems to solve. Additionally, a baseball analogy may be apt here, i.e., three strikes and you're out seemingly applied for a fraction of numeric problems. The raw correct and attempts presented one perspective on student success on auto-graded problems, but the difficulty of numeric problems merits further discussion, which is presented next.

The deliberate practice score, or DP score, was introduced in the methods section as an aggregate of four metrics. While this metric was examined for in-chapter questions for other cohorts [10], quantifying end-of-chapter DP scores, i.e., difficulty, is new here. Lower DP scores were related to high fraction correct as well as low numbers of attempts before correct (Figure 4). Therefore, low DP score questions were easier for students to solve. While DP scores can range from 0 to 8, in-chapter questions scored between 0 and 4 while end-of-chapter questions ranged from 0 to 5. The difference in DP score between the two locations of problems is significant. For example, 83% of in-chapter problems versus 33% of end-of-chapter problems had 0 or 1 DP scores. Thus, 67% of the end-of-chapter problems have DP scores of 2 to 5, which could be interpreted as mild to moderately challenging. Such significant differences in DP scores are not observed for multiple choice questions, which are presented in the next section.

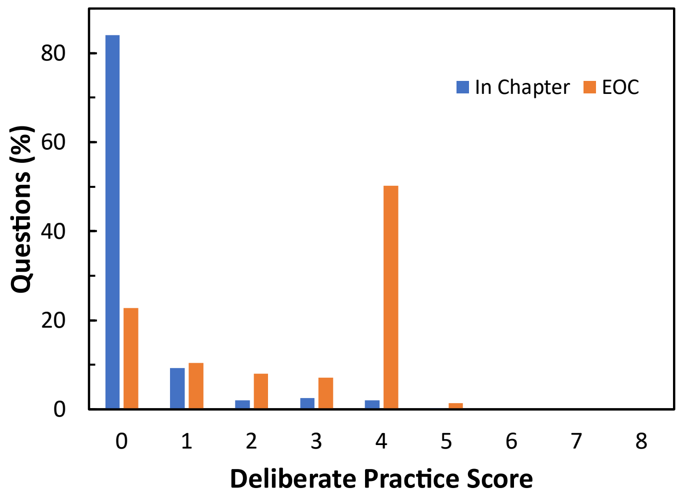

Figure 4. Deliberate practice score for numeric questions located within chapters (n=448) or at the end of a chapter (n=313).

> Research Question 2: How does correctness, attempts, and problem difficulty vary for multiple choice problems based on location - either comprehensive, end-of-chapter questions or more formative in-chapter questions?

Multiple choice problems may be the most common type of auto-graded exercise in engineering education. Here, percent correct was lower (86% median) for in-chapter compared to end-of-chapter problems (100% median) (Figure 5). With a limited number of answer choices, usually 5 or less, and unlimited attempts, very high percent correct was not surprising. While the percent correct showed differences, the attempts before correct were very similar (1.6 median for in-chapter and 1.5 median for end-of-chapter). Extending the baseball analogy here, three strikes are rarely encountered by students attempting multiple choice problems. Specifically, less than a quarter of multiple-choice questions require 2 attempts before answering correctly (1.9 for 3rd quartile in-chapter and 1.8 for 3rd quartile end-of chapter). More thoroughly investigating types of multiple-choice problems may provide better insights on problem type and student behaviors, which is the topic of the next research question.

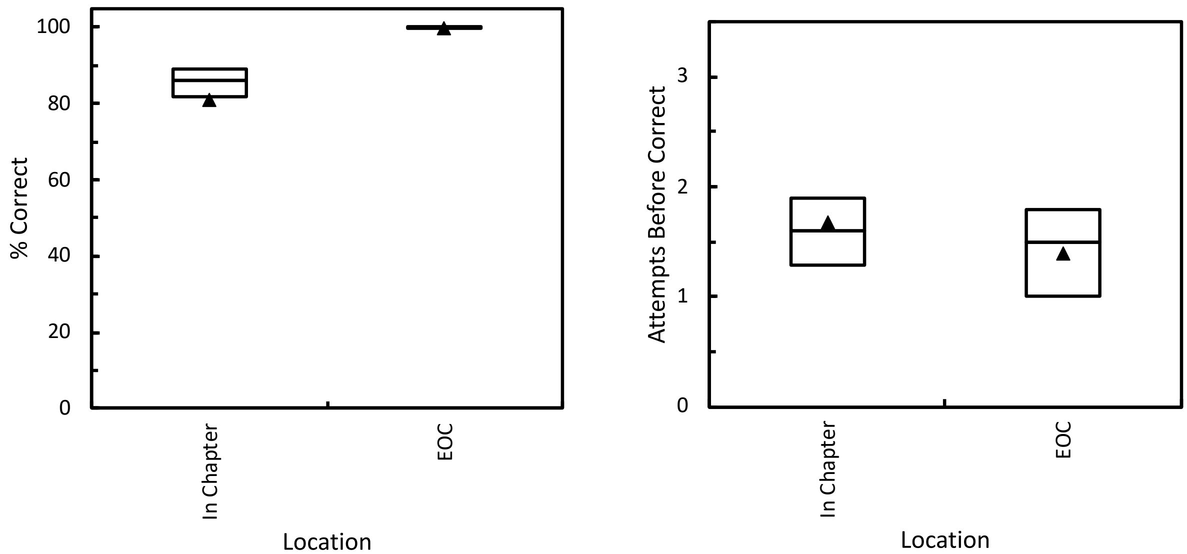

Figure 5. Percent Correct (left) and Attempts Before Correct (right) for multiple choice questions located within chapters (n=112) or at the end of a chapter (n=251).

The DP score allowed for comparison of problem difficulty for multiple-choice problems also. DP scores ranged only from 0 to 2 for multiple choice problems (Table 2), which was a much smaller range than the numeric problems discussed earlier (Figure 4) or another previous study [10]. Here, almost no difference between in-chapter and end-of-chapter multiple choice questions was found. Specifically, 100% of in-chapter and 99% of end-of-chapter problems had a DP score of 0 or 1. Thus, multiple choice questions with unlimited attempts register as easy problems using the DP score framework. Diving deeper into concept and vocabulary multiple choice questions may provide further insights, which are addressed next.

Table 2. Fraction of multiple-choice questions at each deliberate practice (DP) score.

DP Score	In Chapter (%)	EOC (%)
0	99	90
1	1	8
2	0	1
3 - 8	0	0

Research Question 3: How does correctness, attempts, and problem difficulty vary between multiple choice question types – either concept questions or vocabulary?

Multiple choice may be the most studied problem type in education due to its ubiquity and ability to quickly machine grade large numbers of responses, which has been true to decades [31, 32]. While many multiple-choice questions focus on the lowest levels of Bloom's taxonomy, namely remembering, another type of multiple-choice question has been implemented more prominently in more recent years, i.e., the concept question. Concept questions usually fit more into the higher Analyze level of Bloom's taxonomy. In chemical engineering concept questions and inventories have been used in many courses [33-37].

Vocabulary problems match definitions and terms, which usually have 3 to 6 choices. Both terms and choices change when a student answers incorrectly. These problems focus on the remember, or lowest, level of Bloom's taxonomy. Thus, observing 100% correct for at least 75% of these problems is not surprising (Figure 6). Similarly, the attempts before correct were relatively small at 1.6 and 1.9 median attempts before correct for the 3rd quartile. By comparison, concept questions normally only have three answer choices, commonly increase, decrease, or stay the same. These concept questions were normally connected to the multi-concept, end-of-chapter numeric problems and proposed one change to the problem statement that led to an increase, decrease, or no change to another variable, such as a concentration or flow rate. With just three choices, measuring 100% correct for at least 75% of these problems matches the vocabulary questions. However, a broader range of attempts before correct was observed for concept questions compared to vocabulary. Specifically, the 1st quartile is close to 1.0 attempts before correct. Thus, a quarter of the concept questions are immediately answered by students, which would contradict their comfort level in answering conceptual questions - based on informal class feedback over many years or previous research [38, 39].

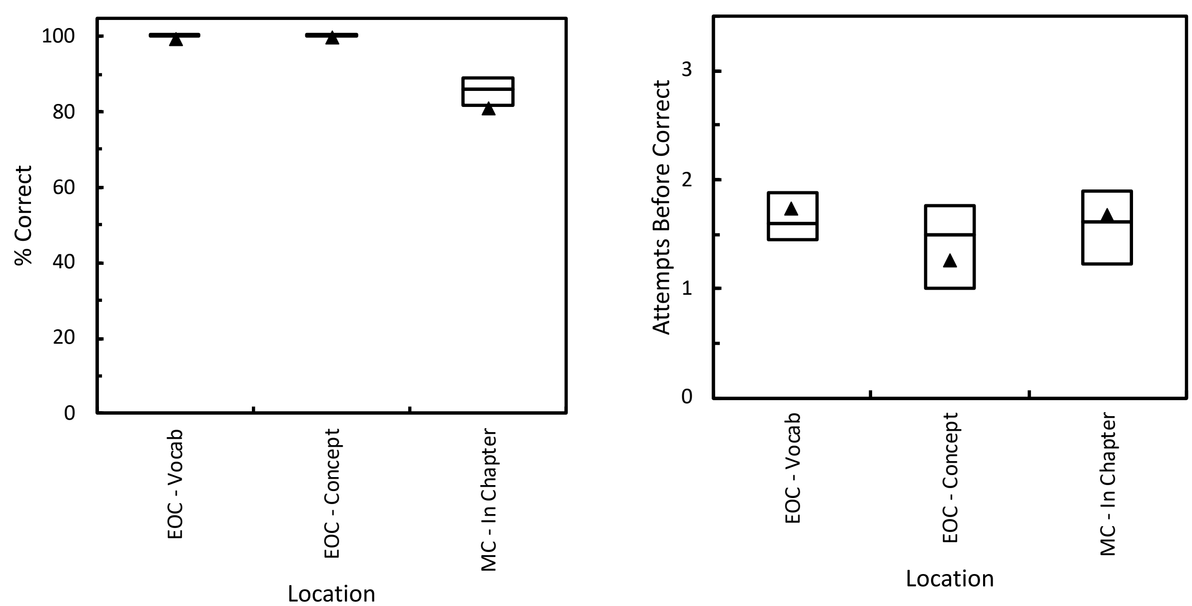

Figure 6. Percent Correct (left) and Attempts Before Correct (right) for end-of-chapter vocabulary (n=86), end-of-chapter concept (n=200), and all in-chapter multiple choice questions (n=112).

Concept and vocabulary problems can be compared to other multiple-choice questions, which were the in-chapter, single concept type. While the percent correct was lower (86% median), the attempts before correct for the middle 50% of students was quite similar (Figure 6). Therefore, no significant differences between vocabulary and concept multiple choice questions was observed for two metrics, namely percent correct and attempt before correct.

Applying the DP score metric found small differences between concept and vocabulary questions. Specifically, 93% of vocabulary questions had a DP score of 0 with the rest scoring 1 (6%) or 2 (1%). For concept questions, 89% of questions had a DP score of 0, 9% score of 1, and 1% score of 2. Therefore, despite having only three choices for concept questions, the 10-11% of question registering a DP score of 1 or 2 is measurably higher than vocabulary or in-chapter questions, 7% and 1%, respectively. Overall, including attempts data to assess multiple choice questions with multiple or unlimited attempts appears to be a promising method.

Conclusion

Auto-graded problems used as formative homework exercises provided rich data sets relating to the fraction of students answering correctly as well as their number of attempts before answering correctly the first time. This study focused on problems in the Material and Energy Balances zyBook, which is a fully interactive textbook. Different problem types and locations were examined for 2 cohorts involving over 100 total students. When offering a small grade incentive, 5% of the overall course grade, the median correct was between 80 and 100% for all types of problems examined. The specific types of problems investigated included: in-chapter numeric, in-chapter multiple choice, end-of-chapter numeric, and end-of-chapter multiple choice. In addition, two types of multiple-choice questions were explored further, namely multiple-choice questions related to either vocabulary or concept questions. Overall, multiple choice showed measurably smaller attempts before correct than numeric entry, which is likely related to the small and finite number of choices for multiple choice compared to any number being a possible correct answer (within a tolerance) for numeric problems. Finally, problem difficulty was captured via a deliberate practice score. The most notable difference in deliberate practice score was between in-chapter and end-of-chapter numeric problems. The end-of-chapter numeric problems had a significantly larger fraction of problems correspond to higher deliberate practice scores.

Overall, the Material and Energy Balances zyBook is configurable tool to serve as the primary text for the first course in chemical engineering. The authors hope other instructors will provide their students with the perpetual, randomized practice that the challenge activities provide, which aligns with many of the best practices of the deliberate practice learning framework.

Acknowledgments

The authors thank assistance from Alex Edgcomb and numerous teaching assistants. This work was completed within the framework of University of Toledo IRB protocols 202214 and 201808.

Disclaimer

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

References

1. Edgcomb, A., 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*. 2015. Seattle, WA. DOI: https://doi.org/10.18260/p.24760
2. Chapman, K.E., M.E. Davidson, and M.W. Liberatore, *Student success and attempts on auto-graded homework across multiple cohorts in material and energy balances.* Chemical Engineering Education, 2021. **55**(1): p. 43-50. DOI: https://doi.org/10.18260/2-1-370.660-123169
3. Kortemeyer, G., *It's All in the Data - But What is It? Learning Analytics and Data Mining of Multimedia Physics Courses.* International Journal of Physics and Chemistry Education, 2019. **11**(1): p. 13-17. DOI: https://doi.org/10.12973/ijpce/17111
4. Kortemeyer, G., E. Kashy, W. Benenson, and W. Bauer, *Experiences using the open-source learning content management and assessment system LON-CAPA in introductory physics courses.* American Journal of Physics, 2008. **76**(4): p. 438. DOI: https://doi.org/10.1119/1.2835046
5. Richards-Babb, M., R. Curtis, Z. Georgieva, and J.H. Penn, *Student Perceptions of Online Homework Use for Formative Assessment of Learning in Organic Chemistry.* Journal of Chemical Education, 2015. **92**(11): p. 1813-1819. DOI: https://doi.org/10.1021/acs.jchemed.5b00294
6. Elias, A.L., D.G. Elliott, and J.A.W. Elliott, *Student perceptions and instructor experiences in implementing an online homework system in a large second-year engineering course.* Education for Chemical Engineers, 2017. **21**: p. 40-49. DOI: https://doi.org/10.1016/j.ece.2017.07.005
7. Lee, J.S. and J. Verrett. *Webwork as an open online homework system in material and energy balances*. in *Canadian Engineering Education Association (CEEA18) Conference*. 2018. DOI: https://doi.org/10.24908/pceea.v0i0.13047
8. Steif, P.S. and A. Dollar, *Study of usage patterns and learning gains in a web-based interactive static course.* Journal of Engineering Education, 2009. **98**(4): p. 321-333. DOI: https://doi.org/10.1002/j.2168-9830.2009.tb01030.x
9. Liberatore, M.W., *Improved student achievement using personalized online homework for a course in material and energy balances.* Chemical Engineering Education, 2011. **45**(3): p. 184-190. DOI: https://journals.flvc.org/cee/article/view/122149
10. Chapman, K.E., M.E. Davidson, N. Azuka, and M.W. Liberatore, *Quantifying deliberate practice using auto-graded questions: Analyzing multiple metrics in a chemical engineering course.* Computer Applications in Engineering Education, 2023. **31**(4): p. 916-929. DOI: https://doi.org/10.1002/cae.22614
11. Couch, J.D., J. Towne, and S. Wozniak, *Rewiring Education: How Technology Can Unlock Every Student's Potential*. 2018: BenBella Books
12. Miller, M.D., *Minds Online: Teaching Effectively with Technology*. 2014: Harvard University Press
13. Ericsson, K.A., R.T. Krampe, and C. Tesch-Römer, *The role of deliberate practice in the acquisition of expert performance.* Psychological review, 1993. **100**(3): p. 363. DOI: https://doi.org/10.1037/0033-295X.100.3.363
14. Mintz, S. *Can Technology Make Grading Fairer and More Efficient?* 2019 [cited 2020 11]; Available from: https://www.insidehighered.com/blogs/higher-ed-gamma/can-technology-make-grading-fairer-and-more-efficient.
15. Lang, J.M., *Small Teaching: Everyday Lessons from the Science of Learning*. 2016, San Francisco, CA: John Wiley & Sons

16. Bullard, L.G. and R.M. Felder, *A Student-Centered Approach To Teaching A Student-Centered Material and Energy Balances. 2. Course delivery and assessment.* Chemical Engineering Education, 2007. **41**(3): p. 168-176. DOI: https://journals.flvc.org/cee/article/view/122438
17. Liberatore, M.W., *Active learning and just-in-time teaching in a material and energy balances course.* Chemical Engineering Education, 2013. **47**(3): p. 154–160. DOI: https://journals.flvc.org/cee/article/view/114520
18. Amos, D.A., C.M. Pittard, and K.E. Snyder, *Active learning and student performance in a Material and Energy Balance Course.* Chemical Engineering Education, 2018. **52**(4): p. 277-286. DOI: https://journals.flvc.org/cee/article/view/106894
19. Ford, L., J. Brennan, K. Dahm, D. Silverstein, L. Landherr, C. West, J. Cole, S. Thiel, B. Vaughen, and M. Jamieson. *How We Teach: Material and Energy Balances*. in *2022 ASEE Annual Conference & Exposition*. 2022. DOI: https://peer.asee.org/41893
20. Liberatore, M.W., Material and Energy Balances zyBook. 2022: Zybooks - a Wiley brand. https://www.zybooks.com/catalog/material-and-energy-balances/
21. Liberatore, M.W., K.E. Chapman, and K.M. Roach, *Significant reading participation across multiple cohorts before and after the due date when using an interactive textbook.* Computer Applications in Engineering Education, 2020. **28**(2): p. 444-453. DOI: https://doi.org/10.1002/cae.22210
22. Yanosko, S. and M.W. Liberatore. *Reproducible High Reading Participation and Auto-Graded Homework Completion across Multiple Cohorts when using an Interactive Textbook for Material and Energy Balances*. in *ASEE Annual Conference*. 2023. Baltimore, MD. DOI: https://doi.org/10.18260/1-2--44109
23. Liberatore, M.W. and K. Chapman. *Reading anytime: Do students complete missed readings after the due date when using an interactive textbook for material and energy balances?* in *ASEE Annual Conference*. 2019. Tampa, FL. DOI: https://peer.asee.org/33224
24. Liberatore, M.W., *High textbook reading rates when using an interactive textbook for a Material and Energy Balances course.* Chemical Engineering Education, 2017. **51**(3): p. 109-118. DOI: https://journals.flvc.org/cee/article/view/104416
25. Jiang, X., K. Chen, Z.K. Hansen, and S. Lowe, *A Second Chance at Success? Effects of College Grade Forgiveness Policies on Student Outcomes*. 2021, National Bureau of Economic Research.
26. Martin, C.C., W.C. Newstetter, and J.M. Le Doux, *Inclusion requires a comprehensive understanding of justice.* Journal of Engineering Education, 2019. **108**(4): p. 453-458. DOI: https://doi.org/10.1002/jee.20296
27. Chapman, K. and M.W. Liberatore. *Can I have More Problems to Practice? Student Usage and Course Success Related to Auto-graded, End-of-chapter Problems in a Material and Energy Balances Course*. in *ASEE Virtual Annual Conference*. 2021. DOI: https://doi.org/10.18260/1-2--36779
28. Chapman, K. and M.W. Liberatore. *Can I have More Problems to Practice? Part 2. Student Success Related to Auto-graded, End-of-chapter YouTube Problems in a Material and Energy Balances Course*. in *ASEE Annual Conference*. 2022. Minneapolis, MN. DOI: https://peer.asee.org/40500
29. Gorbett, L.J., K.E. Chapman, and M.W. Liberatore, *Deliberate practice of spreadsheet skills when using copiable, randomized, and auto-graded questions within an interactive*

textbook. Advances in Engineering Education, 2022. **10**(2): p. 49-79. DOI: https://doi.org/10.18260/3-1-1153-36028
30. Chapman, K.E., M.E. Davidson, N. Azuka, and M.W. Liberatore, *Quantifying deliberate practice using auto-graded questions: Analyzing multiple metrics in a chemical engineering course.* Computer Applications in Engineering Education, 2023. DOI: https://doi.org/10.1002/cae.22614
31. Ch, D.R. and S.K. Saha, *Automatic Multiple Choice Question Generation From Text: A Survey.* IEEE Transactions on Learning Technologies, 2020. **13**(1): p. 14-25. DOI: https://doi.org/10.1109/tlt.2018.2889100
32. Slotta, J.D., M.T.H. Chi, and E. Joram, *Assessing Students Misclassifications of Physics Concepts - an Ontological Basis for Conceptual Change.* Cognition and Instruction, 1995. **13**(3): p. 373-400. DOI: https://doi.org/10.1207/s1532690xci1303_2
33. Friedrichsen, D.M., C. Smith, and M.D. Koretsky, *Propagation from the start: the spread of a concept-based instructional tool.* Educational Technology Research and Development, 2016. **65**(1): p. 177-202. DOI: https://doi.org/10.1007/s11423-016-9473-2
34. Falconer, J., *Combining interactive thermodynamics simulations with screencasts and conceptests.* Chemical Engineering Education, 2015. **50**(1): p. 63-69. DOI: https://journals.flvc.org/cee/article/view/87726
35. Shallcross, D.C., *A concept inventory for material and energy balances.* Education for Chemical Engineers, 2010. **5**(1): p. e1-e12. DOI: https://doi.org/10.1016/j.ece.2009.10.002
36. Miller, R.L., R.A. Streveler, B.M. Olds, and M.A. Nelson. *Concept-Based Engineering Education: Designing Instruction To Facilitate Student Understanding Of Difficult Concepts In Science And Engineering*. in *34th ASEE/IEEE Frontiers in Education Conference*. 2004. Savannah, GA. DOI: https://doi.org/10.1109/FIE.2004.1408673
37. *AIChE Education Division's Concept Warehouse*. [cited 2024 January]; Available from: https://conceptwarehouse.tufts.edu/cw/CW.php.
38. Falconer, J., *Conceptests for a thermodynamics course.* Chemical Engineering Education, 2007. **41**(2): p. 107-114. DOI: https://journals.flvc.org/cee/article/view/122457/
39. Koretsky, M.D. and A.J. Magana, *Using Technology to Enhance Learning and Engagement in Engineering.* Advances in Engineering Education, 2019(Spring): p. 1-53. DOI: https://advances.asee.org/using-technology-to-enhance-learning-and-engagement-in-engineering/