

Analyzing the Impact of Course Structure on Electronic Textbook Use in Blended Introductory Physics Courses

Daniel T. Seaton,^{1,*} Gerd Kortemeyer,² Yoav Bergner,^{1,†} Saif Rayyan,¹ and David E. Pritchard¹

¹*Department of Physics and Research Laboratory for Electronics,
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

²*Lyman Briggs College and Department of Physics and Astronomy,
Michigan State University, East Lansing, Michigan 48825, USA*

(Dated: August 20, 2014)

How do elements of course structure (i.e., the frequency of assessments, as well as the sequencing and weight of course resources) influence the usage patterns of electronic textbooks (etexts) in introductory physics courses? We are analyzing the access logs of courses at Michigan State University and the Massachusetts Institute of Technology, which deployed etexts as primary or secondary texts in combination with different formative assessments (e.g., embedded reading questions) and different summative assessment (exam) schedules. As such studies are frequently marred by arguments over what constitutes a “meaningful” interaction with a particular page (usually judged by how long the page remains on the screen), we are considering a set of different definitions of “meaningful” interactions. We find that course structure has a strong influence on how much of the etexts students actually read, and when they do so: particularly courses that deviate strongly from traditional structures, most notably by more frequent exams, show consistently high usage of the materials with far less “cramming” before exams.

PACS numbers: 01.40.-d, 01.40.Di, 01.40.Fk, 01.50.H-, 01.50.ht

I. INTRODUCTION

A. Log Analysis

The increasing availability of electronic learning resources is providing instructors unique opportunities to explore new course materials and methods of instruction. However, decisions for the adoption of these resources, in particular electronic texts (etexts), are frequently driven by external factors rather than pedagogical considerations: advances in e-reader and web-based technologies,[1, 2] changing student preferences,[3, 4] increasing availability of open educational resources (OERs),[5] the emergence of MOOCs,[6, 7] and the cost of traditional textbooks.[8] Research regarding the effectiveness of physical textbooks in introductory courses shows generally low use and poor correlations with performance.[9–12] It is doubtful that simply switching the medium will result in any positive effects on educational effectiveness if no other changes are made.

While student self-reported surveys [13, 14] have been the primary means of analyzing use of physical textbooks, Learning Management Systems (LMSs) provide an attractive and more reliable alternative in stored digital records of all student interactions with etexts. Each time a student accesses a page, an event is recorded containing the student ID, page accessed, and a timestamp. From

these records, one can describe etext use metrics such as the overall use and the associated temporal pattern, providing an attractive alternative to student surveys. These records provide a unique window into the learning behavior of students [15–19] during activities traditionally performed beyond the instructor’s scope, e.g., reading the textbook before lecture.

The LON-CAPA [20] learning management system provides a platform for disseminating etexts (along with other course resources), while also keeping rich data describing student interactions. An etext in LON-CAPA is a collection of HTML pages containing primarily textual material, typically split into short pages meeting digital standards [21]. Equally important is LON-CAPA’s legacy over more than a decade of having delivered resources in over 7,700 blended and online courses to over 960,000 student course enrollments. Student activity records from LON-CAPA courses have provided in-depth insight into student problem-solving behavior,[17, 22, 23] but little work has been aimed at interactions with etexts.

B. Effects of Course Structure

Course structure is an important feature of any course, encompassing, the types of learning activities, their frequency, sequence, and overall weight toward final grade. Much work has been aimed at elucidating the relationship between course structure/context and student outcomes in introductory science courses (e.g. [24–30]). Improved performance and attitudes have been shown to accompany changes from traditional to frequent exam formats, as well as the replacement of traditional lectures with active-learning activities.[29–34] Ad-

*E-mail: dseaton@mit.edu; now at Davidson College, Davidson, NC 28035, USA

†now at Educational Testing Service, Princeton, New Jersey 08541, USA

ditional research shows positive outcomes when replacing physical-text pre-lecture activities with online media modules.[12, 35–37] Each of these studies involves linking course structure and student outcomes, but with few exceptions (e.g. [24, 25]) overlooks particular aspects of how students adapt their behavior to be successful in a given course. The relationship between course structure and student behavior is a necessary link for understanding learning, as well as in providing instructors information about their teaching strategies.

C. Massed Versus Distributed Practice: “Cramming”

Results from our log analysis on etext usage can have implications for course design: once connections between course structure and learning behavior have been identified, course structures can be modified to foster constructive and discourage counter-productive behaviors. When it comes to reading the textbook, we were not only interested in the question whether the materials were used in the first place, but even more so in the question of when and how they were used. Of particular interest to us was “cramming,” i.e., the attempt to concentrate reading and studying materials into the very last days leading up to an exam. Cramming can be a short-term successful strategy in getting good grades, but it has long been known that cramming has detrimental effects on long-term retention (e.g., studies on “forgetting curves” as early as 1938.[38, 39]) Essentially, cramming (or “massed practice”) stores information in short-term memory with little long-term benefits. The findings have been confirmed over the decades, and strategies for providing opportunities for distributed practice have found their way into classrooms (e.g. Refs. [40, 41]).

D. Challenges of Analyzing Electronic Log Files

While in many respects, analyzing the use of electronic texts is easier and more straightforward than doing the same for physical texts, there are still challenges that make this endeavor more challenging than it first appears: while we know when a learner opened a page, we have no automated way of finding out what he or she did with that page: did the student simply flip past the page, look up a formula or definition, or did he or she carefully read the text? For lack of these insights, timing information becomes a crucial clue, and we need to carefully look at the deployed time windows (“cutoffs”). As we will find, some measures are highly sensitive to the choice of cutoffs (but in a predictable way), while some of the qualitative results are surprisingly robust.

Fortunately, we have a large data set at our disposal. We analyze retroactive data from one Mechanics-reform course at the Massachusetts Institute of Technology (MIT), and nearly a decade of large-lecture courses

from Michigan State University (MSU). Based on the assignment of the etext, exam frequency, and integration of homework within the etext, three categories of course-structure emerge: Supplementary, Traditional, and Reformed. We present background for our categorization and give necessary insight into the courses. After defining our methodologies, we present tools that illuminate both the overall amount and the temporal pattern of etext use. In all subsequent analyses, we highlight differences in etext use through the lens of variations in course structure. Discussion and conclusions follow.

II. ELECTRONIC TEXTS AND COURSE STRUCTURE

We have analyzed etext use in nearly a decade of large-lecture introductory-physics courses at MSU and in one spring semester instance of a Mechanics reform course at MIT. The MSU courses were chosen due to their use of a single etext volume designed by the MultiMedia Physics group at MSU,[42] while the MIT course was analyzed because of the authors’ involvement with implementing its etext as part of the RELATE group’s Integrated Learning Environment for Mechanics.[43]

A. MSU Multimedia Physics Electronic Text

The MMP etext covers traditional physics courses at MSU, whose mechanics (1st semester) volume contains approximately 330 HTML pages, and approximately 30 other e-resources (videos, simulations, etc.). A number of courses at MSU have utilized this etext, and the majority of pages have stayed true to their original design. The etext has an algebra-based and a calculus-based variant for the respective courses. Instructors do have some freedom in which portions of the etext are presented to students. Hence, for certain metrics, we only compare those etext resources that are identical between courses.

B. MIT RELATE Electronic Text

The RELATE etext covers a calculus-based course on introductory mechanics, representing one component of the Integrated Learning Environment for Mechanics.[23, 43] The etext contains 255 HTML pages, and 40 other e-resources (videos, simulations, etc.). The text introduces the Modeling Applied to Problem Solving (MAPS) pedagogy,[44, 45] which provides a framework for building expert-like problem solving skills. Mechanics topics begin with Newton’s Laws rather than kinematics (1D, 2D motion), meaning topics are discussed in a different sequence compared to a traditional introductory mechanics course. Additional effort has been given to adapting the etext around non-traditional course design.[45] This

TABLE I: : Course structure features for all studied courses. Three major categories emerge based on implementation of the etext and overall course structure: Supplementary, Traditional, and Reformed.

Type	Students	Assignment	Pages	Exams	Embedded Assessment	Course Label
MSU algebra-based	898	Secondary	340	3 + Final	No	Supplementary A
MSU algebra-based	911	Secondary	338	3 + Final	No	Supplementary B
MSU algebra-based	808	Secondary	338	2 + Final	No	Supplementary C
MSU calc-based premed	159	Primary	402	2 + Final	No	Traditional A
MSU calc-based premed	190	Primary	383	2 + Final	No	Traditional B
MSU calc-based premed	211	Primary	318	6 + Final	Yes	Reformed A
MSU calc-based premed	209	Primary	295	6 + Final	Yes	Reformed B
MSU calc-based premed	197	Primary	295	6 + Final	Yes	Reformed C
MSU calc-based premed	254	Primary	300	6 + Final	Yes	Reformed D
MIT calc-based	38	Primary	255	12 + Final	Yes	MIT

same etext also provides the base of the MITx Mechanics Review MOOC.[7, 46, 47]

C. Course Structure

1. MSU Algebra-Based Course

The algebra-based course at MSU is a large lecture introductory mechanics course for students throughout the university. A typical course consists of multiple sections with multiple instructors. We study three instances of this course, whose combined enrollment reaches approximately 2400 students. During the three years of these courses, overall structure and etext resources saw minimal changes, however, instructors changed. The course structure consisted of weekly homework, 2 to 3 midterms and a final exam, and no embedded assessment in the reading.

Although the etext was a very visible part of the student experience (available for free in the same system as the free online homework), the syllabus for each course described an optional physical textbook available for purchase at the university bookstore. Hence, we infer that the etext played a supplementary role in the student experience.

2. MSU Calculus-Based for Mostly Premedical Students

The calculus-based physics at MSU, one of two types, caters mainly to premedical students. We have analyzed six fall semester instances of this course, all with a single instructor. The total student population reached approximately 1200 students, and in all courses, the etext was the only textual-based resource provided. During the first two years, these courses followed a traditional course structure: weekly homework, 3 midterms and a final exam, and no embedded assessment in the reading. The final four courses were restructured to 6 midterms and a final, as well as embedded graded assessment within the etext (in addition to the mostly unchanged end-of-

chapter homework). Both embedded and end-of-chapter problems had (for the most part) the same point-values per problem, however, there were about five times more end-of-chapter than embedded problems. Overall, the data sets are providing a distinct transition in course-structure from traditional settings with a homogenous student population over all courses.

3. MIT Course

For comparison with the MSU data, where appropriate, we investigate student interactions with the RELATE etext in one spring mechanics course with an enrollment of 38 students. Typically, this student population did not receive a passing grade in the institute-wide fall-semester mechanics course. Hence, the majority of these students have just finished a full mechanics course, giving them a working knowledge of many basic physics concepts. The course adapts to this knowledge base, focusing more on problem-solving skills and a refined view of mechanics. The RELATE etext is the only textual reference provided to the students. The course structure was flipped [48], and included weekly homework (post-class), weekly reading assignments with embedded assessment (pre-class), weekly quizzes, and a final exam.

The MIT course has both a very different population and pedagogy (see Section II B) than the MSU courses, so it is of particular interest how well results transfer between these educational venues — are the findings limited to the MSU courses or more widely applicable?

4. Categorization

All courses in this study utilized LON-CAPA to deliver an etext and weekly homework assignments. The etext and homework components of the courses were both easily navigable, and students could freely navigate between etext resources and homework, i.e., they were not forced into accessing them in a certain sequence. However, the courses differed in a number of respects, which we will

refer to as “course structure:”

Nature of etext assignment: In most courses, the etext was the primary textual reference, i.e., there was no physical textbook. Alternatively, some of the MSU courses assigned the etext as a secondary text alongside a primary physical textbook, giving students choice on which resource to use. Unfortunately, all courses with primary etext assignments were calculus-based, while all courses with secondary etext assignments were algebra-based; thus, some possible population effects cannot be excluded.

Exam frequency: The number of exams varied from traditional, two or three midterms and a final, to reformed, weekly or bi-weekly midterms and a final. The reformed structure in all cases was chosen to provide more opportunities for formative assessment and remediation.

Embedded assessment: For some courses, online formative assessment was embedded into the etext material, while for others it was not.

To reflect these differences in course structure with respect to etext assignment and assessment, we define the following categories for the purposes of this study:

Supplementary: optional etext alongside optional physical textbook, traditional exam structure, no embedded assessment.

Traditional: primary etext, traditional exam structure, no embedded assessment.

Reformed: primary etext, frequent exam structure, includes embedded assessment.

MIT: the MIT comparison course, which according to our course structure criteria is most closely related to the MSU Reformed courses.

By necessity, this focussed categorization ignores several other aspects of the courses which may be considered “traditional” or “reformed,” for example the style of lecture presentation, the use of peer-instruction, etc. (for an overview of the wider spectrum of course reform, see for example MacIsaac and Falconer [49] and Pollock [28]). Table I details our categorization and provides descriptive statistics for each course; the course label in the rightmost column provides a unique identifier that will be used to reference courses when presenting data.

III. METHODOLOGIES

LON-CAPA generates server logs containing information on all student-resource interactions. Parsing these logs allows one to extract meaningful metrics on resource

use and student behavior.[17, 22, 50, 51] A single student-resource interaction contains at least the student ID, the ID of the accessed resource, and a time-stamp with a resolution in seconds. These tags allow us to count the number of accesses each student has with a given resource. In this study, our focus is on analysis of etext use, but additional data are available for other resource types, e.g., problem interactions allow analysis of student performance (e.g., correctness, number of tries, etc.).

As part of the human subject research procedure, the MSU logs were de-identified and no enrollment or grade book information was available. Without this additional information, we were not able to distinguish students in the course from course personnel and potential guests (i.e., course personnel and guests were just another undistinguishable data point). In order to isolate the student populations, we used attempts on homework and exams to filter out users falling below an adequate minimal level of participation. These filters removed users not attempting at least 25% of all required homework problems, and, if the information was present, not attempting at least 25% of the exams. Note that these filters rely only on participation, not performance. Typically <10% of the users were removed from each course, accounting for early withdrawals, guests allowed to observe the course, and TAs. The MIT course, for which we had access to the actual enrollment table and grade book, provided a means of crosschecking our filters: the students that were filtered out by these participation criteria where indeed students who did not attend the course for credit.

Another limitation of our log file analysis method is that we can only consider pages within etexts that have been accessed at least once by any user over the duration of the course. In other words, we do not know about the existence of any pages that have never been accessed by anybody. This limitation is however not serious, as likely such a page would need to be several links removed from the table of contents.

A. Time Estimates and Cutoffs

In analyzing access logs of course management systems, a persistent question is what constitutes a valid interaction with a resource (e.g. [17]) — simply accessing a page does not necessarily mean that the learner meaningfully interacted with it. Similar to triggers in a high energy detector, we need to filter out noise and non-events. We address this issue by calculating the time spent on a resource as the time difference between two subsequent interactions. As an example, Fig. 1 shows the distribution of measured interaction times Δt with the etext in three of the ten courses that we considered, providing a general means for interpreting the types of etext interactions in a given course.

Referencing Fig. 1, one can make assumptions that extremely low time durations (between 0 and 10 seconds) merely constitute navigational events (e.g., flipping past

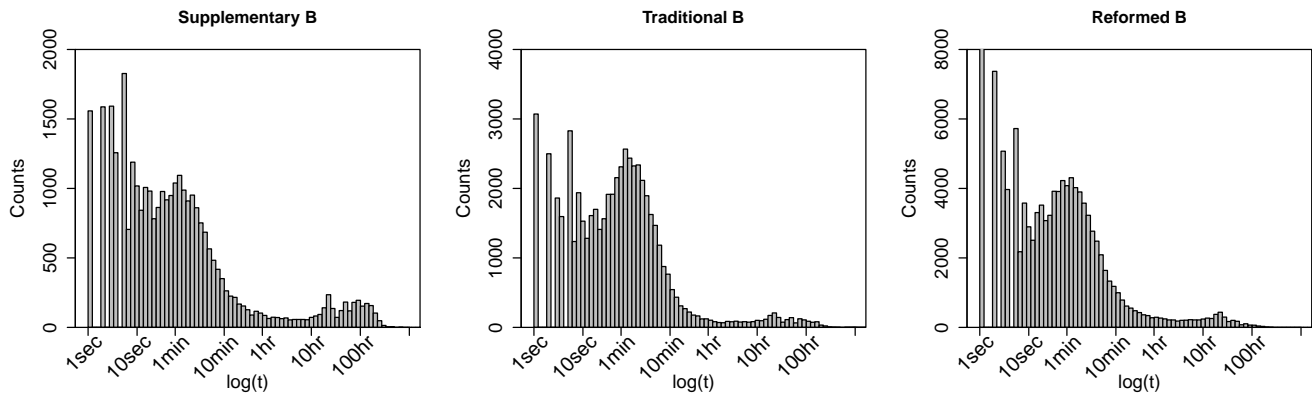


FIG. 1: Distribution of the logarithm of time spent (seconds) interacting with the eText pages. Examples are given for the courses Supplementary B, Traditional B, and Reformed B (see Table I). We indicate three prominent features: large-noisy bins between 0 and 10 seconds, a peaked distribution between 10 seconds and 1 hour, and a delayed peak around 24 hours — students returned to the system on the next day.

a page), and that extremely long times (≈ 24 hours) indicate that the student may have simply walked away from the session or closed the window or tab (for which we do not have log entries, as no server transaction would have been triggered). For intermediate time differences (between ten seconds and about one hour), one could assume that the student actually spent that time reading the content. However, all of these are assumptions, and one can easily find counter arguments: a very short access time could still be meaningful if the student quickly looked up a formula or definition that he or she had already accessed. A presumably meaningful access time of five minutes may in fact be meaningless if the student spent the vast majority of the time getting a coffee or doing social networking (we do not know what happens in other browser windows or tabs). On the other hand, an apparently “too long” interaction time could be meaningful if the student left pages open in a browser window or tab and then proceeded to flip back and forth between those for reference — something we would not see in server logs. Any imposed cutoffs then seem somewhat arbitrary, and we need to consider the influence of these cutoffs on the number of considered events. At best, we would find that the results are robust with respect to these choices.

Our hope is that the results are robust against the choice of cutoff times. To see if that is the case, we analyzed the data with four different cutoff ranges:

No cutoffs For this analysis, we imposed no cutoffs and treated all access events equally. However, as the system logs out any session that is inactive for more than 24 hours, there is an implicit upper limit of one day for the length of a transaction.

Wide cutoffs : For this analysis, we applied cutoffs of $t_{\text{low}} = 3$ sec and $t_{\text{high}} = 3600$ sec. These were motivated by the fact that on the average it takes at least three seconds for a page to completely build up in the browser, and any shorter time indicates

that the student moved past the page without even waiting for it to completely load. The upper limit of three hours is a rough estimate of how much time a student might reasonably spend on uninterrupted study.

Narrow cutoffs: For this analysis, we applied narrower cutoffs of $t_{\text{low}} = 10$ sec and $t_{\text{high}} = 1800$ sec. These cutoffs should filter out all navigational events and also all page accesses that take longer than the reading of any of the pages in the course would require.

Stringent cutoffs: For this analysis, we applied narrower cutoffs of $t_{\text{low}} = 30$ sec and $t_{\text{high}} = 600$ sec. These cutoffs focus solely on the maximum apparent in Fig. 1.

IV. RESULTS

A. Activity

To start exploring the impact of course structure on the usage of the eText, we look at eText usage over time. As expected, particular results will depend on the cutoff ranges used. Figures 2 and 3 show an analysis of time series for the three of the ten courses considered in our study, namely Supplementary B, Traditional B, and Reformed B. These courses are representative of what we find in other courses with the same structure.

Figure 2 shows the daily page-views per student ($A(t)/N$) for different cutoff choices. Note the change in scale between between supplementary courses on the one hand and traditional and reformed courses on the other.

Independent of chosen cutoff range, not only the total number of daily page accesses but also their time-dependence varies greatly with changing course struc-

ture. In Supplementary and Traditional courses, prominent spikes represent activity immediately before exams (exam frequency given in Tab. I), implying that the majority of students access the etext only days before exams. For example, in the Supplementary course, one can clearly see the three spikes corresponding to the three midterms, as well as the fourth spike before the final. The same is true for the Traditional course with its two midterms and a final. This is typical “cramming” behavior, and not very surprising — except for just how clearly the data reflect it! In Reformed courses, “cramming” spikes associated with biweekly exams are also still present, but they are far less pronounced and there is a more constant “background” of activity throughout the weeks. In other words, students were working more consistently also in off-exam weeks.

It is interesting to note the the cutoffs influence the total number of considered page accesses, but the overall time structure of the peaks is independent of the chosen cutoffs — the findings are robust with respect to the cutoff ranges.

Conceptualizing the daily page accesses as a time series, the autocorrelation function (ACF), shown in Fig. 3, provides further illumination of the underlying periodicity in the daily page-view activity. It is given by

$$\text{ACF}(n) = \frac{1}{N-n} \sum_{i=0}^{N-m-1} a(i)a(i+n), \quad (1)$$

where N is the total number of days in the course, $a(j)$ is the number of accesses per student on a given day j , and n is the number of days time lag considered (plotted on the abscissa of Fig. 3).

Supplementary B and Traditional B have significant correlation occurring at 28 and 42 day lags, respectively. Both lags correspond to the approximate time between examinations and both curves lack an instructor’s “expected” periodicity associated with weekly suggested readings. In other words, in these two courses, students read just before exams, not when they are supposed to do the readings — typical “cramming.” Reformed B differs greatly from these two curves, maintaining a consistent weekly periodicity for both small and large lag times. A valid concern is that the weekly peaks might still be a signal of “cramming” before exams that take place only every other week if a holiday or break period introduced one-week shifts in the schedule at some points during the semester (i.e, three weeks between exams instead of two). However, while an uneven number of such week-long shifts could indeed produce a weekly periodicity for long lag times, it would not explain the observed strong weekly peaks for short lag times. Thus, students in the Reformed course are indeed mostly using the etext weekly (as expected by the instructor) versus just immediately before exams.

Of some interest is the decrease in daily activity following the first exam in Supplementary B. This signal occurs in all Supplementary courses (not just the shown

course Supplementary B), as well as in Traditional A. Apparently, in these courses, after the first exam, the students gave up on even the little regular reading they used to do. There may be several reasons for giving up on reading:

- Students might have given up on reading the textbook because they deemed it not worthwhile, as it might not have helped them much with the exams given in those courses; students might be “optimizing” their learning for grades [24].
- Students might have experienced “burnout” after the first wave of midterm exams throughout their courses — if that is the case, providing more lower-stakes midterms might have exempted the reformed courses from burnout. It was found earlier that increasing exam frequency in a physics course indeed leads to higher student satisfaction and a more positive attitude toward the course [29].
- Students in the non-reformed courses might have experienced the typical “epistemology decay” in attitudes and expectations related to diligence [52, 53].

Based on log-analysis alone, the reason for this behavior cannot be determined, however, it is evident that the reformed structure counteracts it. Similar dependencies of course material accesses (lecture slides and course notes) were found in other studies of similar courses from MSU [29], as well as in online courses of varying format.[7, 46]

All courses show peak activity before the final exam, an assessment format common to all of them. One would have hoped that the reformed course would show a less pronounced final sprint, as students hopefully would have gained more confidence through more consistent studying throughout the semester, but the effect is minimal. On the other hand, it is probably too much to expect such short-term effects, and the hope is that through less cramming students carry more physics understanding beyond the final exam.

B. Fractional Usage

Fractional usage, i.e., how much of the etext individual students accessed, provides another indicator of course structure effects. Fig. 4 depicts complimentary cumulative distributions of unique etext pages viewed by students in each of the studied courses (Table I). This method of plotting allows for etext use in all courses to appear in a single plot, where each curve is read as the percentage of students $\%N$ having accessed greater than $\%E$ percentage of the available etext pages.

The influence of course structure (as indicated by the different lines types in Fig. 4) is striking:

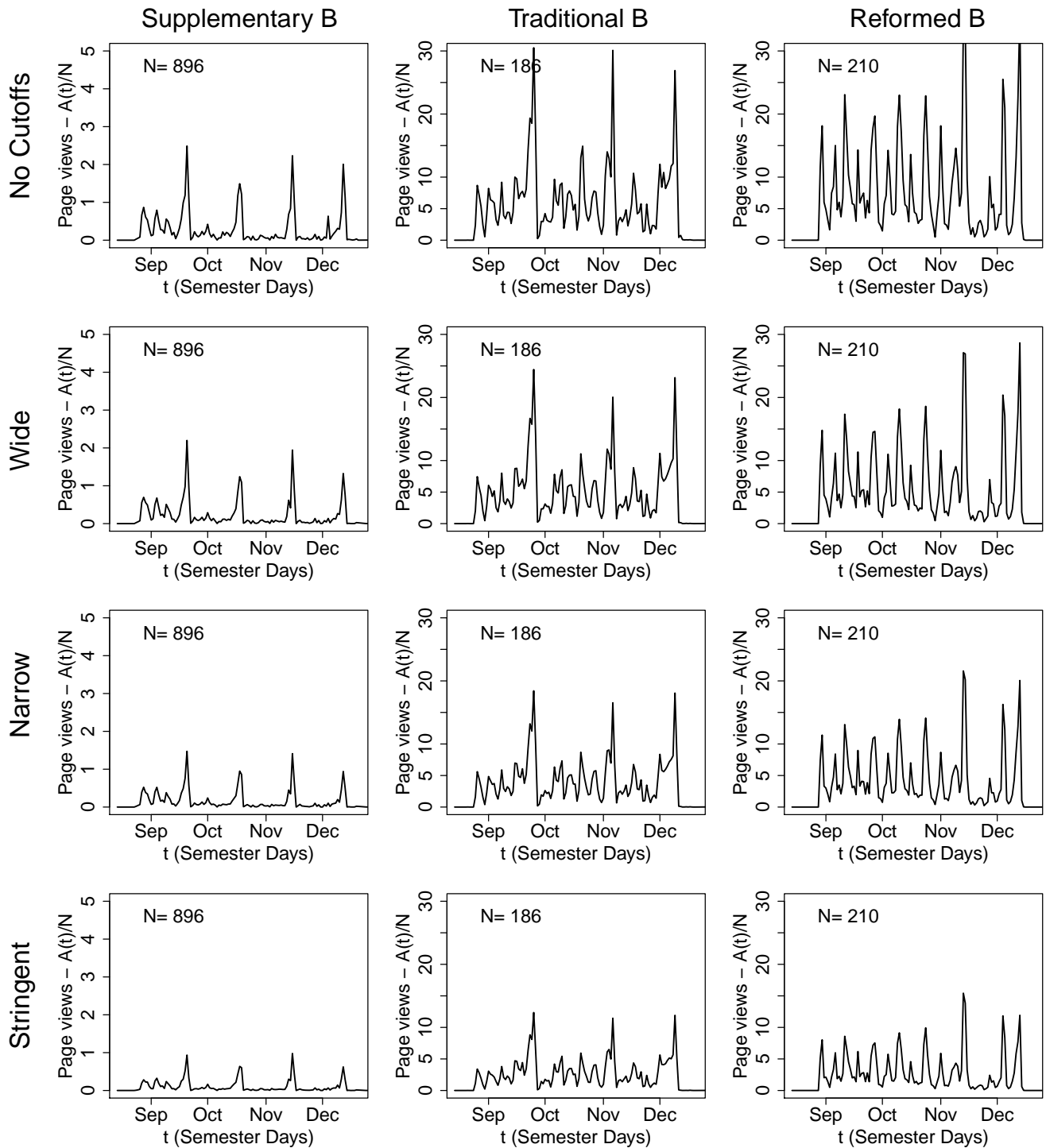


FIG. 2: Daily page-view activity per student $A(t)/N$ for different courses, employing different cutoff ranges.

- Supplementary courses (dotted) have the lowest extent use. For example, only about 10% of the students have accessed more than 20% of the pages.
- Reformed courses (solid black) have the highest overall use, with the majority of the students accessing the majority of the materials.
- Traditional courses (dashed gray) exist between low and high use.
- The MIT course (solid gray) resides entirely within the MSU regime of courses that also had primary etexts, which may be taken as an indicator that the stark distinction between primary and secondary

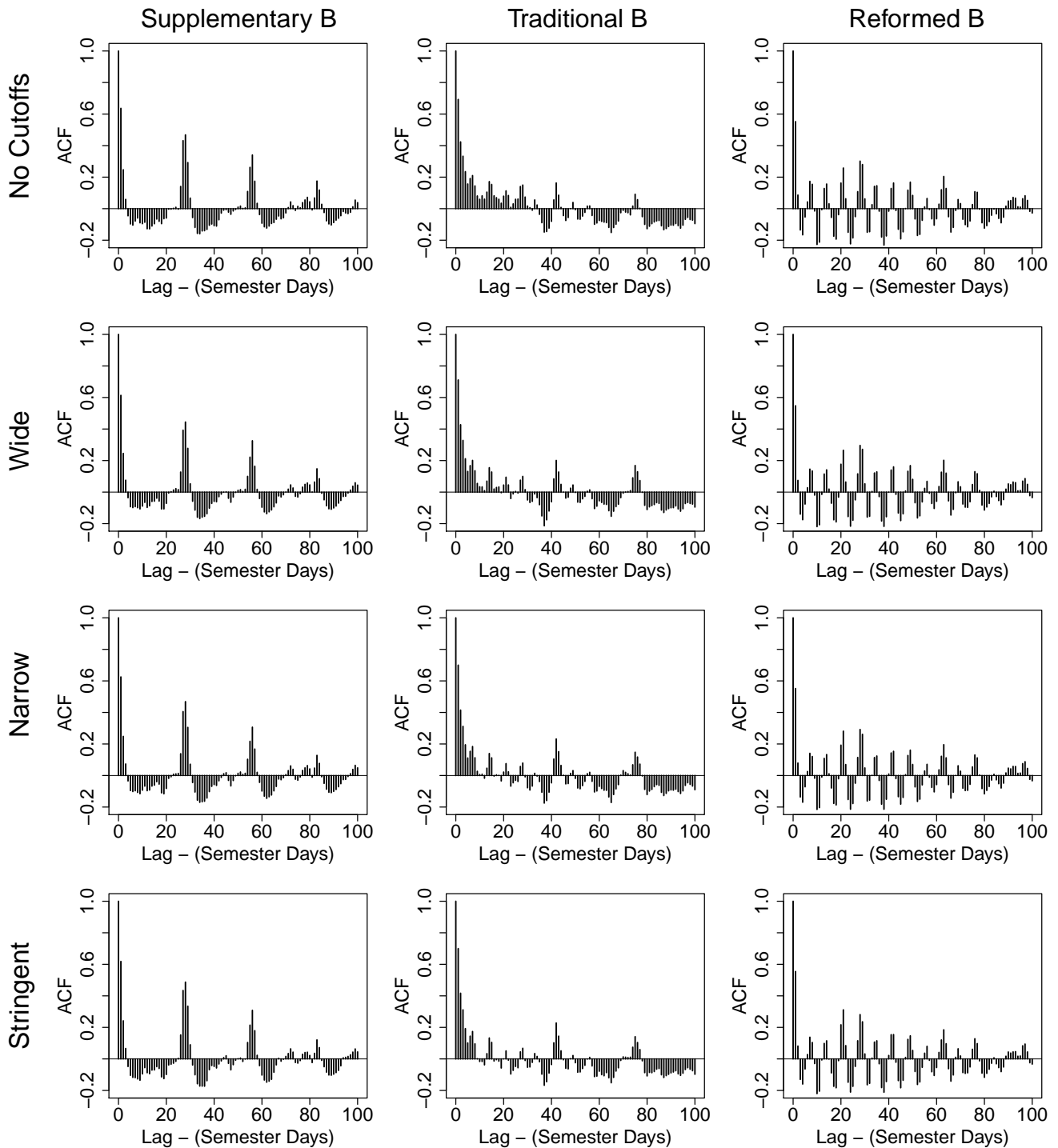


FIG. 3: Autocorrelation function ACF of the daily activity (Fig. 2) for different courses, employing different cutoff ranges.

etexts is not entirely population-dependent (see Sect. II C 4).

The results show remarkable consistency between courses that share the same structure: Supplemental courses show very little variation between semesters, and the regimes of Traditional and Reformed courses are clearly separated — by a wide margin from the Supplemental courses, but

also from each other.

Some of this signal gets lost when applying stricter and stricter cutoff ranges, as can be seen by the curves moving closer together. Particularly the MIT comparison course appears clearly within the regime of Reformed courses when no cutoff is applied, but for stricter cutoffs, this distinction gets lost. However, the overall structure

is encouragingly robust: no matter what criteria you apply for “meaningful” interaction with the pages, the regimes of the different course structures overall remain separated. The fact that stricter “noise” cutoffs lead to *less* signal could mean that even short interactions with a page might have meaning. This could be particularly true in physics, where complex relationships are expressed in succinct formula terms, which do not require a lot of reading time. From experience, for better or worse, even when reading a traditional physics book, eyes frequently gravitate toward the formulas while skimming or skipping the narrative. The MIT etext even supported this by having cascading explanation and derivation boxes that only open on demand, so the gist of the page can be seen at a glance; this feature may have contributed to its sensitivity regarding cutoffs.

C. Cumulative Usage

A common lore among instructors is that students “just don’t read the book.” Can we find evidence of this in the online realm? Hence, we are looking at the data to find out how much of the etext the students actually accessed at least once in a “meaningful” way. As Table I shows, all courses have a large number of pages, but students only look at a fraction of them. As Fig. 5 shows, this percentage strongly depends on the course type. Where the etext is only used as a secondary text, students on the average look at less than 10% of the pages, quite independent of the cutoff ranges set for “meaningful” interactions. For traditional courses with the etext as primary resource, the percentage varies between 30% and 60%, while for reformed courses, the percentage is between 50% and 90%. While the quantitative results depend on cutoff choices, the relative results do not: the realms of the different course structures remain clearly separated for any definition of “meaningful” interaction.

In summary, do students “read the book?” For secondary electronic texts in traditional settings, the answer is indeed “no,” as students on the average only look at about 5% of what is offered. For primary electronic texts, students read about half of it in Traditional courses, and three quarters in Reformed courses.

D. Views per Page

How often do students return to the same page? Do they return to what they have read before? Fig. 6 shows the average number of views per page in the courses, which has to be correlated with how much of the material the students access at least once to begin with (see Fig. 5).

For Supplementary courses, the average number of times a page in the etext is accessed by a particular student is about 1.3 times, i.e., for the most part, students

look at the pages exactly once.

The picture changes when looking at Traditional and Reformed courses, where the etext is the primary resource. On average, course pages are read more than once, students keep returning to the same pages. The separation between Traditional and Reformed courses is essentially non-existent. While overall, students look at less pages in traditional than in reformed courses, they return to these fewer pages about as often as students in the reformed courses: about 2.5 times. Students might be using these pages for review.

Of all considered interaction characteristics, the views per page most strongly depends on the cutoff range chosen. While the results remain qualitatively consistent, in absolute terms, factors of three to four between the range choices are observed. This dependency is dominated by the lower cutoff, which could have two reasons:

- We simply see the effect of “flipping pages back and forth,” students navigating past pages they have already seen — arguably, these events should be discarded.
- We see the effect of quickly looking up a fact or a formula, students coming back to a page exactly knowing what they are looking for — these are meaningful events.

Both categories of events may take place simultaneously, but an indicator for their relative weight may be that of the different course types, the reformed courses, i.e., those with the strongest assessment components, show the strongest dependency (solid black lines in Fig. 6), and students in these courses might more frequently quickly look up details of known relationships (desirable behavior) or hunt for quick answers (undesirable behavior). One might argue that these short page accesses also constitute a type of “cramming” (in this case for homework rather than for exams), however, it is probably too much to expect that students would suddenly become completely intrinsically motivated learners who flip around the pages of their physics text for the sheer joy of learning. Instead, one can hope that this externally motivated frequent searching for information to solve homework is the type of frequent-enough “cramming” that actually constitutes distributed practice.

E. Time-On-Task Reading

The last considered quantity is how much time on average students interact with the etext. Fig. 7 shows the average total time students spent with the etext (left panel) and particular pages in the etext. As expected, this quantity strongly depends on the choice of cutoff ranges, however, the only real outlier is the “no cutoff” result: since timeouts can be up to one day per interaction, the average reaches 300 hours for some courses and 100 minutes per page — this is obviously not how much

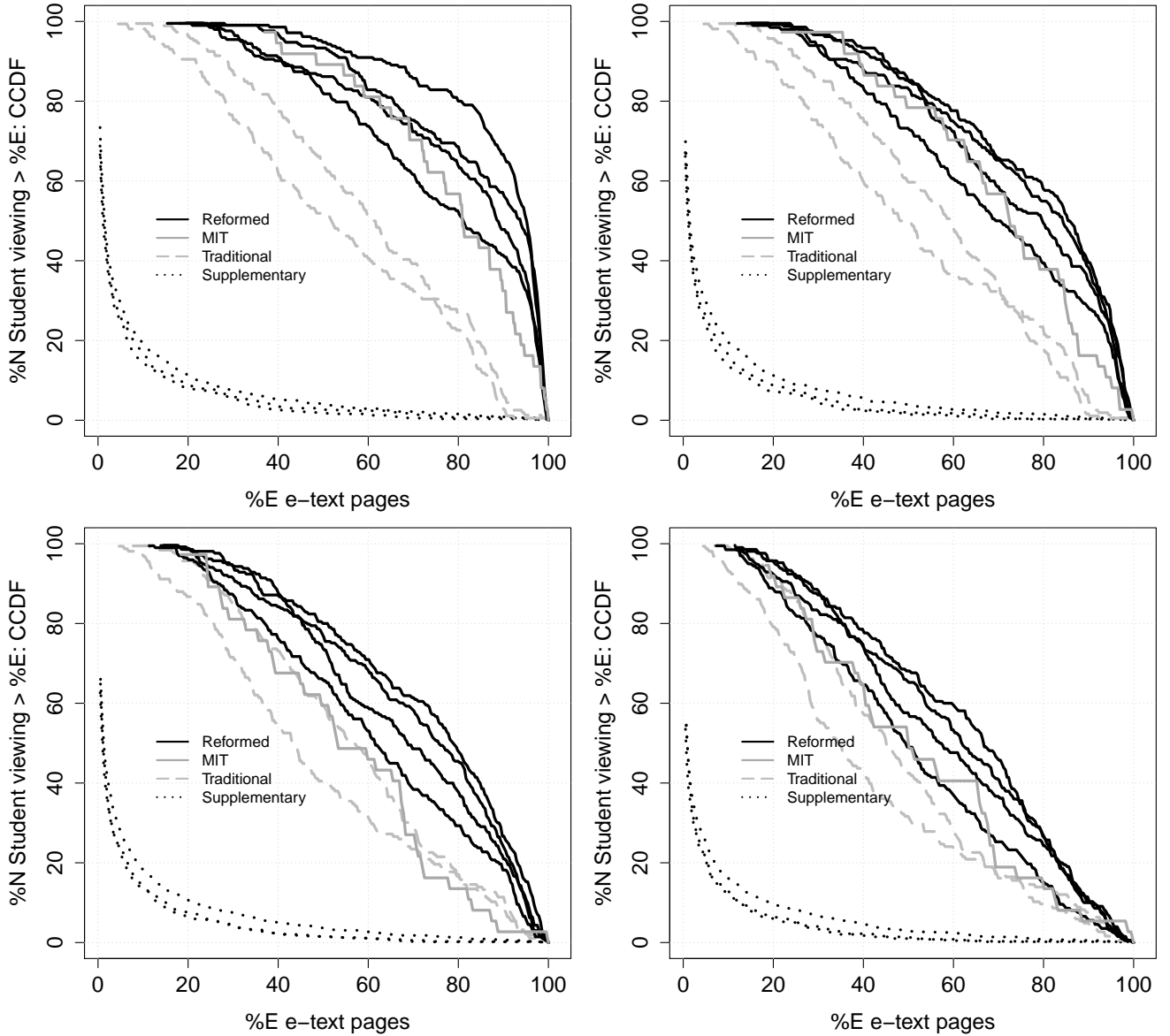


FIG. 4: Figure depicting the %N of students who have accessed greater than %E pages of their respective etext for no cutoffs (upper left panel), wide cutoffs (upper right panel), narrow cutoffs (lower left panel), and stringent cutoffs (lower right panel). The curves correspond to all considered courses (see Table I), and the line types correspond to different categories of courses (see Section II).

time the students actually spent, these averages are artificially inflated by abandoned windows or tabs, or by sessions that only 24 hours later ran against the timeout. As soon as *any* cutoffs are imposed, the results become reasonable, underlining the importance of t_{high} .

Overall, students in supplementary courses spent on the average between one and two hours with the etext, while students in traditional and reformed course invested one magnitude more time — about 10 to 20 hours.

Looking at particular page views, results indicate an average of about two to three minutes per page, independent of the course type. However, somewhat puzzling are the two courses Supplementary A and Supplementary B,

the older ones of the three supplementary courses, where depending on the cutoffs, students spend almost double the time per page. It might be explained by the fact that the students also look at the pages less frequently by about a factor two (see Fig. 6), and thus might spend more time with the page when they do, but that would not explain why Supplementary C does not follow this pattern. All three courses had different instructors, but used the same primary text. The only real difference is that Supplementary A and B had one more midterm than Supplementary C. However, time per page may not be very meaningful in courses where the etext is the secondary resource, as students might use both the physical

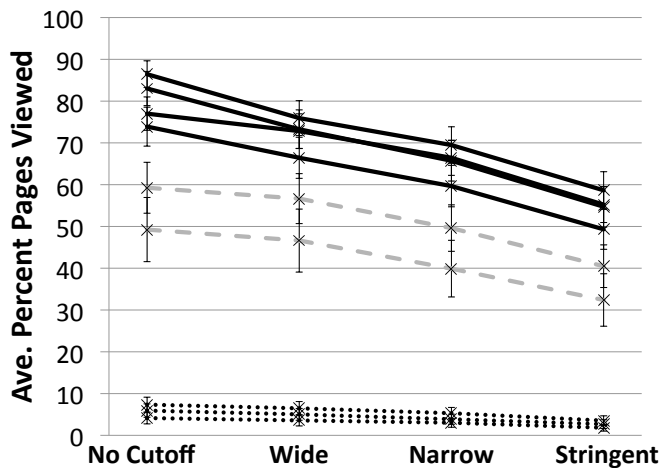


FIG. 5: Average percentage of pages in the online text actually viewed by students. The line types indicate the course types, i.e., Supplementary (dotted), Traditional (dashed gray), and Reformed (solid black). The percentage is given for all considered cutoff ranges (see Subsection III A), and standard errors of the mean are indicated.

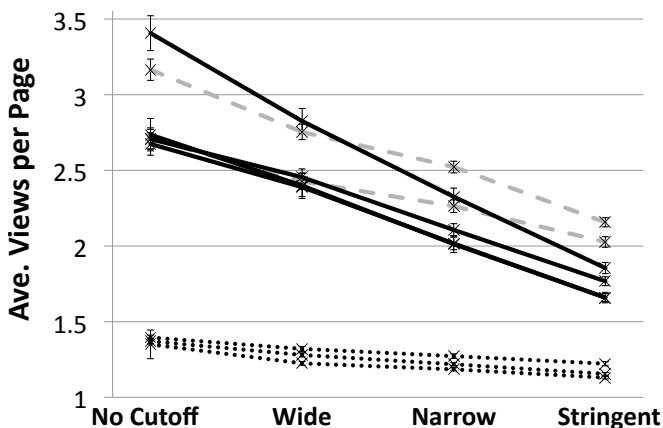


FIG. 6: Average views per page, i.e., how often students returned to the same page. The line types indicate the course types, i.e., Supplementary (dotted), Traditional (dashed gray), and Reformed (solid black). The figures are given for all considered cutoff ranges (see Subsection III A), and standard errors of the mean are indicated.

and electronic text in parallel, and web pages may linger on the screen longer while students are reading the book.

Overall, with those two exceptions, rather independent of cutoff range and course structure, students take about two to three minutes to read a page — which makes sense, given that due to the desire to avoid excessive amounts of scrolling, online pages are rather short. If, however, the cutoff range is chosen too stringently, the average time on task per page will necessarily converge toward the cutoff limits times the average views per page — we already observe the onset of this artifact as the graphs get squeezed in toward more stringent cutoffs.

V. DISCUSSION

At least in the United States, most introductory physics courses are using and following textbooks, yet actual usage of these books is rarely tracked: as they are used outside of classrooms, the only way to assess their usage is learner self-reported surveys which are not always very reliable. The courses we have been analyzing open a window into student-usage of textbook content, as the employed electronic format provides detailed access logs. Those logs still need to be approached with caution, as we do not know how the learners interact with the pages they access, but we are encouraged that our results are robust with respect to any reasonable choice of criteria for “meaningful” accesses.

Our results reveal that course structure is an important factor governing how students interact with textual resources:

Primary versus secondary text: The courses that used the electronic text as secondary text saw hardly any use of this resource (courses Supplementary A, B, C in Fig. 2; note, again, the change of scale). At first glance, this result may not be surprising, however:

- in these courses students were given the *option* of buying a physical textbook (homework was offered for free in the online system). Instructors are often puzzled that students would pay for a physical textbook over a free electronic text, and we acknowledge this to be an interesting area of study.
- the starkness of the result is puzzling since the courses had online homework within the same online system as the secondary electronic text was offered — the students did not even have the inconvenience of having to log into a separate system to access it.

We do not know how much the primary physical textbook was used (or even bought in first place) by the learners; maybe it also saw very little use. However, the result may indicate how much students (and possibly the instructors) cling to the textbook. In most any other part of the World, physics courses are not run based on a particular textbook,[54] and in fact, they do not need to: all of the knowledge required for introductory physics is available in any physics book printed within the last century and also freely available online. Not clinging to a particular textbook, even though it is very costly, may move students in the United States too far outside their comfort zone.

Frequent and embedded assessment: With the available data, it was impossible to decouple the effects of exam frequency and embedded assessment due to their introduction in the same

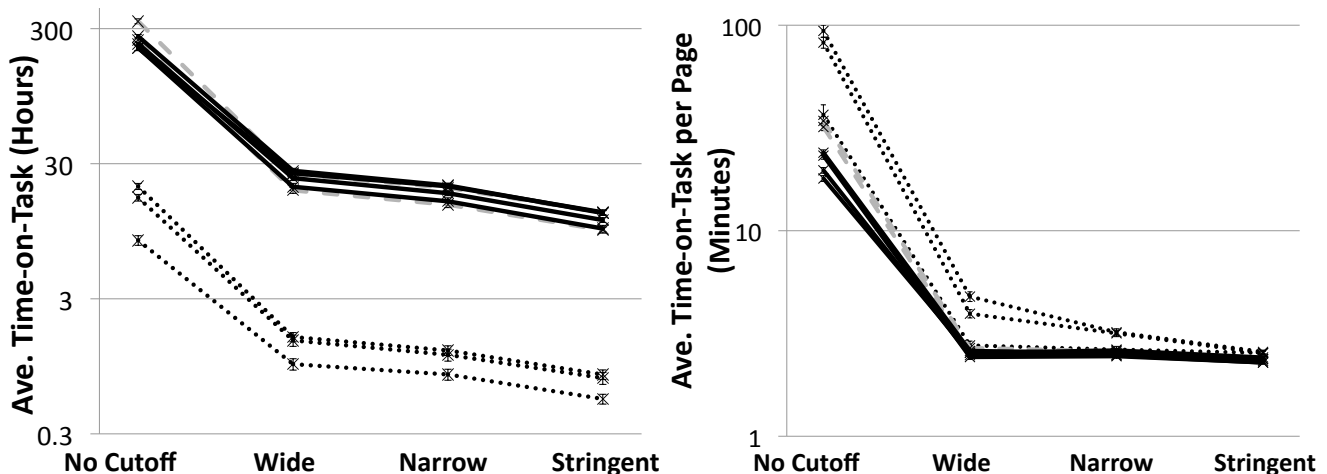


FIG. 7: Average time-on-task in the course (left panel) and per page (right panel). The line types indicate the course types, i.e., Supplementary (dotted), Traditional (dashed gray), and Reformed (solid black). The figures and standard errors of the mean are given for all considered cutoff ranges (see Subsection III A)

course year. Introduction of either or both led to a reduction of “cramming,” which is widely considered an unproductive study strategy with negative effects on coherence and longterm retention of knowledge. While to some level, students were simply forced to abandon last-minute cramming in favor of more constantly keeping up with the course, simply due to the assessment schedule, it is our hope that the effect is deeper, as evidenced also by results showing a link between frequent exams and gains in performance and attitudes in introductory science in general [30, 33] and physics in particular.[29] Embedded assessment may also play an important role, particularly when considering how the assessment is integrated into the software, leading to online peer-teaching.[55] Due to embedded assessment, more students may read more of the pages than they would otherwise (Figs. 4 and 5).

VI. OUTLOOK

We believe that these initial metrics will encourage future researchers and instructors to compare how their students behave within our framework, providing further links between course structure and student behavior in their own courses. We are currently pushing these comparisons further by comparing recent online courses at MIT with blended and online courses from MSU.[56]

A major issue touched upon in this study is the role of the etext in regard to student habits. We have shown that review of the etext occurs before exams in all courses, and in the presence of bi-weekly exams and embedded assessment, students shift to accessing etexts weekly. This points toward a possible connection that student consider the etext to be a review tool. This is particularly in-

teresting when considering instructor perceptions of the textbook, where the majority of classes are prescribed daily or weekly readings in preparation for learning activities. Our results point toward a possible follow-up study analyzing students perception of etext (or textbooks) and their general habits of incorporating them into their learning processes.

We were unable to truly understand the role of electronic texts as supplements to physical texts, since we had no handle on the usage of the latter. Combining student surveys and interviews with tracking data in future courses should shed more light on this issue.

Another interesting feature involves activity plots (Figs. 2 and 3) showing large spikes prior to exams. These peaks are actually indicative of signals seen in analysis of online-social networks,[57, 58] where dynamic classes form according to varying external stimuli. In the case of etext use before exams, we see behavior similar to time-reversed exogenous peaks. Modeling etext activity in this way could allow researchers to gain general perspectives into human behavior and reactions to exam deadlines. An equally interesting feature of the activity plots is the decrease in etext use after the first exam, which could be correlated to a number of factors: students may find the etext less useful after the first major assessment, student course balance equilibrates after the first exam, perspective on book use changes to it being used primarily as a review tool, etc.

When it comes to studying the effect of cutoff ranges, one may argue that applying the same cutoff to all course pages is overly simplistic: some pages are long and complex, other pages are short and straightforward. A future study may incorporate a more data-driven per-page approach, where the average access time for each page is computed and the “meaningfulness” of an access is defined on a per-page base within standard deviations of this average reading time.

VII. CONCLUSIONS

Online logs offer a window into how students are using electronic text resource throughout physics courses, data which is not readily available for physical textbooks. However, even with this readily available data source, some caution is warranted, as not every access to an electronic page indicates that the student actually interacted with the page. Thus, we found it necessary to use different definitions of “meaningful” transactions throughout the study, but were encouraged that the majority of qualitative results are robust against the particular choice of these cutoffs.

We analyzed etext usage in one fall semester instance of an MIT Mechanics reform course and nearly a decade of large lecture introductory physics courses from MSU. We are encouraged that many of the results are independent of the difference in course population and pedagogy between the MSU and MIT courses, which indicates that we are indeed seeing effects of course *structure*. Many of the results are not necessarily surprising *per se*, but what is surprising is the clarity of their signatures: course structure indeed has profound impact on the usage of electronic texts and by extension likely traditional textbooks.

We found that courses where the electronic text was not a primary resource saw very little usage, particularly after the first exam. The students did not appear to find this additional free resource useful, even though buying the primary textbook was optional and the homework was offered in the same system online and for free, and even though a variant of the same online text was used as primary resource in other courses. Students appear to cling to the primary resource.

Where the online text was the primary resource, not surprisingly, usage was a lot higher. However, in courses that were otherwise taught traditionally with conventional assessments (weekly end-of-chapter online homework and a small number of midterms), access logs show

a strong pattern of “cramming” before exams — in spite of weekly homework, the text is not used consistently throughout the semester. Also, students in these courses read the online text selectively, simply ignoring some of the content.

This changes when introducing assessment embedded in the electronic text, combined with frequent summative assessment: students read the electronic text more consistently throughout the semester, signatures of cramming are reduced (except right before the final exam), and overall more students read more of the text. As cramming is associated with steeper “forgetting curves,” the hope is that retention is increased though the reformed course structure. Combined with earlier results that the reformed course structure leads to a more positive attitude toward the course in general and less unproductive problem-solving behavior,[29] more frequent assessment appears to lead to both qualitatively and quantitatively more effective learning environments.

As instructors, we strive to understand how learning materials and course activities should be tailored to maximize individual student learning. Our results show a significant link between elements of course structure and student electronic textbook use, motivating us to plan future studies of how student behavior impacts student outcomes. It is unclear how much of the results for electronic textbooks transfer to traditional textbooks, but it is probably reasonable to assume that learners are also not using these resources in the way that their instructor expects or intends.

Acknowledgments

This work is partially supported, but is not endorsed by, NSF grant DUE-1044294 and Google. We thank Sarah C. Seaton for reviewing our article. We would also like to thank the anonymous reviewers of this journal for their helpful comments and suggestions.

-
- [1] D. Butler, *The textbook of the future*, Nature **458**, 0 (2009), ISSN 0028-0836, URL <http://www.nature.com/doi/10.1038/news.2009.202>.
 - [2] P. Arroway, E. Davenport, and G. Xu, *Fiscal Year 2009 Summary Report* (2010).
 - [3] S. D. Smith and J. B. Caruso, *Key Findings The ECAR Study of Undergraduate Students and Information Technology, 2010 Key Findings*, pp. 1–13 (2010).
 - [4] S. D. Smith and J. B. Caruso, *What Are Students Doing with Technology?* (Educause, 2010), URL <http://anitacrawley.net/Resources/Reports/ECARstudyhighlights.pdf>.
 - [5] G. Kortemeyer, *Ten years later: Why open educational resources have not noticeably affected higher education, and why we should care*, EDUCAUSE Review (2013).
 - [6] J. G. Mazoue, *The mooc model: Challenging traditional education*, EDUCAUSE Review (2013).
 - [7] D. T. Seaton, Y. Bergener, I. Chuang, P. Mitros, and D. E. Pritchard, *Who does what in a massive open online course?*, Accepted - Communications of the ACM (2013).
 - [8] E. F. Gehringer, B. L. Golub, R. Cohen, D. M. Arnow, and C. A. Shaffer, in *SIGCSE '13 Proceeding of the 44th ACM technical symposium on Computer science education* (2013), pp. 633–634.
 - [9] K. C. S. U. Cummings, T. Y. U. French, and P. J. M. U. Cooney, *Student Textbook Use in Introductory Physics*, Proceedings of the Physics Education Research Conference (2002).
 - [10] B. D. Smith and D. C. Jacobs, *TextRev: A Window into How General and Organic Chemistry Students Use Textbook Resources*, Journal of Chemical Education **80**, 99 (2003), ISSN 0021-9584, URL <http://pubs.acs.org/doi/abs/10.1021/ed080p99>.
 - [11] N. Podolefsky and N. Finkelstein, *The Perceived Value*

- of *College Physics Textbooks: Students and Instructors May Not See Eye to Eye*, *The Physics Teacher* **44**, 338 (2006), ISSN 0031921X, URL <http://link.aip.org/link/PHTEAH/v44/i6/p338/s1&Agg=doi>.
- [12] T. Stelzer, G. Gladding, J. P. Mestre, and D. T. Brookes, *Comparing the efficacy of multimedia modules with traditional textbooks for learning introductory physics content*, *American Journal of Physics* **77**, 184 (2009), ISSN 00029505, URL <http://link.aip.org/link/AJPIAS/v77/i2/p184/s1&Agg=doi>.
- [13] T. Peckham and G. McCalla, in *5th International Conference on Educational Data Mining* (2012).
- [14] A. J. Rockinson-Szapkiw, J. Courduff, K. Carter, and D. Bennett, *Electronic versus traditional print textbooks: A comparison study on the influence of university students' learning*, *Computers & Education* **63**, 259 (2013), ISSN 03601315, URL <http://linkinghub.elsevier.com/retrieve/pii/S0360131512002953>.
- [15] P. M. Kotas, J. E. Finck, and M. Horoi, *Homework Habits Of College Physics Students*, **1**, 53 (1998).
- [16] L. Jiang, J. Elen, and G. Clarebout, *The relationships between learner variables, tool-usage behaviour and performance*, *Computers in Human Behavior* **25**, 501 (2009), ISSN 07475632, URL <http://linkinghub.elsevier.com/retrieve/pii/S0747563208002057>.
- [17] G. Kortemeyer, *Gender differences in the use of an online homework system in an introductory physics course*, *Physical Review Special Topics - Physics Education Research* **5**, 1 (2009), ISSN 1554-9178, URL <http://link.aps.org/doi/10.1103/PhysRevSTPER.5.010107>.
- [18] D. J. Palazzo, R. Warnakulasooriya, and D. E. Pritchard, *Patterns, correlates, and reduction of homework copying*, *Physical Review Special Topics Physics Education Research* **6**, 1 (2010), ISSN 15549178, URL <http://link.aps.org/doi/10.1103/PhysRevSTPER.6.010104>.
- [19] R. Phillips, M. Dorit, W. Cumming-potvin, P. Roberts, J. Herrington, G. Preston, M. Elizabeth, and L. Perry, in *Proceedings ascilite 2011* (Hobart, Australia, 2011), pp. 997–1007.
- [20] *Learning online network - computer assisted personal approach* (2013), URL <http://www.loncapa.org>.
- [21] R. Mayer, *Multimédia Learning* (University Press, 2001), ISBN 9780521787499, URL <http://books.google.com/books?id=yMJ9o-w\6WEC>.
- [22] B. Minaei-Bidgoli, Ph.D. thesis, Michigan State University (2004), URL <http://www.lon-capa.org/papers/BehrouzThesisRevised.pdf>.
- [23] R. E. Teodorescu, D. T. Seaton, C. N. Cardamone, S. Rayyan, J. E. Abbott, A. Barrantes, A. Pawl, and D. E. Pritchard, in *AIP Conference Proceedings* (2012), vol. 1413, pp. 81–84, ISSN 0094243X, URL <http://link.aip.org/link/?APCPCS/1413/81/1>.
- [24] H. Lin, *Learning physics vs. passing courses*, *Phys. Teach.* **20**, 151 (1982).
- [25] V. L. Dickinson and L. B. Flick, *Beating the System: Course Structure and Student Strategies in a Traditional Introductory Undergraduate Physics Course for Nonmajors*, *School Science and Mathematics* **98**, 238 (1998).
- [26] E. F. Redish, *Teaching Physics* (Wiley, 2003), ISBN 0-471-39378-9.
- [27] N. Finkelstein, *Learning Physics in Context: A study of student learning about electricity and magnetism*, *International Journal of Science Education* **27**, 1187 (2005).
- [28] S. Pollock, in *AIP Conf. Proc.* (2005), vol. 790, p. 137.
- [29] J. T. Laverty, W. Bauer, G. Kortemeyer, and G. Westfall, *Want to Reduce Guessing and Cheating While Making Students Happier? Give More Exams!*, *The Physics Teacher* **50**, 540 (2012), ISSN 0031921X, URL <http://link.aip.org/link/?PHTEAH/50/540/1>.
- [30] S. Freeman, D. Haak, and M. P. Wenderoth, *Increased course structure improves performance in introductory biology.*, *CBE life sciences education* **10**, 175 (2011), ISSN 1931-7913, URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3105924&tool=pmcentrez&rendertype=abstract>.
- [31] P. W. Laws, *Millikan Lecture 1996: Promoting active learning based on physics education research in introductory physics courses*, *American Journal of Physics* **65**, 14 (1997).
- [32] R. K. Thornton and D. R. Sokoloff, *Assessing student learning of Newtons laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula*, *American Journal of Physics* **66**, 338 (1998).
- [33] D. C. Haak, J. HilleRisLambers, E. Pitre, and S. Freeman, *Increased structure and active learning reduce the achievement gap in introductory biology.*, *Science* (New York, N.Y.) **332**, 1213 (2011), ISSN 1095-9203, URL <http://www.ncbi.nlm.nih.gov/pubmed/21636776>.
- [34] L. Deslauriers, E. Schelew, and C. Wieman, *Improved learning in a large-enrollment physics class.*, *Science* (New York, N.Y.) **332**, 862 (2011), ISSN 1095-9203, URL <http://www.sciencemag.org/content/332/6031/862.abstract>.
- [35] G. M. Novak, A. Gavrin, and C. Wolfgang, *Just-in-time teaching: Blending active learning with web technology* (Prentice Hall PTR, 1999).
- [36] H. Sadaghiani, *Using multimedia learning modules in a hybrid-online course in electricity and magnetism*, *Physical Review Special Topics - Physics Education Research* **7**, 010102 (2011), ISSN 1554-9178, URL <http://link.aps.org/doi/10.1103/PhysRevSTPER.7.010102>.
- [37] H. Sadaghiani, *Controlled study on the effectiveness of multimedia learning modules for teaching mechanics*, *Physical Review Special Topics - Physics Education Research* **8**, 1 (2012), ISSN 1554-9178, URL <http://link.aps.org/doi/10.1103/PhysRevSTPER.8.010103>.
- [38] C. I. Hovland, *Experimental studies in rote-learning theory. i. reminiscence following learning by massed and by distributed practice.*, *Journal of Experimental Psychology* **22**, 201 (1938).
- [39] C. I. Hovland, *Experimental studies in rote-learning theory. vi. comparison of retention following learning to same criterion by massed and distributed practice.*, *Journal of Experimental Psychology* **26**, 568 (1940).
- [40] G. D. A. B. R. Seabrook and J. E. Solity, *Distributed and massed practice: from laboratory to classroom.*, *Applied Cognitive Psychology* **19**, 107 (2005).
- [41] N. Cepeda, N. Coburn, D. Rohrer, J. Wixted, M. Mozer, and H. Pashle, *Optimizing distributed practice: Theoretical analysis and practical implications.*, *Experimental Psychology* **56**, 236 (2009).
- [42] *Multimedia physics* (2013), URL <http://www.pa.msu.edu/~bauer/mmp/>.
- [43] R. E. Teodorescu, A. Pawl, S. Rayyan, A. Barrantes, D. E. Pritchard, C. Singh, M. Sabella, and S. Rebello, in *Proceedings of the Physics Education Research Conference* (2010), pp. 321–324, URL <http://link.aip.org/>

- link/APCPCS/v1289/i1/p321/s1\&Agg=doi.
- [44] A. Pawl, A. Barrantes, D. E. Pritchard, M. Sabella, C. Henderson, and C. Singh, in *Proceedings of the Physics Education Research Conference* (2009), pp. 51–54, URL <http://link.aip.org/link/APCPCS/v1179/i1/p51/s1\&Agg=doi>.
- [45] S. Rayyan, A. Pawl, A. Barrantes, R. Teodorescu, D. E. Pritchard, C. Singh, M. Sabella, and S. Rebello, in *Proceedings of the Physics Education Research Conference* (2010), pp. 273–276, URL <http://link.aip.org/link/APCPCS/v1289/i1/p273/s1\&Agg=doi>.
- [46] C. Fredericks, S. Rayyan, R. Teodorescu, T. Balint, D. Seaton, and D. E. Pritchard, in *Accepted to the Sixth Conference of MIT's Learning International Network Consortium* (2013).
- [47] D. T. Seaton, J. Reich, S. O. Nesterko, T. Mullaney, J. Waldo, A. D. Ho, and I. Chuang, *8.mrev mechanics review mitx on edx course report - 2013 spring*, MITx Working Paper #12 (2014).
- [48] D. Riendeau, *Flipping the Classroom*, *The Physics Teacher* **50**, 507 (2012).
- [49] D. MacIsaac and K. Falconer, *Reforming Physics Instruction Via RTOP*, *The Physics Teacher* **40**, 479 (2002).
- [50] E. Moran, *Logging : The Dark Art of Tracking and Interpreting Interaction with Hypermedia CALL Materials*, *CALL-EJ Online* **6**, 1 (2004), ISSN 1442-438X.
- [51] L. P. Macfadyen and S. Dawson, *Mining LMS data to develop an early warning system for educators: A proof of concept*, *Computers & Education* **54**, 588 (2010), ISSN 03601315, URL <http://linkinghub.elsevier.com/retrieve/pii/S0360131509002486>.
- [52] E. F. Redish, R. N. Steinberg, and J. M. Saul, *Student expectations in introductory physics*, *Am. J. Phys.* **66**, 212 (1998).
- [53] G. Kortemeyer, *The challenge of teaching introductory physics to premedical students*, *The Physics Teacher* **45**, 552 (2007).
- [54] C. Cabolis, S. Clerides, I. Ioannou, and D. Senft, *A textbook example of international price discrimination*, *Economics Letters* **95**, 91 (2007).
- [55] G. Kortemeyer, *Correlations between student discussion behavior, attitudes, and learning*, *Phys. Rev. ST-PER* **3**, 010101 (2007).
- [56] D. Seaton, Y. Bergner, and D. Pritchard, in *6th International Conference on Educational Data Mining* (2013).
- [57] R. Crane and D. Sornette, *Robust dynamic classes revealed by measuring the response function of a social system*, *Proceedings of the National Academy of Sciences of the United States of America* **105**, 15649 (2008), URL <http://arxiv.org/abs/0803.2189>.
- [58] R. Crane, F. Schweitzer, and D. Sornette, *New Power Law Signature of Media Exposure in Human Response Waiting Time Distributions*, *Physical Review E - Statistical, Nonlinear and Soft Matter Physics* **81**, 4 (2010), URL <http://arxiv.org/abs/0903.1406>.