Online Tutoring System (MoFaCTS) for Anatomy and Physiology: Implementation and Initial Impressions

Amanda M. Banker, MS[1], Philip I. Pavlik Jr, PhD[2], Andrew Olney, PhD[3], Luke G. Eglington, PhD[4]

[1]Southwest Tennessee Community College, 737 Union Ave, UNS 118G, Memphis, TN, 38103, ambanker@southwest.tn.edu (corresponding author)
[2]University of Memphis, 365 Innovation Dr., Suite 303, Memphis, TN 38152, ppavlik@memphis.edu
[3]University of Memphis, 365 Innovation Dr., Suite 303, Memphis, TN 38152, aolney@memphis.edu
[4]Amplify, 81 Albion St., #3, Medford, MA 02155, luke.eglington.mail@gmail.com

Abstract

The Mobile Fact and Concept Textbook System (MoFaCTS) is an individualized online tutoring system designed to increase information comprehension and retention. It is being implemented in community college anatomy and physiology (A&P) courses for further system development. A&P was selected because it is a very challenging and highly in demand course. MoFaCTS was used to create Cloze (fill-in-the-blank) questions from the course text which students use to practice the course material. The system provides correction and feedback to the student, repeating questions to improve recall. The system also produces detailed progress reports for both faculty and students. A survey given to students showed a moderately positive impression of the systems, with A&P II students responding more positively on some survey items.
https://doi.org/10.21692/haps.2022.012

Key words: Adaptive learning, reading comprehension, memory, mental model, artificial intelligence

Introduction

Anatomy and physiology (A&P) courses are notoriously challenging. They are also required courses for many fields of study. For example, nursing, physical therapy, dietetics, sports sciences, radiological technician, medical assistant, and more require this course as part of the program or as a prerequisite. A&P courses are high enrollment courses with low success rates, a situation which impacts the future academic goals of some students. Many programs require a minimum grade of C in A&P and successful acceptance into a program is necessary for students to continue in their intended field of study. Additionally, students who fail also risk losing financial aid, a consequence that can negatively impact their future academic progress. While there is variability across multiple institutions, an estimate of course success rates (a grade of C or better) ranges between 55 and 65% (Gultice et al. 2015; Harris et al. 2004; Hull et al. 2016). As such, effective new methods to improve reading comprehension and retention, and ultimately course success, could be valuable.

These courses have an impact on the workforce as well. A&P is a key prerequisite for nursing programs. The United States Registered Nurse Workforce Report Card and Shortage Forecast projects an increasing lack of RNs leading to nationwide shortages by 2030, especially in southern and western areas of the country (Juraschek et al. 2019). The Bureau of Labor Statistics expects Registered Nursing (RN) to see huge growth, with 1.09 million job openings by 2024 (Hogan and Roberts 2015). There will be a great need for students who have successfully completed two A&P courses and low A&P completion rates may create a shortage of nurses as demand grows. Methods for improving course success would greatly impact individual students and, more broadly, the allied health workforce. For these reasons, A&P faculty are always searching for ways to improve student success.

The Mobile Fact and Concept Textbook System (MoFaCTS) is an online practice system developed to improve recall and retention of course material with personalized learning tailored to individual student ability and knowledge. It is currently being implemented in community college A&P classes to improve its efficacy, both generally and specifically in this area of study. It is an adaptable system that can use a textbook to generate questions and then present these questions with correction and feedback.

Background

While various reasons for low success rates in A&P have been proposed and evaluated, the most malleable cognitive factors are reading skill and background knowledge in foundational science topics (Harris et al. 2004; Hull et al. 2016). Reading is a perennial problem for college students. Even though 69% of high school graduates attend college, only 37% of high school graduates are proficient readers as defined by the National Assessment of Educational Progress (National Center for Educational Statistics 2015; US Bureau of Labor Statistics 2022). It is no wonder, then, that A&P students struggle with their textbooks, especially when the textbooks

continued on next page

themselves require a high reading level. For example, *Hole's Human Anatomy & Physiology*, a common A&P text also used at Southwest Tennessee Community College (Southwest), has a Lexile score of 1260L, placing it at the upper end of the 12th-grade reading level. Since only approximately half of first-year college students are proficient readers (Olney et al. 2017; National Center for Educational Statistics 2015; US Bureau of Labor Statistics 2022), content-area reading deficiencies pose a serious problem for postsecondary education.

Content-area reading deficiencies start at an early age. They are often marked by a sudden drop in reading scores, particularly for students from low-income families (Chall and Jacobs 1983), as students transition to increased reliance on textbooks for learning (Hirsch 2003; Moss 2005). These early deficiencies often cause affected students to read less, which delays both fluent reading and vocabulary growth, and negatively impacts reading comprehension (Mol and Bus 2011; Torgesen 2004). To address this problem, researchers have called for reading comprehension practice to be embedded in the learning of content areas (National Institute of Child Health and Human Development 2000; Snow 2002). Aligning reading comprehension practice with content area learning can help students less familiar with the content area vocabulary, background knowledge, and grammatical style, all of which can cause serious comprehension deficits (Cromley and Azevedo 2007; Fang 2006; Kintsch 1998; Laufer 2013; Nagy and Townsend 2012).

Cloze questions, also known as fill-in-the-blank questions, are well established for vocabulary and comprehension practice and assessment (Fang 2006; McKeown 1985; National Institute of Child Health and Human Development 2000). When no answer alternatives are provided, filling in the blanks requires a student to look carefully at the context and draw on background knowledge to construct an answer. While it is possible that students could use shallow strategies, which would reduce deep meaning-based learning from the Cloze questions, we have designed MoFaCTS to personally adapt to the student (described below) in a way that emphasizes deep processing of items rather than rote memorization strategies. We expect this deeper processing will promote both fluency and understanding of the content.

MoFaCTS System Development

MoFaCTS was originally developed in 2006-2007 at Carnegie Mellon University (Pavlik Jr. et al. 2007). MoFaCTS was designed to address the lack of scientifically backed learning tools for content that needs repetition. One original context for MoFaCTS development was Chinese language learning which has many words and characters (Pavlik Jr. et al. 2008). Spaced practice is often recommended in such situations and heuristics exist like the Leitner method for flashcards in which frequency of practice is linked to practice success for each card (Leitner 1972).

However, the MoFaCTS system uses a quantitative memory model based on a scientific theory of memory (ACT-R) to estimate learning as the practice progresses. Inferred from this student practice performance, the model estimates optimal decisions about when to practice items for maximal learning (Eglington and Pavlik Jr 2020; Pavlik Jr and Anderson 2008). While comparisons with all possible heuristic algorithms are not possible, optimal learning (using a student model) is expected to be more effective than heuristic algorithms in most situations. The system's goal is to speed simple learning tasks, saving the student time and allowing faster acquisition of key facts in a domain.

In Fall 2019, a three-year grant was awarded through the Institution for Education Sciences (IES) to develop the MoFaCTS System, culminating in a pilot study of the system in the 2022 Academic Year. The current report describes our development progress and student attitudes toward the system. It is useful to note that the work described in this paper represents the first two years of our IES project, and the Department of Education mandates that we focus on developing the system for efficacy testing during this period. We look forward to eventually reporting on the fully developed system and the efficacy test.

A&P at Southwest Tennessee Community College

Southwest is an open-enrollment, two-year community college that offers associate degrees, technical certificates, courses for transfer, and prerequisite courses for many career programs. The Southwest student body is diverse, consisting of 61% Black/African American, 7% Hispanic, 26% White, and 6% Other (TBR: The College System of Tennessee, 2021a, b). Most Southwest students are female (67%), and many qualify as low income (41%) (TBR: The College System of Tennessee, 2021a, b, c). As with many colleges, students are often academically underprepared for college courses. The average American College Testing (ACT) results of first-time, first-year students is consistently below the threshold and therefore requires learning support coursework, which must be completed before registering for A&P (Office of Institutional Research 2018).

Southwest offers many programs in allied health, including nursing, physical therapy assistant, radiological technician, and funeral services. All of these require at least one semester of A&P. As a component of many programs, A&P courses are in high demand. On average, Southwest offers 65 sections of Human A&P I and 47 sections of Human A&P II per year, although not all participated in this study. The students taking these courses often have difficulty succeeding. For example, Southwest Community College's internal research found that for the 2017-2018 academic year, A&P I students had a 35% rate of D, F, or W (withdrawal) grades and A&P II students performed only a little better at a 30% D/F/W rate. These educational challenges are not unique to Southwest students. The development of the MoFaCTS system at Southwest reaches a student population with challenges widespread among college students.

continued on next page

Methods

Content Generation

MoFaCTS creates automatically generated Cloze questions based on text input into the system (Figure 1). Text analysis methods identify key sentences and select words for fill-in-the-blank questions. These methods decide what sentences and words are most important and form the 'backbone' of the text, both at the chapter section level and the overall chapter level (Pavlik Jr. et al. 2020). This backbone is operationalized by identifying the key concepts used in an interconnected way across the text. It defines interconnection as occurring when a concept appears in multiple sentences and when a given sentence includes three or more such concepts. Therefore, these interconnecting sentences bridge key concepts, and their importance can be defined in terms of the importance of the concepts they bridge, which is defined by the number of sentences in which the key concepts occur. Although this operationalization is quite straightforward, it can be challenging to implement automatically because of the complexity of natural language text.

The system uses a technology called coreference resolution to identify the same concept across the text regardless of its precise wording. For example, the first sentence in a paragraph may mention *the human immunodeficiency virus* but then refer to that concept later on as *HIV, the virus,* or *it*. Coreference resolution allows the proper calculation of concept importance by recognizing all these mentions as referring to the same concept. Once the important sentences have been identified, Cloze items are generated by deleting important concepts and replacing them with a blank. The important concepts are identified as previously discussed, but they are further expanded using syntactic and semantic information in the important sentences, specifically syntactic and semantic arguments.

For example, the sentence, *"The unit of genetic information is a gene, which encodes a protein,"* may include *gene* and *protein* as key concepts identified by coreference, but we additionally add *genetic information* as the object of a preposition. Thus, we would create three Cloze items for this sentence, one for each key concept. Once the questions have been generated, instructors can add, delete, or edit questions. The system also automatically generates paraphrased versions of some of the questions (Olney 2021b). These paraphrased versions are intended to reduce rote memorization from the text and improve comprehension. The automatically generated content can be selected or edited for inclusion in a class assignment.

The system provides several ways of organizing content to facilitate effective content navigation. Faculty may view automatically generated questions in the order they appear in the source text, ranked by sentence importance, coreference, or availability of paraphrased versions (Figure 1). Sorting by sentence importance allows faculty to sort the items by the AI (artificial intelligence) algorithms measure of their centrality to the text meaning. This sorting allows quick identification of the least important items, so that faculty may easily remove items that may not be necessary. Coreference sorting finds all cases where the AI algorithm detects an ambiguity in the sentence due to coreference phenomena like pronouns, e.g., *"It arises from the surface of the tibia, passes __________ over the distal end of the tibia, and attaches to bones of the foot."* This item was automatically corrected to *"The tibialis anterior arises from the surface of the tibia, passes __________ over the distal end of the tibia, and attaches to bones of the foot."* However, such correction is imperfect, and the sorting allows faculty to quickly identify those items that need manual checking for deletion of items or reversion of the correction. Finally, the system also paraphrases some of the sentences and offers these paraphrases as additional items.

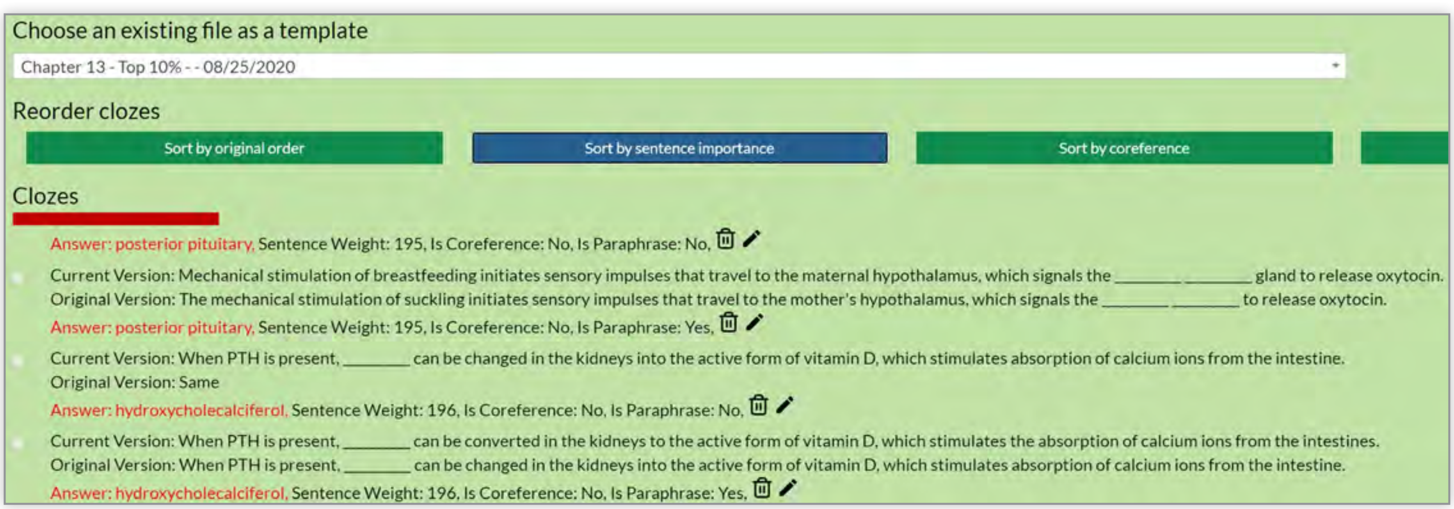

Figure 1. *Faculty content generation screen available to professors that shows the sorting options, editing tools, and paraphrased questions.*

continued on next page

Faculty can further customize their content by selecting the percentage of material made available to the students. A smaller body of questions could be more approachable for students and the percentage of material covered could be increased in updated assignments. To assign the material to the students, faculty members create a class, then assign chapters to each class. Once the material is saved and assigned to the class, it is available for students.

Student Usage

Currently, students log onto the system through the MoFaCTS website using their student username and password. This single sign-on authentication requires coordination with the Information Technology department at the participating school but makes it simpler for students to access the system. Once a student is logged in, they select their professor and course. Before beginning the content questions, students see directions for answering the questions and a consent form. This project was approved by the Institutional Review Board of the University of Memphis, and informed consent was obtained from all participants.

Students were asked if they had read the chapter and encouraged to do so if they had not. Rather than looking up answers, they were instructed to answer as best they could from memory. This practice was intended to move students away from hunting through the text for answers and toward recalling information from memory. Students had 30 seconds to answer a question before the question timed out and moved to the next one. Students were logged out of the system after two time-outs to prevent the system from running unattended and collecting invalid data. In the current course implementation, students were required to complete 30 minutes of practice in the chapter. Thus far, classes have required that practice be completed before the lecture exam on that material. Between fall 2019 and summer 2021, various models of credit were implemented in the classes, including assigning different numbers of chapters, requiring system use, and offering system use for extra credit.

Once a student begins practicing a chapter, the question sequence is controlled using an adaptive algorithm for optimal learning. This algorithm tracks student progress and uses a mathematical model to optimize when to repeat questions to improve comprehension and retention (Pavlik Jr and Anderson 2008; Pavlik Jr. and Eglington 2021a; 2021b). The algorithm uses a student's prior performance in the mathematical model to estimate how well a student knows each Cloze item response. This estimation uses many aspects of their prior learning, including how well they have done, overall, how well they have done in prior repetitions of the Cloze, and the recency of last repetitions for each Cloze. Most importantly, it includes a representation of long-term learning as a function of the difficulty of the items since medium difficulty items (50% correct per trial) have been shown to result in the most learning (Cao et al. 2019; Pavlik Jr. et al. 2019). While the model says this is true, we also have an adjustment because errors are costlier for students (restudy takes time), which leads us to predict optimal learning at about 75% correct. During the learning session, the algorithm tries to give practice for previously seen Cloze items at this level of correctness, using the estimates for all the items to make the selection (Pavlik Jr. and Eglington 2021a). If all prior items are above this level, the algorithm can introduce new items into the currently practiced set.

If a question is answered incorrectly, the student receives feedback in one of the following forms: the correct answer, definitional feedback, refutational feedback, definitional dialogue, and refutational dialogue (Figures 2 and 3). Correct answer feedback tells the student that their answer was incorrect and provides the correct answer to fill in the blank. All other feedback types include this information but also provide additional instruction as described in the next two paragraphs.

Definitional feedback provides the student with the definitions for both the correct concept (the one they missed) and the concept they chose incorrectly. The rationale for providing information on both concepts is that the student's error reveals three kinds of misunderstanding: the correct concept, the incorrect concept, and the relationship between them. The definitional feedback only provides information on the first two kinds of misunderstanding by drawing definitions from the textbook glossary.

Refutational feedback typically includes definitional information but focuses more on the third type of misunderstanding, the relationship between concepts. For example, if the student responded with *digestive tract* instead of *digestive system*, the refutational feedback would explain that the digestive tract is just one part of the digestive system. Refutational feedback is dynamically generated using deep learning and so has the advantage of being closely tailored to the student's error but also the disadvantage of providing good but imperfect answers (Olney 2021a).

In addition to being available in paragraph form, both definitional and refutational feedbacks are available in tutorial dialogue form. The tutorial dialogue form is created by taking each sentence in the corresponding paragraph form and generating questions. For example, consider the questions: "*What is a part of the digestive system?*" or "*What breaks food down into small water-soluble molecules?*". During the dialogue, 2-3 such questions are selected and presented based on the student's ongoing answers, and feedback is given each time the student answers one of these dialogue questions. After the tutorial dialogue feedback is complete, the student returns to Cloze item practice for as long as they wish to continue.

continued on next page

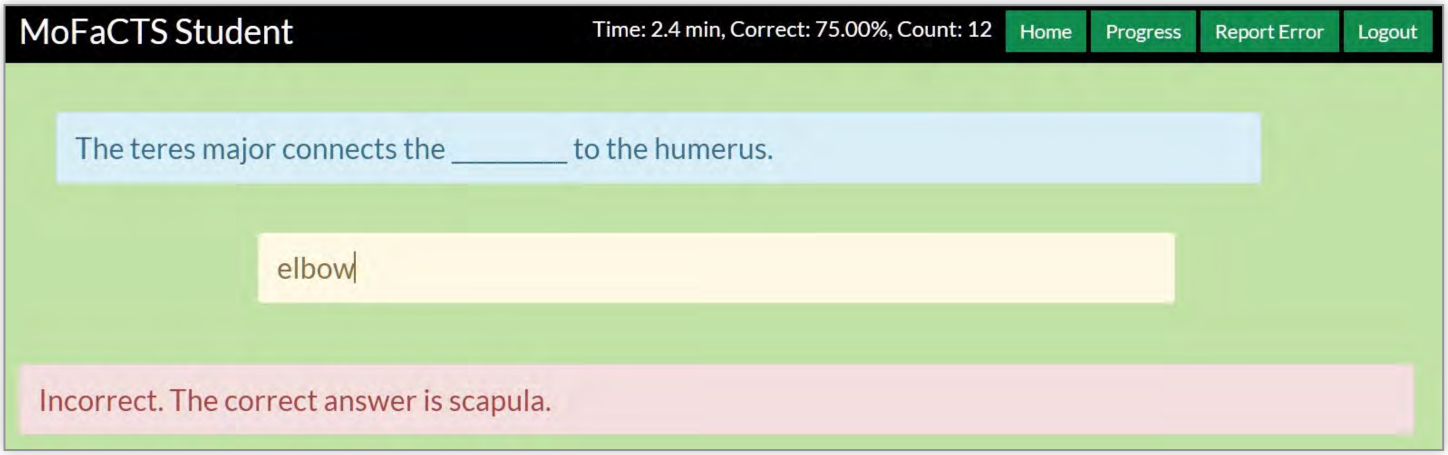

Figure 2. *Cloze item generated from course textbook with incorrect answer and MoFaCTS automatic answer correction.*

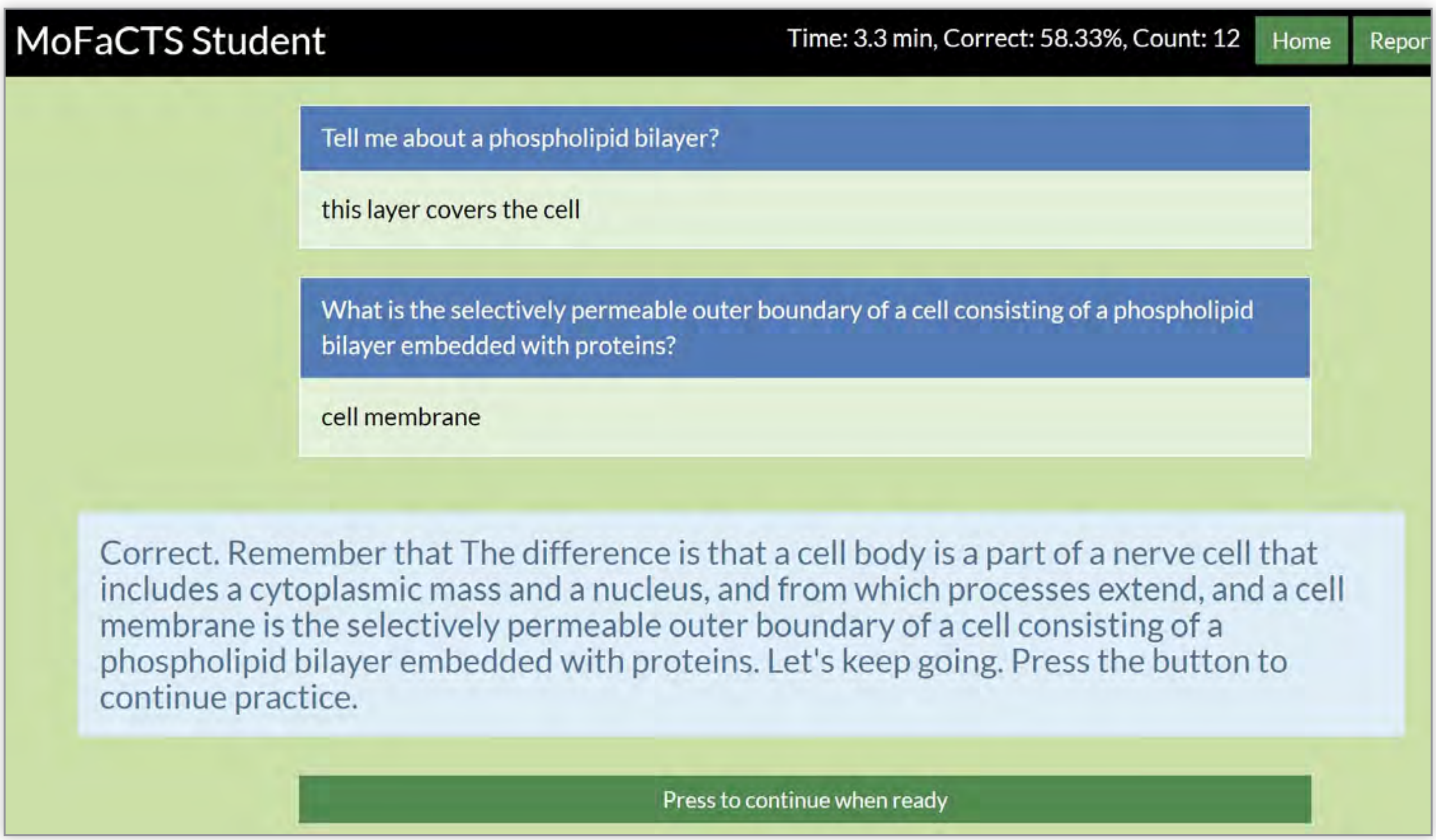

Figure 3. *Cloze item generated from course textbook that was answered incorrectly. The correct answer is cell membrane, but the student answered incorrectly with cell body (not shown). Figure shows automatically generated dialog error correction explaining the error and difference between the two structures.*

continued on next page

Once they have completed a practice session, students can track their progress with a report generated by MoFaCTS (Figure 4). They can see the percent of items answered correctly, the number of practice items completed, and time spent in a particular section. The student report includes the "Correctness Across Repetitions of the Same Fill-in", which is a chart showing student progress as questions are repeated for reinforcement. It also has a "Current Estimate of Recall" chart showing an estimate of recall ability for a particular question, which they can see by moving the mouse over the bar in the chart. This information gives students both numerical and graphical feedback on their progress.

Student Survey Feedback

In continuing to develop the MoFaCTS system, student feedback is very important for us to understand how our development decisions affect students. Each semester students were offered the opportunity to take a survey about their experience with MoFaCTS. The survey asked students to rank their responses to 17 questions on a 6-point Likert scale ("Strongly Agree", "Mostly Agree", "Not Sure, Guess Agree" and the three corresponding disagree statements, with a final uncoded "Did not use" response). Most importantly these results can reveal insights into the student experience depending on other variables such as the course (A&P I or A&P II) in which they used the system.

According to the survey results, the perceptions of the system were marginally positive. A&P I and A&P II were analyzed separately due to differences in knowledge and education experience between two groups of students, differences that may impact how they perceived the system. Small differences in N by question were due to a handful of students who selected "did not use" for some items. It is important to note, when comparing the feedback from A&P I and A&P II students, that the A&P II students have demonstrated success initially because they were required to pass A&P I to be able to enroll in A&P II. Seven survey questions showed significant differences between A&P I and A&P II student responses (Table 1). The analysis reported below comes from a two-way ANOVA in which we controlled for the term during which the survey was submitted, since we did not want differences in the systems deployed across the two years of development to influence the comparison.

For six of these questions, students in A&P II showed significantly higher opinions regarding the system's effectiveness. The only significant measure for which the A&P I students agreed more strongly was the statement, *"I found the MoFaCTS practice items were too often about unimportant details"* (Table 1). This result may reflect the fact that completing A&P I gave students more experience with studying the material and identifying important information

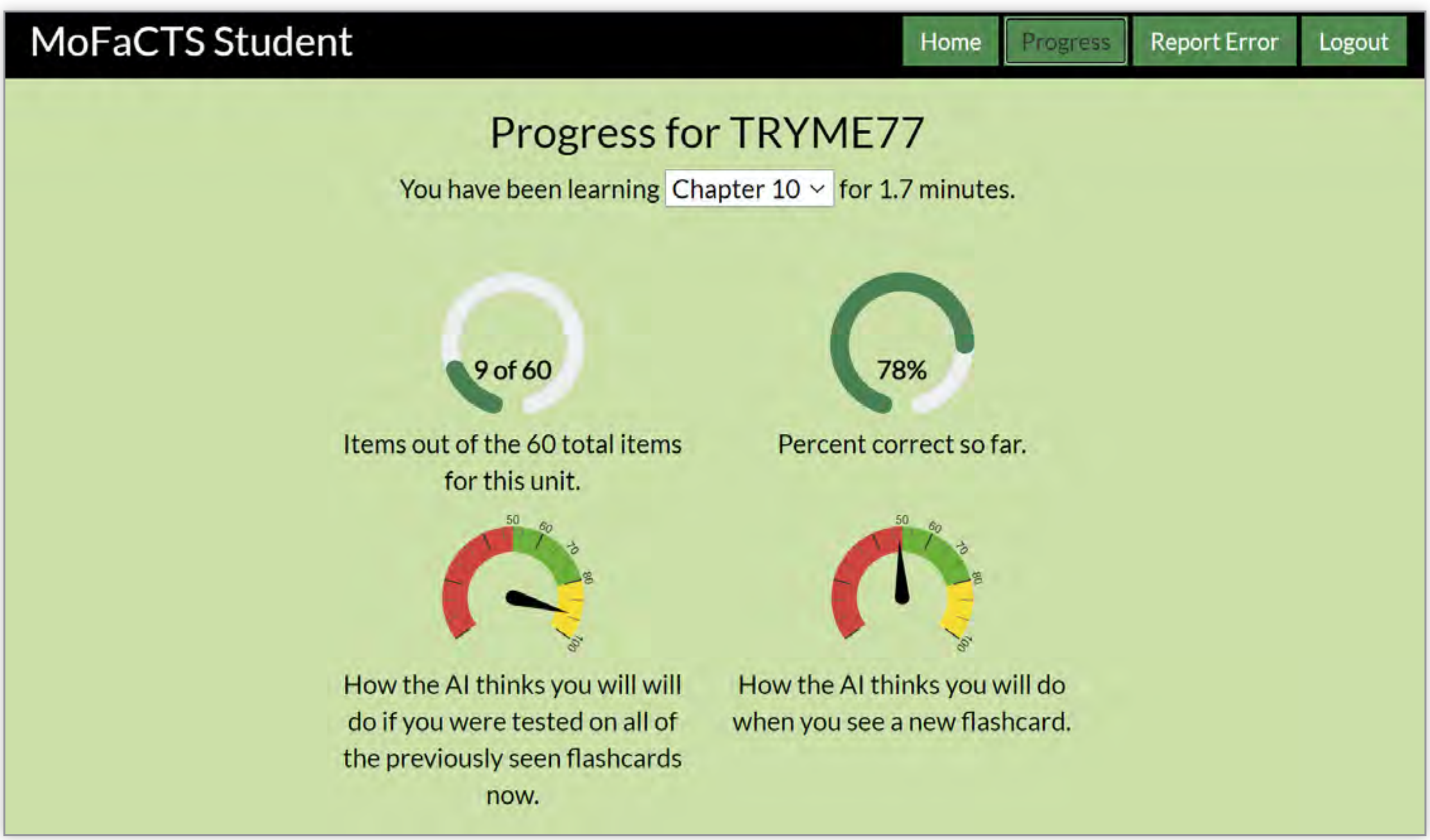

Figure 4. *Progress report for an individual student, including the number of items answered correctly, percent correct, predicted test score, and predicted chance of answering the next item correctly.*

continued on next page

for exams, resulting in their evaluating the system as more useful once in A&P II. Additionally, students who academically struggle in A&P I do not progress and would not be represented in the A&P II student population.

While this survey was not a controlled experiment, it did provide useful feedback on student experiences with the system. It is encouraging that the scores were generally positive, with A&PII students rating the system higher, as mentioned above. This indicated that the system was likely serving A&P II students better, and there is clearly a need to improve the experience for less experienced students. Since A&P I students have less prior knowledge and academic experience, they likely need the support of a tool such as MoFaCTS. This result also agrees with our overall analysis of the difficulty of practice which indicated that it was below what might be considered optimal (Pavlik Jr. and Eglington 2021a). This converging evidence that the algorithm results in practice that is too difficult has caused us to make a substantial adjustment for the final year of system testing.

Faculty Reporting

Faculty can view the summary tables showing class data, including the number of practice items completed, the percent correct, and the time spent by each student. Student work can be viewed by individual sections, usually chapters, and combined over the semester. Faculty can also select an individual student to see a more detailed report on their practice, either by chapter or for all content (Figure 5). The more detailed report compares an individual student's performance with class averages in percent of questions answered correctly, practice item count, and total time spent on practice. These reports allow faculty to keep track of student progress within the system. The faculty reporting capabilities are currently being improved with feedback from previous semesters. In Spring 2022, faculty will have the ability to view completion dates for chapters, making incorporating due dates for chapter practice more feasible.

Question	Term							
	AP I			APII			Comparison	
	N	M	SD	N	M	SD	F-value	p-value
Using the A&P MoFaCTS practice exercises in my class enabled me to learn more quickly.	61	3.48	1.58	27	4.11	1.48	4.94	.029
Using the A&P MoFaCTS practice exercises enhanced my effectiveness in my class.	60	3.61	1.45	27	4.14	1.43	5.48	.022
Using the A&P MoFaCTS practice exercises made it easier to learn.	62	3.61	1.54	27	4.18	1.44	6.45	.012
I found the A&P MoFaCTS practice exercises useful in my class.	61	3.51	1.59	27	4.15	1.54	6.91	.010
I was able to make meaningful connections to learn more deeply using the A&P MoFaCTS practice exercises.	61	3.61	1.60	27	4.19	1.49	4.94	.029
The feedback the MoFaCTS practice exercises gave was adequate to help me learn what things mean.	61	3.90	1.47	26	4.46	1.39	5.45	.022
I found the MoFaCTS practice items were too often about unimportant details.	60	3.5	1.58	27	2.62	1.39	5.71	.019

Table 1. *End of class survey results (AP I and AP II) comparing student perceptions of the MoFaCTS system. Results given only for questions with significant differences (p<0.05). Includes sample size (N), mean (M), and standard deviation (SD).*

continued on next page

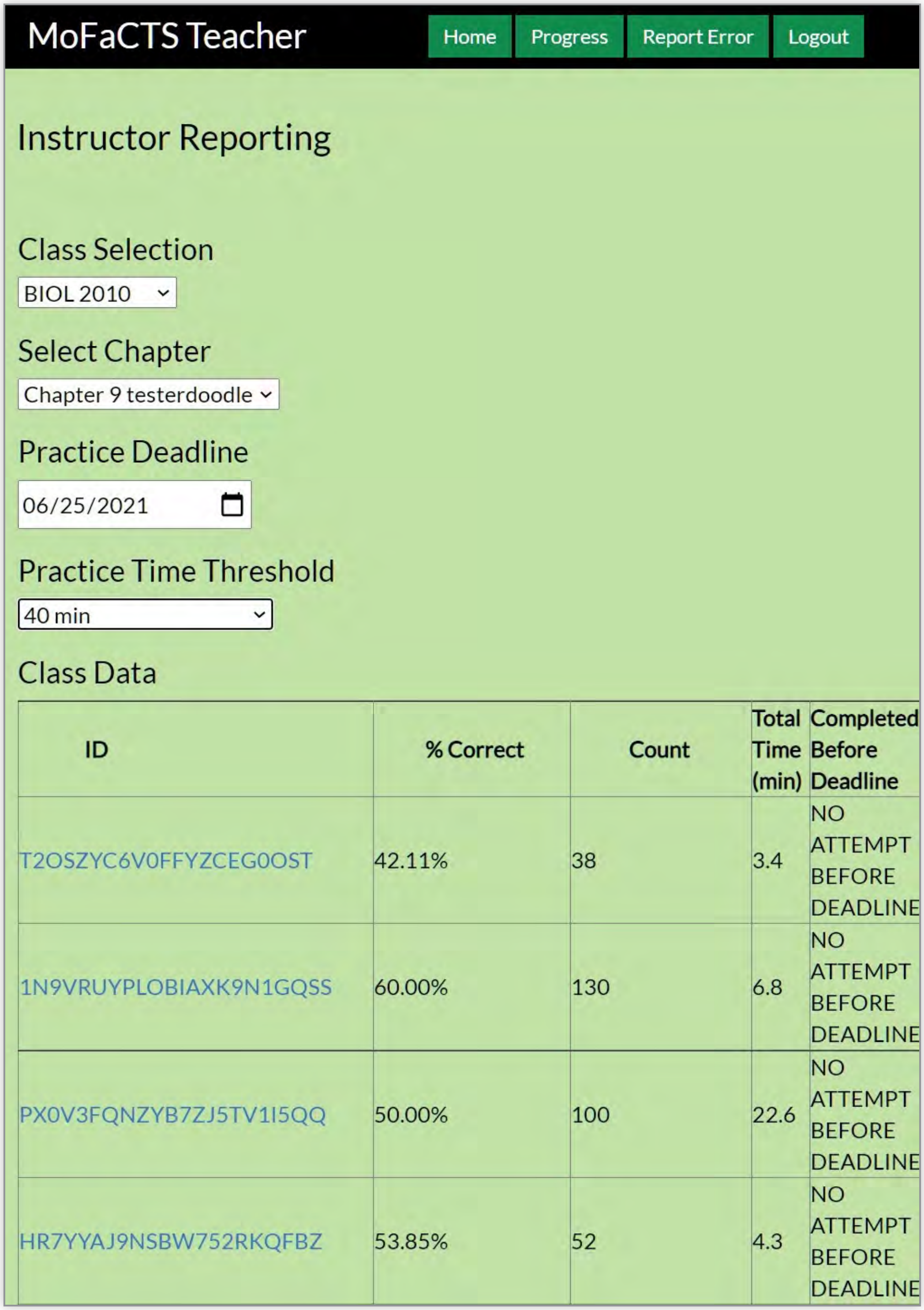

Figure 5. *Summary faculty report showing date, class, and chapter selection options, as well as percent of questions answered correctly, number of questions attempted, and time spent on practice for that chapter for each student.*

continued on next page

Future Development

Improvements in usability for faculty and students will be made in response to continued feedback from participants. In addition to the personalization conferred by the model-based practice scheduling, we intend to further personalize practice for individual students using self-report measures collected from surveys. Survey data collected from students has shown to be predictive of student performance within the system and course exams. We found that refitting past data to include survey responses in the learner model improved accuracy. For instance, a portion of survey questions concerned the students' self-reported reasoning for taking the course, their interest in the content, and the challenge they experienced studying anatomy and physiology. Including their answers to these questions as covariates in the learner model improved predictive accuracy.

As a concrete example, individual differences predicted different learning gains from practice and thus could inform model predictions and pedagogical decisions. Student self-reported ACT scores also predicted student performance. Students with higher self-reported ACT scores learned faster (possibly due to higher levels of prior knowledge and different study habits). Accounting for this in the model could lead to more appropriate pedagogical decisions by MoFaCTS. Students also completed an author recognition survey to estimate their reading behavior, in which they were asked to determine if presented author names were real or fake. Including their score on this test in the model also improved fit and we believe would improve the system's efficacy. In future semesters we aim to collect this data before the students begin using the system to personalize their practice further. This approach may be especially useful early in the students' use of the system when there is little other data yet available.

We are also continuing to improve our artificial intelligence techniques for creating Cloze items from text, paraphrases, elaborated feedback, and tutorial dialogue. Our goal is to closely match the quality of materials that A&P instructors would produce if they had the time to create such materials by hand. We have also recruited nurses to provide feedback on materials in terms of both correctness and usefulness for understanding a concept. Our ongoing efforts are focused on collecting example materials and feedback from experts that will allow us to identify weaknesses in our models so that we can be sure to provide high-quality results that are ideal for learning A&P.

Conclusion

The student survey showed that students had a positive opinion of the MoFaCTS system. A&P II students perceived the system as more valuable and more focused on important information. Over the two years of use of MoFaCTS by students, valuable information has been gained that will lead to improvements in the future. As system development continues, student perceptions of value will hopefully increase. We are planning efficacy testing in fall of 2022 to measure any direct effects of MoFaCTS on student performance. With the substantial challenges facing A&P students, effective ways to improve performance are greatly needed.

Acknowledgment

This work was supported by the Institute of Education Sciences (IES; R305A190448). Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of IES. This work was also supported by the University of Memphis Institute for Intelligent Systems and Southwest Tennessee Community College. Special thanks to Southwest faculty members Robert Blaudow, Gayathri Kaushik, Darrick Slaughter, Marjorie Rothschild, Koushik Roy, and Sheila Bouie for their participation.

About the Authors

Amanda M. Banker, MSc, is an associate professor at Southwest Tennessee Community College, teaches anatomy and physiology, and is the A&P online course coordinator. Philip I. Pavlik Jr., PhD, is an associate professor at the Institute for Intelligent Systems and Psychology at the University of Memphis. His primary research interests include mathematical modeling of learning and memory. Andrew M. Olney, PhD, is a professor at the Institute for Intelligent Systems and Psychology at the University of Memphis. His primary research interest is natural language interfaces. Luke G. Eglington, PhD, is a postdoctoral fellow at the Institute for Intelligent Systems at the University of Memphis. His research focuses on adaptive instructional systems.

Literature Cited

Cao M, Pavlik Jr PI, Bidelman GM. 2019. Incorporating prior practice difficulty into performance factor analysis to model Mandarin tone learning. In: Lynch C, Merceron A, Desmarais M, Nkambou R, editors. *Proc Educ Data Mining Conf (EDM2019)*, Montreal, QC.

Chall JS, Jacobs VA. 1983. Writing and reading in the elementary grades: Developmental trends among low SES children. *Lang Arts* 60(5):617-626.

Cromley JG, Azevedo R. 2007. Testing and refining the direct and inferential mediation model of reading comprehension. *J Educ Psychol* 99(2):311-325. https://doi.org/10.1037/0022-0663.99.2.311

Eglington LG, Pavlik Jr PI. 2020. Optimizing practice scheduling requires quantitative tracking of individual item performance. *npj Sci Learn* 5(1):15. https://doi.org/10.1038/s41539-020-00074-4

Fang Z. 2006. The language demands of science reading in middle school. *Int J Sci Educ* 28(5):491-520. https://doi.org/10.1080/09500690500339092

continued on next page

Gultice A, Witham A, Kallmeyer R. 2015. Are your students ready for anatomy and physiology? Developing tools to identify students at risk for failure. *Adv Physiol Educ* 39(2):108-115. https://doi.org/10.1152/advan.00112.2014

Harris DE, Hannum L, Gupta S. 2004. Contributing factors to student success in anatomy & physiology: Lower outside workload & better preparation. *Amer Biol Teach* 66(3):168-175. https://doi.org/10.2307/4451650

Hirsch ED. 2003. Reading comprehension requires knowledge – of words and the world. *Amer Educ* 27(1):10-29.

Hogan A, Roberts B. 2015. Occupational employment projections to 2024. *Monthly Labor Reviews,* U.S. Bureau of Labor Statistics. https://doi.org/10.21916/mlr.2015.49

Hull KM, Wilson SWG, Hopp R, Schaefer AF, Jackson J. 2016. Determinants of student success in anatomy and physiology: Do prerequisite courses matter? *HAPS Educ* 20(2):38-45. https://doi.org/10.21692/HAPS.2016.005

Juraschek SP, Zhang X, Ranganathan V, Lin VW. 2019. United States registered nurse workforce report card and shortage forecast. *Amer J Med Qual* 34(5):473-481. https://doi.org/10.1177/1062860619873217

Kintsch W. 1998. Comprehension: A paradigm for cognition. Cambridge (MA): Cambridge University Press.

Laufer B. 2013. Lexical thresholds for reading comprehension: What they are and how they can be used for teaching purposes. *TESOL Quarterly* 47(4):867-872. https://doi.org/10.1002/tesq.140

Leitner S. 1972. So lernt man lernen. Der weg zum erfolg. Freiburg im Breisgau (Germany): Verlag Herder.

McKeown MG, Beck IL, Omanson RC, Pople MT. 1985. Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Read Res Q* 20(5):522–535. https://psycnet.apa.org/doi/10.2307/747940

Mol SE, Bus AG. 2011. To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychol Bull* 137(2):267-296. https://doi.org/10.1037/a0021890

Moss B. 2005. Making a case and a place for effective content area literacy instruction in the elementary grades. *The Read Teach* 59(1):46-55. https://doi.org/10.1598/RT.59.1.5

Nagy W, Townsend D. 2012. Words as tools: Learning academic vocabulary as language acquisition. *Read Res Q* 47(1):91-108. https://doi.org/10.1002/RRQ.011

National Center for Education Statistics. 2015. The nation's report card: Mathematics & reading at grade 12. Available from: https://nationsreportcard.gov/reading_math_g12_2015/\#reading/scores

National Institute of Child Health and Human Development. 2000. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Washington (DC): National Institute of Child Health and Human Development.

Office of Institutional Research. 2018. Southwest Tennessee Community College Fall 2009-2018 Historical Fact Book. https://www.southwest.tn.edu/ie/docs/2009f-2018f-historical-fact-sheet.pdf

Olney AM. 2021a. Generating response-specific elaborated feedback using long-form neural question answering. *Proc Eighth ACM Conf Learn @ Scale*:27–36. https://doi.org/10.1145/3430895.3460131

Olney AM. 2021b. Paraphrasing academic text: A study of back-translating anatomy and physiology with transformers. In: Roll I, McNamara D, Sosnovsky S, Luckin R, Dimitrova V, editors. Artificial intelligence in education. AIED 2021. Switzerland AG: Springer International Publishing. p. 279–284. https://doi.org/10.1007/978-3-030-78270-2_50

Olney AM, Walker B, Davis RN, Graesser A. 2017. The reading ability of college freshmen In: Hu X, Barnes T, Hershkovitz A, Paquette L, editors. Proceedings of the10th International Conference on Educational Data Mining. 396-397. https://eric.ed.gov/?id=ED596352

Pavlik Jr PI, Anderson JR. 2008. Using a model to compute the optimal schedule of practice. *J Exp Psychol: Applied* 14(2):101–117. https://psycnet.apa.org/doi/10.1037/1076-898X.14.2.101

Pavlik Jr PI, Bolster T, Wu S-M, Koedinger KR, MacWhinney B. 2008. Using optimally selected drill practice to train basic facts. In: Woolf B, Aimer E, Nkambou R, editors. Proceedings of the 9th international conference on intelligent tutoring systems. Montreal, QC. p. 593–602. https://doi.org/10.1007/978-3-540-69132-7_62

Pavlik Jr PI, Cao M, Eglington L. 2019. Mathematically modeling the optimal desirable difficulty. *Proc 60th Ann Meet Psychonom Soc*, Montreal, QC.

Pavlik Jr PI, Eglington L. 2021a. The mobile fact and concept textbook system (MoFaCTS) computational model and scheduling system. *Proc AIED 2021: Third Workshop Intelligent Textbooks*, 2895:1-15. http://ceur-ws.org/Vol-2895/paper07.pdf

Pavlik Jr PI, Eglington L. 2021b. Modeling the EdNet dataset with logistic regression. *AAAI 2021 Workshop AI Educ*:1-5. https://doi.org/10.48550/arXiv.2105.08150

Pavlik Jr PI, Olney AM, Banker A, Eglington L, Yarbro J. 2020. The mobile fact and concept textbook system (MoFaCTS). *Proc AIED 2020: Second Workshop Intelligent Textbooks* 2674:35–49. https://eric.ed.gov/?id=ED616456

continued on next page

Pavlik Jr PI, Presson N, Dozzi G, Wu S-M, MacWhinney B, Koedinger KR. 2007. The fact (fact and concept training) system: A new tool linking cognitive science with educators. In: McNamara D, Trafton G, editors. Proceedings of the twenty-ninth annual conference of the cognitive science society. Mahwah, NJ: Lawrence Erlbaum. p. 1379–1384. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.68.9072&rep=rep1&type=pdf

Snow CE. 2002. Reading for understanding: Toward an R&D program in reading comprehension. Santa Monica (CA): Rand Education. https://www.jstor.org/stable/10.7249/mr1465oeri

TBR: The College System of Tennessee. 2021a. General enrollment trends. Available from: https://app.powerbi.com/w?r=eyJrIjoiNTY5MTMzZWYtYTQ3Yy00ZTc2LWEzMmUtNTIyMzkxNGY5MmViIiwidCI6Ijc4ZTkwNWIzLTE4ZWEtNGE5MS04YjlmLTMzZTRmZTNjYTQ4YSIsImMiOjN9

TBR: The College System of Tennessee. 2021b. Enrollment by status and demographics. Available from: https://app.powerbi.com/w?r=eyJrIjoiMDJmMGQ2YWQtYjcxMi00NzcxLTgwZjQtMjU5Mzk1OTQyMGIzIiwidCI6Ijc4ZTkwNWIzLTE4ZWEtNGE5MS04YjlmLTMzZTRmZTNjYTQ4YSIsImMiOjN9

TBR: The College System of Tennessee. 2021c. Low income. Available from: https://app.powerbi.com/w?r=eyJrIjoiNTI0YzExYTktMTE2ZS00ZDI3LThkZDgtMmE0YzgxMzYwNmNiIiwidCI6Ijc4ZTkwNWIzLTE4ZWEtNGE5MS04YjlmLTMzZTRmZTNjYTQ4YSIsImMiOjN9

Torgesen JK. 2004. Avoiding the devastating downward spiral: The evidence that early intervention prevents reading failure. *Amer Educ* 28(3):6-19. https://www.aft.org/periodical/american-educator/fall-2004/avoiding-devastating-downward-spiral

US Bureau of Labor Statistics. 2022. College enrollment and work activity of high school graduates news release. Available from: https://www.bls.gov/news.release/hsgec.htm

Back to TOC