Can Motivated Students Do More Activities?

Arun-Balajiee Lekshmi-Narayanan[1,*], Michael W. Asher[2], Peter Brusilovsky[1] and Paulo F. Carvalho[2]

[1]Intelligent Systems Program, University of Pittsburgh, Pittsburgh, PA
[2]Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA

Abstract

The *"Doer Effect"* is the empirical phenomenon observed as a stronger correlational relationship between students who complete more activities and their course learning outcomes compared to those who complete fewer activities or watch fewer videos. In this paper, we extended prior evidence of a *"Doer Effect"* to investigate how doing more can be related not only to better learning outcomes but also to motivational ones. Specifically, we investigated persistence as the student's willingness to continue working on course activities. We used secondary analyses of data from MOOC that taught Advanced Placement (AP) Introductory Java Programming to high school students using the digital textbook platform RuneStone. Although we failed to identify a doer effect in learning outcomes, our analyses do suggest that completing more activities is related to longer persistence in the course than reading more pages or watching more videos. This effect does not appear to be limited to highly motivated students.

Keywords

Digital Textbooks, Introduction to Programming, JAVA, Doer Effect, Student Motivation

1. Introduction

Several studies have provided empirical evidence from laboratory experiments [1], secondary data analyses [2], and classroom studies [3] that students learn more by practicing than watching videos or reading text. For example, Koedinger and colleagues [4, 5] showed that completing more reading causes students to do about 0.39 times more across 3 online courses. Thus, this indicated that *doing more* was a better predictor of student success at the end of the course than reading more pages or spending more time watching videos, often referred to as the *"Doer Effect"*.

Several researchers have demonstrated the *"Doer Effect"* using many datasets on different domains. Carvalho and colleagues [3] presented this in the context of psychology, computing, and other courses [2]. Hou and colleagues [6] presented this in the context of a digital textbook for E-Learning Design Principles and Methods. Asher and colleagues [7] experimentally demonstrated a causal relationship between the students doing activities and learning better. Researchers confirmed the *"Doer Effect"* using student performance and grades.

Although to date comparisons between doing more and reading/watching more has been limited to its relationship to learning outcomes, recent work by Asher and colleagues suggests that practice-only instructions reduced student motivation to pursue the course in the future [8] and can motivate the students to try harder [1]. So, students' motivation plays a role in their effort and the number of activities they do, beyond students' course performance and grades.

Inspired by the work by Asher and colleagues [8], in the current study we expand research on the *"Doer Effect"*to include motivational outcomes. Specifically, we studied the relationship between doing and motivation by investigating the short- and long-term impacts of practice versus reading/watching videos on students' persistence and interest. We used secondary data analyses to process student

CSEDM'25: 9th Educational Data Mining in Computer Science Education (CSEDM) Workshop, July, 2025, Palermo, Sicily, Italy
[*]Corresponding author.
arl122@pitt.edu (A. Lekshmi-Narayanan); masher@andrew.cmu.edu (M. W. Asher); peterb@pitt.edu (P. Brusilovsky); pcarvalh@cs.cmu.edu (P. F. Carvalho)
https://a2un.github.io/ (A. Lekshmi-Narayanan); https://michaelasher.com/ (M. W. Asher); https://sites.pitt.edu/~peterb/ (P. Brusilovsky); https://www.paulocarvalho.me/ (P. F. Carvalho)
0000-0002-7735-5008 (A. Lekshmi-Narayanan); 0000-0002-1006-8813 (M. W. Asher); 0000-0002-1902-1464 (P. Brusilovsky); 0000-0002-0449-3733 (P. F. Carvalho)

interaction data collected from an online introductory programming course offered in the form of an online digital textbook. We ask the following research questions:

RQ1: Is there a *"Doer Effect"* in the context of a digital textbook for programming?

RQ2: Does "doing" influence the persistence of students in a CS course?

RQ3: Does "doing" influence students' interest in CS?

Our analyses will help build and design better course materials, as well as MOOCs that support student–driven learning. Additionally, studying the relationship between student motivation and the *"Doer Effect"* is an important problem that will encourage adaptive personalized sequencing of course content to prevent less–confident students from being discouraged by students who do more activities and achieve success.

2. Related Work

Several open digital textbook platforms developed and deployed by other researchers allowed us to investigate *"Doer Effect"* on anonymized and openly available log data of MOOCs that blend reading, video watching and doing activities. Since the days of early evolution in adaptive hypermedia, digital textbooks have been developed for programming, such as ELM-ART [9]. Walker and colleagues [10] considered student needs in the design of digital textbook platforms integrated with interactive content. A group of researchers has explored the implementation of digital textbooks in Computer Science Education by integrating interactive learning activities [11]. With the evolution of programming languages used, such as Python, Ericson and colleagues [12] developed new kinds of content for these programming digital textbooks, such as Parson's puzzles, and programming exercises that students can work on in addition to reading through the material. Pollari-Malmi and colleagues [13] investigated the value of adding interactive activities and problems to digital textbooks and found an increase in the usage of digital textbooks by students who persisted until the end of the course.

2.1. The Doer Effect and Digital Textbook Implementation

The *"Doer Effect"* [4] is the empirical observation that completing practice activities is associated with larger post-test gains than reading more or watching more video lectures. The authors categorized the content features of the course into 2 broad categories – passive components that students could read, watch, or study and active components that students could solve like problems such as quizzes, discussion forums that encourage participation, and interactive activities with targeted feedback. The focus of the work was to identify features in the course design that encourage student participation in more activities offered by the course.

Follow-up studies have worked with larger sample sizes (N=1000) over 2 courses and showed that students solving more problems learn better than students who re-read to prepare for better performance at the end of the course [3]. In this work, the authors focused on, in addition to students learning by doing, whether the variability in the practice items also improves their performance at the end of the course. The work also focused on the larger aspects of learning outcomes, such as memory and retention.

Studies have also analyzed and demonstrated the effect in other datasets and at scale [14, 15]. In these systems, the VitalSource platform offers digital textbooks on various topics for students to learn online. The motivating goal of replication research in these works was to design and implement better curricula in the format of digital textbooks with high student learning outcomes. Furthermore, in follow-up to the second original work on causal effects of the doer effect [5], the authors also could confirm at scale the causal effects of students' better learning outcomes from doing more activities over reading.

2.2. Understanding the Doer Effect Using Student Persistence

Often student motivation at the beginning of the course affects the extent to which they engage in practicing by solving problems in a course. Asher et al. [8] presented conclusive evidence that students who receive opportunities for practice with feedback can effectively replace attending lectures by completing more activities. They tested these hypotheses with an in-classroom study and showed a clear interaction between student interest growth in the practice-only condition as opposed to the lecture-only condition. In another follow-up to Koedinger and colleagues [5], Asher and colleagues [7] also show casual connections between students working on more activities as mock-tests with feedback lead to mastery with some mediating effect by repeating test attempts leads to an improvement in their final exam performance, suggesting a causal effect.

3. Dataset

Our study used the data from *CSAwesome* digital textbook, comprised of approximately 16 content units (see Figure 1). Additional units focus on mock exams to prepare students for the Advanced Placement (AP) level exam on JAVA programming. Students took the exam after completing this course. As students progress through the course, they read texts, watch videos, and work on activities (see Figure 1). The system records all student interactions in a DataShop–specific format [16]. In each iteration of the course, the students start by responding to a survey on their confidence in learning JAVA and whether they will pursue a career in computer science, as shown in Figure 2[1].

The current study included data from 1,060 students who interacted with the *CSAwesome* digital textbook in the year 2018 - 2019. The students worked on 2 kinds of activities – Programming Activities and Content Activities. Programming activities included *Parson's problems, CodeLens, and ActivityCode* [17]. In Parson's problems, students moved and rearranged blocks of code to generate the correct solution. In CodeLens, students debugged and analyzed the code. In ActiveCode, students wrote code to solve a problem. Content Activities included Multiple Choice Questions (MCQs) and Short Answer Questions (SAQs).

In addition to these activities, students could watch YouTube videos using links embedded in the course material. The average time spent by students watching these videos was 25.65 minutes (SD = 20.73). Students could refer to the reading material while working with activities. The current dataset is a log of time-stamped student interactions with the digital textbook system. The log records for reading activities include open and scroll events; video interactions include click and pause events; activity interactions encompass short answer responses for SAQs, choice selections for MCQs, solution success or failure along with error feedback; notably for *ActiveCode, Parson's Problems, and CodeLens.*

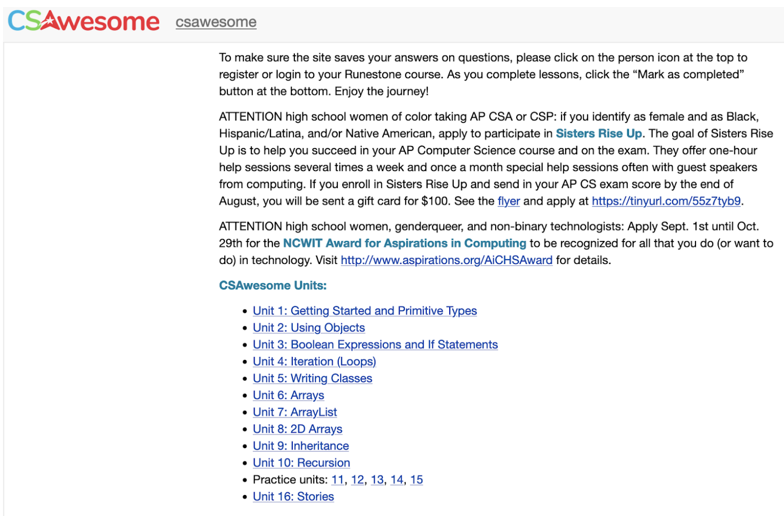
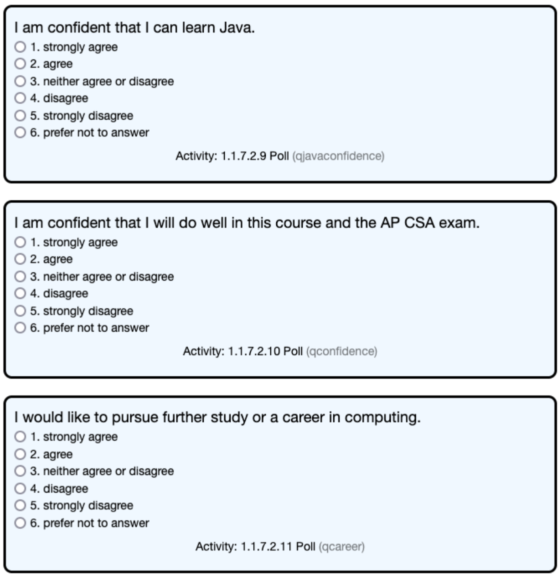


Figure 1: Table of Contents showing the sequence of topics covered in the digital textbook.

[1] https://runestone.academy/ns/books/published/csawesome/Unit1-Getting-Started/survey.html



I am confident that I can learn Java.
1. strongly agree
2. agree
3. neither agree or disagree
4. disagree
5. strongly disagree
6. prefer not to answer
Activity: 1.1.7.2.9 Poll (qjavaconfidence)

I am confident that I will do well in this course and the AP CSA exam.
1. strongly agree
2. agree
3. neither agree or disagree
4. disagree
5. strongly disagree
6. prefer not to answer
Activity: 1.1.7.2.10 Poll (qconfidence)

I would like to pursue further study or a career in computing.
1. strongly agree
2. agree
3. neither agree or disagree
4. disagree
5. strongly disagree
6. prefer not to answer
Activity: 1.1.7.2.11 Poll (qcareer)

Figure 2: Pre–survey questionnaire: in this figure the student has to respond with their confidence to learn JAVA and interest in pursuing a career in CS.

4. Measures and Analyses

4.1. Measures

Preprocessing As a first step, we filtered students' interactions that included pre-test scores. For both pre-and post-tests, we considered the students' first attempts to calculate the scores. We aggregate Multiple sessions of trace log. We winsorized the outliers [2]. The final usable version of the dataset had an effective sample size of 694. The logs contained page views, activity completions, and video-watching events. We then sorted the columns using student IDs and timestamps in that order. After sorting the data by student IDs and timestamps, we took the difference between the timestamp of the current row event and the timestamp of the row for the previous row event, which provided us with the duration of the activity or action in each given row. To address our research questions, we filtered the dataset in 3 ways. 694 students completed the pre-test. Ninety-two students completed both the pre-test and the post-test.

Pre- and Post-tests We collected the first attempts of all students who completed the respective tests. We measured the reliability of the pre- and post-test questions using the Cronbach's α, a metric that estimates using a missing data imputation method. Students' responses to the items in the pre- and post-tests is used to evaluate the reliability. (Cronbach's $\alpha = 0.87$, $\alpha = 0.81$ respectively).

Resource Use To calculate the total time spent in each resource, we summed the time for all rows for each student in each resource (i.e., text, activities, or videos). To eliminate outliers (possible situations where a student might have started an activity and left their browser open while away from the computer), we replaced the bottom 5% of event durations with the 3rd quantile (75%) value of the distribution of all values of event durations. In addition to calculating the total time spent on each resource, we also calculated the number of pages read, videos watched, and activities completed by each student. We labeled these as "Page Read", "Video Watched" and "Activity Done" counts. All counts represent repeated increments as present in the student interaction of a given page in the data.

Persistence For us, students' persistence is a measure of how far a student progresses through the course, that is, the furthest page up to which a student reads, watches videos, or solves problems / activities from the start of the course. As observed in the data and reported in Table 1, it appears that students continue with reading, watching, or doing up to 75.7 pages on average.

Course Engagement Course Engagement is the number of unique activities done / videos watched

[2] https://search.r-project.org/CRAN/refmans/DescTools/html/Winsorize.html

/ content read by a student up to the page the student progressed through the course. We calculated this measure by taking the ratios between the number of *page read*, *video watched*, and *activity done* counts, respectively, and dividing them by the maximum of pages, videos and activities that the student could complete up to the page they reach the furthest from the start.

Self-reports of Interest and Students' Instruction Finally, using the embedded questionnaires, we calculated for each student their career interests and confidence. We summarized the measurements used for testing the hypotheses in Table 1.

Amount of Doing We defined one categorical variable for students who read more pages, watch more videos, and engage in more activities. Many students read, fewer watched, and even fewer did activities (see Figure 3). Following the first paper on this topic [4], we used the medians to differentiate "readers" from "non−readers"; "watchers" from "non−watchers"; "doers" from "non-doers". This measure would help us analyze the extent to which a specific type of instruction, as defined by the kind of interaction – reading, watching or doing – that students engage in more often when progressing through the digital textbook-based MOOC, in combination with students' prior motivation affects their interest to pursue a career in computer science at the end of the course.

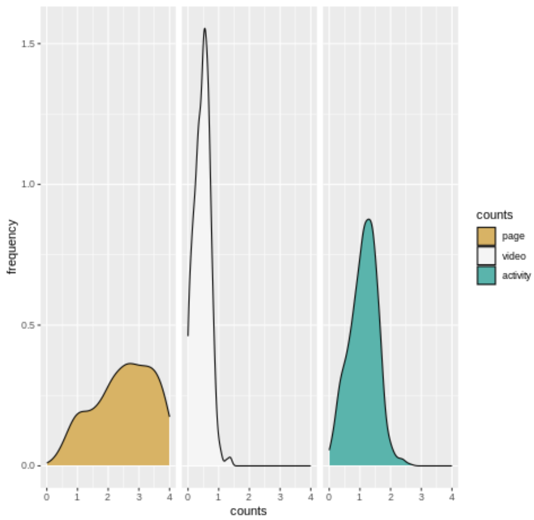


Figure 3: Distributions for the standardized Reading, Video and Activity Counts completed by students. We observed that the per page read counts with maximum number of students is between 3 and 4, for video counts is between 0 and 1 and for act counts is between 1 and 2.

4.2. Analyses

To address our 3 research questions, we used multiple linear regression analyses. For all regression analyses, we normalized all continuous predictors and outcomes by z-scoring the raw values, allowing us to compare regression estimates as effect sizes.

To answer **RQ1**, we evaluated the *"Doer Effect"* in the data using the outcome and predictors as described in the original work. We used multiple linear regression analyses with page, video, and activity counts / times as the predictors. The predicted outcome was students' post-test scores while controlling for their pre-test scores.

To answer **RQ2**, we studied the relationship between student motivation and doing more activities in a course. We used the standardized resource use and baseline student confidence to learn JAVA (see Table 1) as predictors. The results would help demonstrate that student motivation is associated with *"Doer Effect"*.

To answer **RQ3**, we finally analyzed whether students who did more were interested at the end of the course in pursuing a career in computer science (CS) when controlling for their interest to pursue a career in CS at the beginning of the course. We used regression analysis to predict their career interest

	N	mean	SD
Total # of Students	1060	-	-
(for RQ1) N Completed Posttest	92	-	-
(for RQ 2,3) N Completed Pretest	694	-	-
# Pages in the Book	96	-	-
# of activities	674	-	-
# of videos	62	-	-
Students' Page Read Counts	-	547.3	798.32
Students' Video Watched Counts	-	27.11	15.98
Students' Activity Done Counts	-	655.5	301.85
Students' Read Time (in hrs)	-	220.85	242.45
Students' Video Time (in mins)	-	25.65	20.73
Students' Activity Time (in hrs)	-	11.1	6.80
Students' Persistence	-	75.70	17.63
Standardized_Page_Counts = Page_Read_Counts / Persistence	-	6.91	9.4
Standardized_Video_Counts = Video_Watched_Counts / Persistence	-	0.46	0.25
Standardized_Activity_Counts = Activity_Done_Counts / Persistence	-	1.12	0.45
Confidence to Learn JAVA	-	1.35	17.63
Career interest in CS (Pre)	-	1.02	1.05
Career interest in CS (Pre)	-	1.12	1.24
# Readers	72	-	-
# Watchers	30	-	-
# Doers	38	-	-
# Readers or Watchers or Both	47	-	-
# Doers or Watchers or Both	81	-	-
# Readers or Doers or Both	48	-	-
# Readers and Watchers and Doers	180	-	-

Table 1
Measures / Features present and created for the purposes of analyzing the dataset to answer RQ 1, 2 and 3.

at the end of the course. We used the baseline confidence to learn JAVA and its interaction with the baseline career interest as predictors.

5. Results

5.1. RQ1 Is resource use related to better learning outcomes?

When using activity, video, and page read times as predictors and post-test scores as the outcome while controlling for the pre-test as a covariate, we could not demonstrate the *"Doer Effect"*. Similarly, we could not demonstrate the *"Doer Effect"*with the page, video, and activity counts as predictors of post-test scores. The coefficient for page counts had a small significant negative effect on the post-test scores when controlling for pre-test scores (see Table 2).

5.2. RQ2: Is resource use related to persistence in the course?

We found that both reading a larger percentage of the available pages and completing a larger percentage of the available activities were related to achieving more of the course, that is, persisting longer, controlling for baseline confidence (see Table 3). Notably, the relation between persistence and completing more of the available activities was 2.3 times stronger than the relation between persistence and completing more of the available reading pages, suggesting a *"Doer Effect"*.

5.3. RQ3: Is resource use related to interest in CS careers?

We found that there was a positive relationship between higher baseline confidence in JAVA and post-test interest in following a CS career ($\beta = -0.76, t = 0.12, p < .0001$), which is not surprising, as

Independent Variable	Estimate	SE	t
Pretest Score	0.04	0.11	0.46
Activity Time (in s)	-0.08	0.11	-0.63
Video Time (in s)	0.10	0.13	0.75
Read Time (in s)	-0.04	0.11	-0.40
Pretest Score	0.07	0.10	0.71
Activity Counts	-0.09	0.15	-0.59
Video Counts	0.27	0.16	$1.67^{\cdot}$
Read Counts	-0.27	0.11	-2.38^{*}

Table 2
Model Parameter Estimates for the *"Doer Effect"* with independent variables on Activity Time / Counts, Video Time / counts and Read Time / Counts, with the top and bottom portions of the table focusing on the estimates for the model with times and counts as predictors respectively. $^{***}p < 0.001,\ ^{**}p < 0.01,\ ^{*}p < 0.05$

Independent Variable	Estimate	SE	t
Confidence to Learn JAVA	-0.4	0.00	-0.85
%age of Pages Read	0.08	0.04	2.07^{*}
%age of Videos Watched	0.08	0.05	1.54
%age of Activities Done	0.19	0.05	3.73^{***}

Table 3
Model Parameter estimates using the counts and student persistence as proxy for the *"Doer Effect"* $^{***}p < 0.001,\ ^{**}p < 0.01,\ ^{*}p < 0.05$.

more confident students are more likely to have more prior knowledge in programming. We found no relationship between the type of interaction the students had with the course and their final interest in pursuing a CS career or any interaction (see Table 4).

Independent Variable	Estimate	SE	t
Career Interest	-0.76	0.12	-6.24^{***}
Confidence to Learn JAVA	0.03	0.26	-0.76
Confidence × Career	-0.09	0.13	0.52

Table 4
Model Parameter Estimates for the effect of proxy measures for student motivation, self–efficacy and goals on the $^{***}p < 0.001,\ ^{**}p < 0.01,\ ^{*}p < 0.05$

6. Discussion & Conclusion

Although the *"Doer Effect"* has been identified in many courses and datasets before, we failed to identify it in the current paper when predicting learning outcomes. There are multiple possible reasons for this. The number of students contributing data to our analyses is likely too small to identify the effect. Prior work has used larger sample sizes. It is also possible that the use of the text and the video was substantially different in the current sample than in previous studies. Regardless, the numeric direction of the data is consistent with prior doer effect results.

Notably, the results presented here extend existing evidence of a doer effect to examine motivational outcomes rather than learning outcomes. Motivated by previous work [8], our second analysis revealed that students who do more activities also continue for longer than those who watch more videos or read more pages. Therefore, the *"Doer Effect"* might extend beyond learning outcomes to engagement and motivation. That is, completing more activities might not only lead to better learning through processes or retrieval and generalization associated with completing practice and receiving feedback [3], but it might also lead to greater interest and engagement. Thus, it is possible that completing more practice has the indirect effect of leading students to persist and continue learning, ultimately leading to more learning [7].

Despite this greater engagement among students who completed more activities, our third and final analysis found that the type of interaction that students had with the course or their initial confidence in learning JAVA did not affect their interest in pursuing a career in CS at the end of the course. It is not the case that more confident students are more likely to do more or more likely to want to pursue a career. Therefore, the relationships observed between doing more and engagement are unlikely to be limited to highly motivated students, although further research is needed.

Researchers have extensively studied students' persistence in MOOCs [4, 18] to help reduce student dropout. Understanding the relationship between the *"Doer Effect"* with student engagement/persistence will help design MOOCs that encourage student persistence. Evans and colleagues [18] find that the use of specific terms in video lectures increases student engagement and interaction, which we could not confirm in this work. However, when we encourage students to engage in more activities than just watching video lectures or reading texts, we can motivate them to exhibit greater persistence. This result is consistent with prior results [7].

Acknowledgments

Thanks to Brad Miller, DataShop @ CMU team for the dataset. Annie Wang, Ryan Nuqui for LearnSphere integration. This work was supported by the National Science Foundation under Grant Nos. 2418655 and 2418656, to Peter Brusilovsky and Paulo F. Carvalho.

Declaration on Generative AI

We did not use Generative AI for this work. Grammar correction was done with human interventions.

References

[1] M. W. Asher, F. Sana, K. R. Koedinger, P. Carvalho, Students can learn more efficiently when lectures are replaced with practice opportunities and feedback, in: Proceedings of the Annual Meeting of the Cognitive Science Society, volume 46, 2024, pp. 1–9.

[2] P. F. Carvalho, M. Gao, B. A. Motz, K. R. Koedinger, Analyzing the relative learning benefits of completing required activities and optional readings in online courses, in: the 11th International Conference on Educational Data Mining, 2018, pp. 418–423.

[3] P. F. Carvalho, E. A. McLaughlin, K. R. Koedinger, Varied practice testing is associated with better learning outcomes in self-regulated online learning., in: Journal of Educational Psychology, volume 114, American Psychological Association, 2022, pp. 1723–1742.

[4] K. R. Koedinger, J. Kim, J. Z. Jia, E. A. McLaughlin, N. L. Bier, Learning is not a spectator sport: Doing is better than watching for learning from a MOOC, in: Proceedings of the Second (2015) ACM conference on Learning at Scale, 2015, pp. 111–120.

[5] K. R. Koedinger, E. A. McLaughlin, J. Z. Jia, N. L. Bier, Is the doer effect a causal relationship? How can we tell and why it's important, in: Proceedings of the Sixth International Conference on Learning Analytics and Knowledge, 2016, pp. 388–397.

[6] X. Hou, P. F. Carvalho, K. R. Koedinger, Drinking our own champagne: Analyzing the impact of learning-by-doing resources in an e-learning course, in: Companion Proceedings of the 11th International Conference on Learning Analytics & Knowledge LAK21, 2021, pp. 1–3.

[7] M. W. Asher, J. D. Hartman, M. Blaser, J. F. Eichler, P. F. Carvalho, Test, review, repeat: Mastery-based testing and its benefits for student engagement and performance in a general chemistry course, in: OSFPreprints, 2024, pp. 1–32.

[8] M. W. Asher, F. Sana, K. R. Koedinger, P. F. Carvalho, Practice with feedback versus lecture: Consequences for learning, efficiency, and motivation, in: Journal of Applied Research in Memory and Cognition, Educational Publishing Foundation, 2025, pp. 1–44.

[9] P. Brusilovsky, E. Schwarz, G. Weber, Electronic textbooks on WWW: from static hypertext to interactivity and adaptivity, Educational Technology Publications, Englewood Cliffs, New Jersey, 1997, pp. 255–261.
[10] E. Walker, R. Wylie, A. Danielescu, J. P. Rodriguez III, E. Finn, Balancing student needs and learning theory in a social interactive postdigital textbook, in: End-user considerations in educational technology design, IGI Global, 2018, pp. 141–159.
[11] G. Rößling, T. Naps, M. S. Hall, V. Karavirta, A. Kerren, C. Leska, A. Moreno, R. Oechsle, S. H. Rodger, J. Urquiza-Fuentes, et al., Merging interactive visualizations with hypertextbooks and course management, in: Working group reports on ITiCSE on Innovation and technology in computer science education, 2006, pp. 166–181.
[12] B. J. Ericson, I. YeckehZaare, M. J. Guzdial, Runestone interactive ebooks: A research platform for online computer science learning, in: Proceedings of SPLICE Workshop in the International Computing Education Research Conference (ICER'19), 2019, pp. 1–7.
[13] K. Pollari-Malmi, J. Guerra, P. Brusilovsky, L. Malmi, T. Sirkiä, On the value of using an interactive electronic textbook in an introductory programming course, in: Proceedings of the 17th Koli Calling International Conference on Computing Education Research, 2017, pp. 168–172.
[14] R. Van Campenhout, B. G. Johnson, J. A. Olsen, The doer effect: Replicating findings that doing causes learning, in: Proceedings of eLmL2021, The Thirteenth International Conference on Mobile, Hybrid, and On-line Learning, 2021, pp. 1–6.
[15] R. Van Campenhout, B. Jerome, J. S. Dittel, B. G. Johnson, The doer effect at scale: investigating correlation and causation across seven courses, in: LAK23: 13th International Learning Analytics and Knowledge Conference, 2023, pp. 357–365.
[16] K. R. Koedinger, J. Stamper, P. F. Carvalho, P. Pavlik, L. Eglington, Sharing and reusing data and analytic methods with learnsphere, in: Proceedings of the 10th International Conference on Educational Data Mining (EDM) 2017, volume 2, 2019, pp. 501–520.
[17] B. J. Ericson, M. J. Guzdial, B. B. Morrison, Analysis of interactive features designed to enhance learning in an ebook, in: Proceedings of the eleventh annual international conference on international computing education research, 2015, pp. 169–178.
[18] B. J. Evans, R. B. Baker, T. S. Dee, Persistence patterns in massive open online courses (MOOCs), in: Journal of Higher Education, volume 87, Taylor & Francis, 2016, pp. 206–242.