We sought to determine if replacing lectures with an online adaptive learning platform resulted in changes to academic performance and student evaluations in a biomedical systems course. Using a mixed method, retrospective study approach, we evaluated a biomedical systems course using either lectures or an online adaptive learning platform for content delivery for subsequent classes. While both classes exhibited similar demographics, adaptive learning produced significantly higher score gains on pretest/posttests compared to lectures but more critical student feedback and no change in overall course performance.

Keywords: student evaluation, lecture, adaptive learning, professional, learning

INTRODUCTION

Traditionally, the education of health care professionals involves many hours of lectures in the early years of training followed by clinical experiences towards graduation. This generally follows experiential learning theory where lectures begin the learning cycle by communicating the problem and preparing students for subsequent preclinical and clinical experiences. Lectures are followed by simulation exercise and supervised practice that give learners concrete experiences, followed in turn by instructor feedback that allow students to reflect and conceptualize the material (Kolb, 1984). Students may prefer learning at each stage of this cycle, giving rise to corresponding learner styles (Activists, Reflectors, Theorists, and Pragmatists) (Honey & Mumford, 1995). The initial emphasis on lectures progressing to clinical experiences also corresponds to the escalating level of learning goals in Bloom’s taxonomy. As students progress through professional school, the initial learning goals are centered on simply recalling facts, followed by progressively higher goals for knowledge application and ultimately achieving analytical and creative skills as beginning practitioners (Anderson et al., 2001). Therefore, lectures often serve to disseminate foundational knowledge at the start of health sciences education where the recall of biomedical science facts is emphasized.

However, some opinion leaders in education disparage lectures as ineffective, since they only support passive learning and fail to use educational and technological advances (Reese, 2013). One of the problems with passive learning is the lack of student engagement with the material. In one study of medical students, student engagement, as measured through student heart rate, continually dropped during a lecture-based class, suggesting a lack of student engagement. In contrast, student heart rates increased when students were engaged in discussions or other forms of active learning (Darnell & Krieg, 2019). In another study of medical students, students did not see lectures as useful for learning, but they attend lectures to meet social norms of professionalism or gain insights into what to study (Klender & Notebaert, 2019). A further criticism of lectures is that they are not supportive of achieving higher levels of learning such as the analysis and evaluation skills as described in Bloom’s taxonomy system of learning (Anderson et al., 2001).
Yet, most didactic content in medical education is still delivered through lectures (Association of American Medical Colleges, 2019) as lectures offer several advantages. Lectures can deliver information to large groups of students with relatively little preparation and few resources, which makes this an attractive mode of teaching for instructors short on time and administrators short of instructors. Lectures can also be delivered in ways to encourage interaction and student engagement by not just passively delivering content but focusing on the student learning needs (Saroyan & Snell, 1997). Lectures may also fulfill a social need as most humans prefer learning by hearing from other humans in a social context, which may trace back to ancient humans gathering in a cave planning their next foraging expedition (Charlton, 2006).

The COVID-19 pandemic has accelerated the movement away from in-person lectures to reduce the chance of disease transmission (Zayapragassarazan, 2020). Prior to COVID-19, other factors led to an increasing use of online instruction, such as the necessity of teaching medical science courses in a multicampus system (Green & Hughes, 2013), the need to offer flexible learning opportunities to diverse students in rural and urban settings (Penman & Thalluri, 2014), and the potential of elearning systems to facilitate learning in health science students with disabilities (Rodrigo & Tabuenca, 2020).

In the context of health care education, online teaching generally seems to be as effective as in-person lectures (Richmond et al., 2017; Spickard et al., 2002), but the evidence is still limited to small studies of limited scope. As a subset of online learning, adaptive learning platforms deliver didactic material tailored to an individual student’s previous knowledge as measured with small knowledge checks. Properly implemented, adaptive learning has the potential to lift students out of negative states such as boredom and frustration and keep them engaged with the material (Luckin, 2018). The potential of adaptive learning is seen in recent studies using adaptive learning to teach intellectually disabled students, where this technology increased engagement, albeit with no significant increase in achievement (Standen et al., 2020). For health science students, there is not enough data for systematic reviews (Fontaine et al., 2017), and current studies evaluate the use of adaptive learning for the training of specific techniques.

Adaptive learning platforms available in the health sciences typically use Bayesian networks to analyze student responses to instructor-generated questions. In these Bayesian networks, possible student answers are mapped out and linked to topics and probabilities whether a student has mastered or not mastered the material. Based on a student’s answer, the platform calculates a probability whether a student has achieved mastery in a given topic and feeds the student more material and questions on topics that the student has not mastered yet. For example, prototypes of this type of platform have been evaluated at the University of Chicago Dental School and Thammasat University Dental School for review of basic medical sciences in preparation for board testing (Alwadei et al., 2020; Suebnukarn, 2009). As an example of a more complex and commercially available platform, HumanDx Global Morning Report is a platform for teaching formulation of differential medical diagnoses using case vignettes and vetted diagnoses submitted by physicians. The software provides the user a score and feedback on how well the user’s list of diagnoses lines up with the correct diagnosis, and current research evaluates how this score can be used to guide learning in medical students and residents (Huffman et al., 2018).

In the context of health science education, adaptive learning technology may have several benefits. For example, in a trial of medical students learning radiology techniques, adaptive learning platforms were better received by students and enhanced learning compared to simple online teaching (Wong et al., 2015). Another study of medical students learning cervical pathology found that a well-designed adaptive learning platform can be received positively by students (Samulski et al., 2018). Medical students learning electrocardiogram interpretation also seem to learn faster if an adaptive learning platform is used (Kellman & Krasne, 2018). Similarly, adaptive learning enabled medical residents to learn effective tobacco counseling strategies faster when compared to nonadaptive online learning (Warner et al., 2020). In nursing students, adaptive learning may produce significantly better learning of a procedure compared to a traditional on-campus course using lectures (Morente et al., 2014). While
the existing evidence is encouraging, there is no data on the effectiveness of adaptive learning in the context of an entire biomedical systems course in dental education.

Since adaptive learning in biomedical sciences is still in its infancy, and alternatives to in-person teaching are desired, we aim to provide additional data on the effectiveness of adaptive learning versus in-person lecturing. The primary aim of this study is to determine if the level of factual knowledge attained in a basic medical science systems course for first year dental students differs significantly when course content delivery is switched from lectures to an adaptive learning platform. A secondary aim of this study seeks to evaluate how the content and sentiment of the student feedback given in course evaluations changes in response to switching from lectures to an adaptive learning platform.

MATERIALS AND METHODS
This is a mixed method, retrospective analysis of student learning, academic performance, and student feedback in a course given two consecutive years starting in 2018. The study was approved as “exempt” by the local Institutional Review Board (protocol number X19/TRB/060).

Subjects
This study evaluated the effect of this course on two groups of first-year dental students taking DMD 5175 Blood and Lymphatics System Course in the spring semester of 2018 and spring 2019.

Overall Course Design
This course is a systems-based course that teaches in an integrated, seamless fashion anatomy, histology, biochemistry, physiology, pathology, microbiology, immunology, pharmacology, and medicine topics as they relate to the blood and lymphatic tissues and the practice of dentistry. Teaching in this course is done with a variety of teaching methods such as lectures, case-based learning, independent study, and seminars (see Figure 1).

Student knowledge was assessed with various class projects (2 seminars, 1 peer graded case analysis), an ungraded pretest survey at the beginning of class on the first day of the course, a graded final written exam, and an ungraded posttest survey at the very end of the course. New exams had been created for each year prior to 2018 to discourage students from forming a test question bank, and previous exams were closely guarded against question release. There is no evidence of

---

Figure 1. Teaching Methods for DMD 5175 Blood and Lymphatics System Course.
Note. Both the 2018 and 2019 course iterations employed the same course layout and sequence of activities, shown as gray blocks and displayed in chronological order. Knowledge assessments (pretest, posttest, and final) are shown in white. These were either graded (final), shown as a black bar, or ungraded (pretest, posttest), shown as a box with rounded corners. The only change from 2018 to 2019 was the replacement of lectures with online adaptive learning modules (dark gray).
students possessing old test questions from this course, and there are no positive trends in exam performance overall or for individual questions suggesting otherwise. For the 2018 and 2019 course, the final exam was identical, as well as the pretest/posttest surveys. Scores from both exam and pretest/posttest surveys were used as data for this study. While the pretest/posttest surveys were ungraded, completion of both surveys was incentivized with a small credit towards the course grade (5%), resulting in near complete participation. Students also received small incentives for professionalism and completion of reflection exercises (5% each). Scientific literature assignments for each class assignment were different to prevent plagiarism, and as a way to cover emerging topics in hematology and infectious diseases for each year.

**Intervention**

The course design was kept the same in 2018 and 2019 except for the delivery of didactic material in the introductory sessions before each case. In 2018, didactic material was delivered using scripted, 1–2-hour long lectures given in a traditional lecture hall setting. For 2019, the lecture material was converted word-for-word using the same illustrations into an online, adaptive learning platform (Realizelt, https://realizeltlearning.com/).

To ensure completion of online modules, students were told that their learning progress was monitored by Realizelt, unlike in past lecture courses, and that completion of online modules was mandatory for passing the 2019 course. Realizelt recorded study times between 1 to 5 hours for each module, with the heaviest usage taking place 1 week prior to the final exam.

**Pretest/Posttest Survey**

The pretest/posttest surveys were administered 3 weeks apart at the beginning of the course and at the end of the course. The survey consisted of ten multiple-choice items that were designed to test detailed recall of the various disciplines connected to the blood and lymphatic system.

Pretests/posttests were scored individually against a key and recorded as pretest/posttest survey score for each student in both years. The pretest/posttest surveys seem valid and reliable tools to measure knowledge gain in this course. We have used these surveys for five years and measured consistent average increases each year (average ranges from 1- to 3-point increase on a scale of 0 to 10 between 2014 and 2019). Gains between pretest and posttest surveys correlate weakly with final exam scores in the 2018 and 2019 course (2018: R² = 0.21; 2019: R² = 0.35).

**Exam**

For a more comprehensive summative, one-time assessment of knowledge, we used a 55-item multiple choice exam modeled after the National Dental Board Exam (NDBE), questions that are used in the licensure process of dentists. In this exam, knowledge is assessed with multiple choice questions featuring a detailed stem and four short possible answers. The questions typically are designed in a way that candidates must recall biomedical science information and be able to apply them to a clinical scenario or evaluate their clinical significance. Exam questions tested application, evaluation, and judgment skills (29 of 55 questions); basic recall (15 of 55 questions); and deeper understanding (11 of 55 questions). The basic recall questions were similar in style to the pretest/posttest questions but tested different material. The exam was identical for both years with an assessment score reliability (KR-20) of 0.7.

**Data collection**

We obtained from the college’s Office of Academic Affairs the number of students in each class, along with the percentage of males, average age, ethnicity/race, incoming GPA, and incoming DAT scores. We also obtained pretest and posttest survey results with the number of correct answers, and acquired exam grades and course grades.

**Qualitative Analysis**

Where applicable, we followed the Consolidated Criteria for Reporting Qualitative Studies (COREQ) guidelines (Tong et al., 2007) for designing and reporting qualitative analysis of student course evaluations. For course evaluations, a female college staff member (college-educated, with seven years experience collecting these surveys but not possessing a dental background and not otherwise engaged in teaching) sent an email instructing students to fill out an online form powered by Qualtrics XM (https://www.qualtrics.com/experience-management/). Students received this survey in the same form for other courses and were accustomed to taking this type of survey from previous courses. The students had previous
exposure to RealizeIt as it was introduced in 2018 in other courses and extensively used in 2019. All students of both classes submitted these course evaluation surveys, as completing these surveys was required for course completion.

The relevant course evaluation survey questions for the qualitative analysis portion of this study were “Please provide any additional constructive feedback regarding the instructor’s teaching,” and “Please provide any additional feedback regarding any of the learning activities presented in this course.”

The course evaluation survey content was analyzed for concepts using context analysis principles as described by Berelson (1952) at the word level. Concepts were coded with a predefined set of categories and for frequency of the concept (Table 1). Text was coded as it appears, allowing for minor grammatical variation (i.e., addition of “s” to verbs if the subject is singular, upper/lower case spelling variants, etc.) and only low-level implication (i.e., “enjoy” coded, but not a sentence describing enjoyment without using the word “enjoy”) (see the example for Table 2). This reduced the need for interpretation and ambiguity and kept the coding process organized and consistent for validity. Irrelevant information (i.e., words such as “and” and “he”) was ignored.

Table 1. Categories and Coding Used for Content Analysis of Student Comments.

<table>
<thead>
<tr>
<th>Category</th>
<th>Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enjoyment</td>
<td>“enjoy,” “good,” “like,” “love,” “great”</td>
</tr>
<tr>
<td>Usefulness</td>
<td>“helpful,” “useful,” “relevant,” “clear,” “logical,” “well organized,” “well presented”</td>
</tr>
<tr>
<td>Negative</td>
<td>“poor,” “not helpful,” “disorganized,” “did not like,” “hate,” “dislike,” “useless,” “frustrating”</td>
</tr>
<tr>
<td>Professor-specific</td>
<td>[the professor’s name]</td>
</tr>
<tr>
<td>Lecture</td>
<td>“Lecture”</td>
</tr>
<tr>
<td>RealizeIt</td>
<td>“RealizeIt,” “online”</td>
</tr>
</tbody>
</table>

The occurrences of code words were normalized against the total number of words in all comments and tabulated using Microsoft Excel as described by Mayring (2014). In addition, the total number of comments, comments related generally to course content, and or comments specific to the professor teaching the course were counted. We assume that data saturation was reached given the large number of responses and the amount of feedback provided by each student.

Table 2. Coding and Additional Rules Used for Content Analysis of Student Comments

<table>
<thead>
<tr>
<th>Category</th>
<th>Coding</th>
<th>Additional Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enjoyment</td>
<td>enjoy</td>
<td>“did not count “like” when it meant “as”; did not count “did not like”</td>
</tr>
<tr>
<td></td>
<td>good</td>
<td></td>
</tr>
<tr>
<td></td>
<td>like</td>
<td></td>
</tr>
<tr>
<td></td>
<td>love</td>
<td></td>
</tr>
<tr>
<td></td>
<td>great</td>
<td></td>
</tr>
<tr>
<td></td>
<td>helpful</td>
<td></td>
</tr>
<tr>
<td></td>
<td>useful</td>
<td></td>
</tr>
<tr>
<td></td>
<td>relevant</td>
<td>excluded “not relevant”</td>
</tr>
<tr>
<td>Usefulness</td>
<td>clear</td>
<td>excluded “were not clear,” “unclear,” “would be more clear”</td>
</tr>
<tr>
<td></td>
<td>logical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>organized</td>
<td>Must be preceded by “well”</td>
</tr>
<tr>
<td></td>
<td>presented</td>
<td>Must be preceded by “well”</td>
</tr>
</tbody>
</table>

Example: comment #2 2018: “I enjoyed this class, especially the aspects of case studies. That is our careers in a nutshell. I think that it makes material relevant for us and tries to place those specific blood disorders and disease in front of us in a very real and relatable way. I really liked this course”

Frequency content analysis for comment #2: enjoy (1), relevant (1), like (1).

No other code found. Categories: Enjoyment (1), Usefulness (1), Negative (0), Professor-specific (0), Lecture (0), RealizeIt (0)

Note: Example of context analysis for one student comment. Codes and any additional rules are listed. For the negative, professor-specific, lecture, and RealizeIt categories, there were no additional rules. The code had to appear as is. This was applied to comments as shown here for comment #2 of the 2018 cohort, with each word matching the code highlighted in bold. This produced the raw data for frequency counts, and a count for each category matched as shown.

To further validate the manual coding, machine-based textual analysis was used to capture prominent themes and sentiments using NVivo 12 Plus (https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home/). For this, student responses were imported as PDF files and coded using the auto-code wizard functionality.

Statistical analysis
For statistical calculations and data representation, we used the R statistical package (https://www.r-project.org/). Student average ages, incoming GPA, incoming DAT, and pretest/posttest survey means were compared with the unpaired student t-test, while student gender percentages, ethnicity, and course grades were compared with Fisher’s exact test. Pretest/posttest survey scores and exam scores were found to be parametrically distributed and compared using the student t-test. Keyword count proportions among student comments for the 2015 and
2016 years were compared using chi-square proportion testing.

RESULTS

Since the aim of this study was to determine the effect of replacing lectures with adaptive learning, we first determined if the two groups of students taking this course were comparable. As shown in Table 3, the 2018 and 2019 groups taking this course were similar in size and did not differ significantly in respect to key demographics, incoming grade point average (GPA), and dental aptitude test (DAT) scores. Therefore, we assumed the 2018 and 2019 cohorts to be comparable.

For the purposes of this study, academic performance was measured in three ways: First, as difference in performance on a 10-question, ungraded pretest/posttest survey testing basic recall; second, as performance on the single 55-question multiple choice final exam testing deeper understanding, application, analysis, and evaluation skills; third, as overall course grade that also included assessment of other skills such as scientific literacy and collaboration in addition to knowledge retention.

Both the lecture (2018) and adaptive learning (2019) course iterations lead to increased recall-type knowledge as evidenced by the significantly increased performance on the posttest compared to the pretest (p < 0.05; both courses) suggesting that either method supports student learning (see Figure 2.).

We also noted that the adaptive learning (2019) cohort had a slight, but significantly higher, average pretest score than the 2018 students (2018: 3.01±1.31; 2019: 3.69±1.56; p < 0.05). To test whether lecture and adaptive learning were

![Table 3. Comparison of the Lecture Course and the Adaptive Learning Course.](image)

![Figure 2. Learning as Measured by the Pretest/posttest Surveys in the 2019 Course Using Adaptive Learning](image)

Note: The increase in scores was significant (*)p < 0.05, student t) suggesting the effectiveness of adaptive learning.

![Figure 3. Comparison of Pretest/posttest Score Gains for both Courses](image)

Note: Gains in pretest/posttest scores are significantly higher with the online adaptive learning course (*, p < 0.05, student t), although some students did poorly.
equally effective, we compared pretest/posttest score gains and found that the 2019 course using adaptive learning produced a significantly higher average gain than the lecture-based course (Figure 3, p < 0.05). For individual student performances we noted a large degree of diverse outcomes. With the lecture course, some students achieved exceptionally large pretest/posttest increases with most students gaining little improvement. In contrast, with the course using adaptive learning, a small number of students did exceptionally worse at the posttest while most students achieved more solid improvement compared to the lecture course.

When using the exam score as measure of academic performance, there was no significant difference between the lecture and adaptive learning iterations of the course (Figure 4, p > 0.05).

![Figure 4. Exam Performance for the Lecture Course and the Adaptive Learning Course](image)

Note. Exam performance was essentially the same for both the lecture-based and adaptive learning-based course (p = 0.05, student t). The student who performed excessively poorly (dot) also was one of the students who also experienced a significant decrease in scores with the adaptive learning platform.

The spread in individual exam performances was similar in both years. When using the course grade as the broadest measure of academic performance, the change from lecture to adaptive learning had no obvious impact (2018: A:B = 52:17; 2019 A:B = 51:18; p > 0.05). On a subjective level, we also did not notice any changes in student behavior, expressed themes, or shown sentiment while observing students during peer-grading, group activities, and seminars in this course. There also was no difference in attendance or professionalism ratings for each class.

To explore other impacts created by the change from lecture to adaptive learning, we performed a qualitative analysis of student comments using manual coding (Table 4). While the 2019 class provided more than twice as much comment volume than the 2018 class, it also seemed subjectively more critical of the online approach on a subjective level. This was also mirrored in the qualitative evaluation of the comments, where there was a significantly higher use of keywords associated with “enjoyment” and “usefulness” in the lecture-based version of the course (p < 0.05). There was a slight, but significant decrease in professor-specific comments in 2019 (p < 0.05). While there is no mentioning of the adaptive learning platform in 2018, “lecture” was a more common keyword in 2019 comments than in 2018, as many comments from the 2019 cohort voiced preference of lectures versus adaptive learning (p < 0.05).

<table>
<thead>
<tr>
<th>Category</th>
<th>2018 (Lecture based)</th>
<th>2019 (Online based)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enjoyment</td>
<td>23 (1.19%)</td>
<td>15 (0.36%)</td>
</tr>
<tr>
<td>Usefulness</td>
<td>13 (0.67%)</td>
<td>7 (0.17%)</td>
</tr>
<tr>
<td>Negative</td>
<td>3 (0.16%)</td>
<td>9 (0.22%) ns</td>
</tr>
<tr>
<td>Professor-specific</td>
<td>14 (0.72%)</td>
<td>10 (0.24%)</td>
</tr>
<tr>
<td>Lecture</td>
<td>25 (1.29%)</td>
<td>31 (0.75%) ns</td>
</tr>
<tr>
<td>Realized/online</td>
<td>0 (0%)</td>
<td>63 (1.53%)</td>
</tr>
<tr>
<td>Word count</td>
<td>1935</td>
<td>4126</td>
</tr>
</tbody>
</table>

* p < 0.05 (chi-square)

ns not statistically significant.

Note. There are significant more “enjoyment,” “usefulness,” and professor-specific hits in 2018 versus 2019, and more significant hits related to realized/online in 2019. Overall, the 2019 class provided many more comments, and several of these were related to using realized.

Machine-based coding confirmed the shift in sentiment to more negative evaluations (Figure 5), although at weak statistical significance (p < 0.1), and confirmed the impression obtained from manual coding. As seen with manual coding, the use of adaptive learning triggered more expression of sentiments in general, and there was marked increase in negative sentiment.
In this study, we found that compared to lecture-based teaching, use of an adaptive online learning platform may enhance recall-type learning. However, in this setting, neither mode of content delivery was superior in attaining higher levels of learning as evidenced by similar performances on the final exam that also tested knowledge application, analysis, and evaluation of clinical scenarios. What was concerning from a faculty perspective was the significant shift towards negative sentiment about the course with the introduction of online adaptive learning as evident in both manual and machine-based analysis of course evaluations.

The findings are similar to a previously published study on an adaptive homework response system for college-level general chemistry (Richards-Babb et al., 2018). The adaptive homework response system improved course grades, whereas in this study, adaptive learning showed significantly better pretest/posttest score gains but not course grades. One explanation for this apparent contradiction is that the adaptive learning platform in this study improved basic recall skills but not other skills. The exam in our course tested predominantly higher-level skills such as application, evaluation, and judgment. For the homework response study, the learning platform (ALEKS, https://www.aleks.com/) enabled repetitive practice of general chemistry principles that could readily be applied to American Chemical Society questions (American Chemical Society Chemistry Olympiad Examinations Task Force, 2018). In retrospect, the online adaptive learning modules in this study were geared to recall. Clinical application was taught in the small group/large group, case-based learning activities of this course, which were unaltered between 2018 and 2019. Consequently, any recall benefits from our adaptive learning platform were likely too small to be reflected in the exam or course grades.

Another explanation for the difference in exam performance may be the design of the adaptive learning module. In this study, the modules created with RealizeIt for this course used eight to 12 questions on specific concepts to test learning progress, whereas the commercial adaptive homework response system (ALEKS) used in the chemistry study utilized 20 to 30 questions to reinforce general concepts. We believe that the higher number of questions in ALEKS makes it a more effective learning platform as it is more likely to identify individual student weaknesses. While ALEKS seems to be effective for selected basic sciences, there is no ALEKS product that is relevant for medical education. Therefore, it may be useful to test in a further study if, similar to ALEKS, the addition of 20 to 30 questions to the RealizeIt platform can improve its effectiveness in the medical sciences.

While the exam tested a variety of Bloom’s revised taxonomy skills (Anderson et al., 2001), the pretest/posttest survey tested exclusively recall type skills. We did measure a significantly higher gain in pretest/posttest scores with the adaptive learning enabled course, but it did not correlate with a higher exam performance even though survey and exam performance are related. This seems to indicate that adaptive learning may be more adept than lecturing for fostering recall-type skills. In contrast, other components of the course (i.e., case-based learning, seminars, independent study) were not changed, resulting in similar exam scores in both years. Therefore, it is plausible that adaptive learning may lead to a more noticeable

**DISCUSSION**

Note. The change from a lecture to an online adaptive learning format in this course produced significantly more sentiments, and a weakly statistically significant shift to more negative sentiment (p < 0.1), as determined by machine-coded textual analysis of student evaluations.
performance improvement in courses that only utilize lectures and assessment of recall skills.

We noted that the average 2019 student using the adaptive learning platform spent about 6 hours per module and per week during the course. This appears significantly higher than the 2–3 hours of study time suggested for a typical 1-hour college level class (see The EOP Study Formula, [https://eop.ucsc.edu/advising/Study%20Formula.html] and the Study Time Calculator, [http://www.opt.uab.edu/retentioncouncil/Exhibits/study_time_calculator.html]). It is possible that the study hours recorded by RealizeIt may be inflated if students conduct other computer activities such as social media while running the adaptive learning platform. However, a class representative did verify that classmates spent these hours actively engaged with the adaptive learning platform to absorb “everything” for the exam. It is possible that the increased study hours are a coping mechanism to overcome the reduced media richness of the online material (Daft & Lengel, 1986). Lectures are a rich medium as they contain a presenter’s body language, nonverbal clues, and other signals that may indicate the relative importance of the material presented to students. In addition, students typically asked 4–5 questions per lecture hour in this course, whereas students did not utilize interaction tools within RealizeIt. To overcome this challenge on adaptive learning platforms, online content could feature a “talking head” video of the presenter along with the presented content, or visual clues such as bolding and highlights to emphasize key concepts. It may also be useful to provide a study guide and supplement online material with frequent virtual office hours.

While adaptive learning seemed to benefit most students, we noted that there were two students in the adaptive learning course who performed substantially worse than the worst performing students in the lecture course. While not statistically significant, it raises the possibility that adaptive learning may not be appropriate for some students. Further research needs to clarify if there are student characteristics that necessitate availability of alternative learning strategies.

In this study, students felt poorly about adaptive learning. This seems different from what has been reported elsewhere (Wong et al., 2015). It may be a general theme as adaptive learning as a form of learner-centered teaching usually elicits student resistance (Weimer, 2008). In previous studies, adaptive learning tended to be optional, allowing for self-selection of motivated students, whereas in this course it was mandatory. As pointed out by Barbi Honeycutt in Flipping the Classroom: Practical Advice from Faculty, a better practice resulting in less student resistance would have been to provide alternative media formats, such as short videos of lecture segments (Honeycutt, 2016). A third explanation for the critical feedback may be that the 2019 cohort of students was the first to have to use the RealizeIt adaptive learning platform in most of their coursework.

Yet, the adaptive learning platform already had been introduced to the same group of students in August of the previous year in two other courses, and the students had experience with the platform for 6 months already. Even though negative feedback may be caused by lack of familiarity, frustration with navigating new software likely is not a cause of the negative feedback.

An alternative explanation may be that the use of the platform was mandatory. Mandatory requirements are known to produce more critical student feedback (Nilson, 2020), and other courses using RealizeIt taken by the same cohort of students also received unusually critical evaluations. Recent research suggests that students are more likely to view adaptive learning technology positively if it is given as optional study aid in addition to other forms of instruction (White, 2020).

One difficulty with applying the findings from this study to a wider context is that the learning effectiveness of adaptive learning may vary across different knowledge levels, subjects, and learning styles. Existing knowledge is a major determinant for learning in learning theories (Ertmer & Newby, 2013), and in this study the largest gains in knowledge were seen among students who had pretest scores in the second quartile. In pharmacology students, adaptive learning produced much greater knowledge gains in chemistry than mathematics, where use of adaptive learning was not correlated at all with knowledge gains (Liu et al., 2017). Adaptive learning systems seem to be much more effective for individuals possessing a Theorist learning style, who easily grasp abstract concepts, than Reflectors, who prefer to learn from
concrete examples (Dounas et al., 2019). Therefore, implementation of adaptive learning likely is best done in the context of a well-defined overall learning strategy that considers learner’s diversity and tailors adaptive learning units to specific learning tasks.

A shortcoming of this study may be the use of incentives such as awarding a small amount of credit towards the course grade for completion of pretests and posttests or making submission of a course evaluation mandatory for course completion. As indicated before, this may color student perception to the activity and the course in general, and grade incentives are best avoided in course design. However, since this applied equally in both course iterations, this should not cause differences in performance and sentiment between the course iterations.

In summary, it appears from this and previous studies that adaptive learning may help the rote memorization of facts but not acquiring higher learning skills. As such, current adaptive learning platforms may be valuable as an optional study aid for factual material such as anatomic landmarks, medical terminology, or the characteristics of microbes, cells, or medications. However, the acquisition of higher learning skills likely requires a more advanced complex learning platform that can mimic the interaction that would take place between a health profession student and a clinical faculty. At a minimum, this complex adaptive learning platform must be able to analyze a variety of input, including text and images, and provide rich feedback that guides the student towards clinical competence. This complex learning platform also must be much more intuitive and user friendly than existing programs so that using an adaptive learning platform is not seen as a chore but as a benefit to students.

CONCLUSION

In this study we observed a potentially significant increase in recall-type knowledge when changing a series of six lectures into an equivalent series of online adaptive learning modules. In the context of a system course with multiple graded activities assessing a variety of skills, this change produced similar performances on assessments and course grade distribution. We also noted a significantly more critical student response in the course evaluation when the lectures were replaced with adaptive learning modules. We believe that this can be minimized with the proper orientation of students highlighting the benefits of adaptive learning, by offering adaptive learning as optional resource, and by providing online material that is both rich in media content that offers frequent real-time interaction with the instructor. Future research will need to elucidate how adaptive learning can be optimized when considering student characteristic, subject matter, and course design. Specifically, further research is needed to identify whether the number of knowledge checks is related to improved recall when using adaptive learning platforms.


