KJEP 18:1 (2021), pp. 87-107
Doi: 10.22804/kjep.2021.18.1.005

Adaptive learning system in a statistics course: An experience in Korea and its implications*

Dongseok Kim
KDI School of Public Policy and Management, Korea

Abstract

Adaptive learning is attracting much attention in recent years as a promising alternative teaching and learning method, and is expected to spread to many countries in the near future. The purpose of this paper is to test the effect of an adaptive learning system introduced in a statistics course in a graduate school in Korea, which was the first adoption of a commercial adaptive learning system in a regular credit-bearing course in Korea. Specifically, this paper studies its effect on academic achievement, distribution of test scores, and the relationship between the time spent inside the system and the test scores, using rigorous statistical methodologies. The paper concludes with recommendations for instructors, universities, and the government, to maximize the effect of adopting this learning system.

Keywords: adaptive learning, ALEKS, artificial intelligence, Korea, statistics course

* I am deeply grateful to the professors who have taught the Quantitative Methods course at the KDI School of Public Policy and Management, and I am also grateful to our students who have taken the course. I am particularly grateful to the Teaching and Learning Division of the KDI School. I thank the anonymous reviewers for their comments and suggestions.

Introduction

Adaptive learning is attracting much attention in recent years as a promising alternative teaching and learning method in the era of the Fourth Industrial Revolution. It is said that in the new era, talent and technology will be more important production factors than unskilled labor and physical capital. For this reason, a consensus on the need to foster a workforce equipped with stronger creativity by a more effective education system is being formed. In particular, the perception that standardized uniform education has not been effective in the past decades has led to a strong request for a new effective education system, in particular in Asian countries. A serious concern about the big number of 'su-po students,' the Korean acronym for students who give up on mathematics, in Korea is a good example.

There is no single widely accepted definition of adaptive learning. Kerr (2016) defined it as a way of delivering learning materials online in which the learner's interaction with previous content determines the nature of materials delivered subsequently, with a process that is automated, dynamic, and interactive, and with a purpose of generating a personalized learning experience. Oxman and Wong (2014) defined adaptive learning as a learning process where the content taught or the way content is presented changes, or adapts, based on the responses of individual students.

While the concept of a customized learning experience was suggested decades ago, it is only in recent years that it has been widely implemented thanks to the rapid development of information technology, artificial intelligence (AI), the internet, and various related advances in recent decades.

The essential component of adaptive learning is to assess the knowledge level of students as precisely and quickly as possible, using AI algorithms and big data, and to determine the next questions so that students can follow an optimal customized learning path. Each student follows a different learning path inside an adaptive learning system. This is why the advancement of AI is crucial in the development of adaptive learning systems.

Modern adaptive learning systems are computer-based; students experience the learning process with the help of computers. Also, those systems are web-based, so instructors and students only need a computer connected to the internet, without having to purchase expensive hardware or software. This feature has two important implications. First, web-based systems can be beneficial not only to the users but also to the system itself, since every time a student answers a question, the record is transmitted to the system and the database is updated. In this sense, adaptive learning system shows strong returns to the number of users. Secondly, web-based systems can be useful in urgent or extreme situations. For example, pandemics do not affect web-based systems, so students' learning does not have to be stopped. Also, it can be a good remote learning solution in scarcely populated regions or countries. It is also possible that computer- and web-based learning systems can be useful in the education of gifted children.

Several commercial adaptive learning software have appeared in the education market in recent years, with the two most widely used being ALEKS (Assessment and LEarning in Knowledge Spaces) by McGraw-Hill and MyLab by Pearson. Considering that McGraw-Hill and Pearson are two of the biggest commercial publishing companies in the world textbook market, it can be understood as a typical example of vertical integration.

Due to strong interests and expectations, combined with the recent COVID-19 pandemic, adaptive learning is likely to spread rapidly into many schools at various levels in many countries, including Korea, in the near future. It is highly important and necessary, therefore, to test its effects on various aspects of teaching and learning processes as rigorously as possible.

The goal of this paper is to test the effect of ALEKS introduced into a statistics course in a graduate school in Korea, which was the first adoption of a commercial adaptive learning system in a regular credit-bearing course in Korea. Specifically, this paper studies its effect on academic achievement, distribution of test scores, and the relationship between the time spent inside the system and the test scores, using rigorous statistical methodologies.

The paper is organized as follows. The literature in the field will be reviewed and a brief background history of the adoption of ALEKS in the above-mentioned school will be given. We will test the effects of ALEKS before concluding the paper with policy implications.

Literature review

The field of adaptive learning has grown quickly in the past two decades, resulting in a vast literature. The US Department of Education (2010, 2013) provided macro perspectives regarding the role of adaptive learning, the directions to promote technology-powered education, and recommendations to a variety of stakeholders. Oxman and Wong (2014) provide a brief introduction to adaptive learning systems and their structure, a brief history of the field, and penetration into various educational institutions, types of adaptive learning systems, etc.

Out of various components and aspects of adaptive learning, the concept of customization has attracted the most attention, for it has been regarded as the key to the success of future education. For this reason, many types of customization have been proposed. For example, the US Department of Education (2010) classified customization into three types: the same goals for all students with different speeds ('individualization'), the same goals with different methods ('differentiation'), and different goals, methods and speeds ('personalization'). UNESCO (2012) also provides detailed information about personalized learning.

The rest of this section is dedicated to the review of empirical studies on the effects of adaptive learning, primarily on academic achievement.

Taylor (2006) tested the effect of ALEKS in a college algebra course. The treatment group consisted of 54 students, who were enrolled in an intermediate algebra course using ALEKS, while 39 students in the control group enrolled in a traditional lecture course. All students took the First Year Algebra Test of the National Achievement Test before and after the semester, which were used in testing the effect of ALEKS. Average test score of the students in the treatment group increased from 16.56 in the pretest to 20.26 in the posttest, with increment 3.7, while that in the control group increased from 13.89 in the pretest to 19.67 in the posttest, with increment 5.78, and Taylor concluded that the traditional lecture gave better performance.

Carpenter and Hanna (2006) studied the relative predicting power of students' scores from the initial ALEKS assessment and the Math ACT (American College Testing) score regarding students' success, defined as letter grade C or better, in the calculus course.

According to the authors, Louisiana Tech University introduced ALEKS in MATH340-241 courses, in which students were encouraged to spend at least three hours and make at least 6% progress each week. Carpenter and Hanna found a strong positive correlation among the initial and final ALEKS assessments, total time spent in ALEKS and the final letter grade of the course, and also found that the chance of success in the course doubled if the time spent in ALEKS exceeded 23.5 hours.

The empirical study by Hampikian et al. (2007) was conducted in a more direct way, in which the performance of students enrolled in a precalculus course with the aid of ALEKS was compared with that of students enrolled in the same course without ALEKS. Specifically, out of 440 freshmen in the engineering school at Boise State University, 84 enrolled in a non-compulsory course where ALEKS was used to help students with a Precalculus course, a required course for all engineering students. An additional 37 non-freshman engineering students also enrolled in the course. Hampikian and colleagues found that the average score of ALEKS students was higher than non-ALEKS students in Precalculus although the difference was not significant, and found a positive Pearson correlation coefficient between the time spent in ALEKS and Precalculus score, 0.18, and between ALEKS score and Precalculus score, 0.48. They also reported that 63% of the students said that the course with ALEKS was helpful, and that 66% of students said that ALEKS was helpful in learning mathematics.

The setting of Xu et al. (2009) is similar to that of our study, in that the authors studied the effect of ALEKS in a statistics course in a graduate school. Statistical Methods Applied to Education I course, called 'Stats I,' is required for all graduate students in the College of Education in a mid-south urban university in the US. In this university, the Stats I course was offered with the traditional teaching method in 2005, but was offered in a hybrid format with ALEKS in 2006. The control and the treatment groups in the study of Xu et al. consist of 45 and 41 students enrolled in Stats I in 2005 and 2006, respectively. The authors found a positive effect of ALEKS on test scores, which was statistically insignificant in a t-test, with p-value of 0.20.

Hagerty et al. (2010) studied the case of Black Hills State University in which mathematics courses were completely redesigned in order to enhance the passing rates and enrollment in advanced courses, with the adoption of ALEKS playing a central role. They found that (i) the passing rate in the algebra course improved by 21 percentage points, from 54% in 2002 to 75% in 2006, (ii) students' performance in the CAAP (Collegiate Assessment of Academic Proficiency) test, a national standardized test, increased by 1.1 points, (iii) daily attendance rate increased from 50% in 2001 to 76% in 2006, and (iv) the enrollment in advanced math courses increased.

Nwaogu (2012) evaluated the effectiveness of ALEKS on university students' mathematics achievement. The sample consisted of 80 students in five 5-week mathematics courses, with 16 students in each class. There was no control group in this study since it was not allowed in the university, which means that the study was a one-group pretest-posttest experimental design in which the difference between the scores from both tests was used to study the effect of treatment.[1] The means of the pretest and posttest were 41.09 and 75.64, with standard deviations 23.43 and 19.93, respectively. The p-value of the paired sample t-test was smaller than 0.1%, and the null hypothesis that ALEKS has no effect was rejected. Nwaogu also used the Pearson correlation coefficient method and found that (i) the relationship between weekly times spent in ALEKS and weekly assessments showed a declining trend, with 0.117 in the first week and -0.077 in the last week, and (ii) there is a weak negative relationship, -0.103, between total time spent in ALEKS and the posttest scores, although all correlation coefficients were not statistically significant.

Padilla-Oviedo et al. (2016) compared the effects of three teaching methods—traditional lecture, teaching with ALEKS, and CCA-FOCUS, a teaching program with targeted support and just-in-time content-specific remediation—on students' performance, measured by the final grade in a College Algebra course. CCA-FOCUS was found to have the biggest effect, followed by ALEKS and the traditional teaching method. They also studied the difference in the effects of gender and students' college division, but did not find any statistically significant difference.

The effects of ALEKS in secondary education also have been extensively studied. LaVergne (2007), in a pretest-posttest study on 98 ninth and tenth grade students, found that ALEKS increased the RIT (Rasch UnIT) score[2] of Algebra 1A course by 2.7 points, which was bigger than the district average of 1.0 points and the national average of 1.6 points. Sabo et al. (2013) assessed the effect of two adaptive learning systems, Cognitive Tutor and ALEKS, on high school mathematics performance in algebra and arithmetic. Based on two-way ANOVA, the authors confirmed the statistical significance of the effect of both systems on students' math scores and the statistical insignificance of the difference between the effects of both systems.

Yilmaz (2017) conducted a sophisticated empirical study to assess the effect of ALEKS on students' mathematics performance in the 2014-2015 academic year. The sample consisted of 1,110 students in fifth to ninth grade from two school districts, 555 students from one district in the treatment group and 555 students from the other district in the control group. The students in the treatment group used ALEKS during the academic year for about 45 minutes a day, while the students in the control group followed the traditional school program. The effect of ALEKS was measured at the end of the year based on students' scores of NWEA MAP, a standardized assessment test score. The Mean RIT scores of the treatment and the control groups were 229.13 and 224.61, respectively, with the difference of 4.52. The F-statistic from the ANCOVA was 64.43 with p-value almost zero. Yilmaz reject ed the null hypothesis, and concluded that, after controlling forstudents' initial knowledge level, ALEKS has a significant positive effect on students' mathematics achievement.

Muralidharan et al. (2019) investigated the impact of Mindspark, an adaptive learning system developed by an Indian firm, on middle school students' performance in mathematics and Hindi. In this study, the students with access to the system, the treatment group, scored 0.37σ (standard deviation) higher in mathematics and 0.23σ higher in Hindi, than the students in the control group after only 4.5 months.

We can observe from the above studies that the effect of ALEKS is significantly positive when it is used as an intelligent tutoring system (ITS), that is, when it is used *in addition to* instructor's lectures. This is not surprising because students receive customized tutoring and spend more time in addition to lectures, while students in the control group only receive lectures. We can also observe that the effect of ALEKS is not significant when it is used as a computer-aided instruction (CAI), that is, when it is provided *in place of* instructor's lectures. This is not surprising either because it implies that computer-based teaching has not developed to the extent that it can replace instructor's lectures.

The study of Hu et al. (2011) is surprising in this sense. The authors investigated the performance of the students in two groups in an after-school program intended to improve mathematical skills, where one group of students were taught by certified teachers while the other group worked with ALEKS. Students' performance was measured in a standardized math exam, and the performance in both groups did not show a statistically significant difference.

The effect of adaptive learning systems on outcomes other than mathematics achievement has also been investigated. Hu et al. (2008) demonstrated that ALEKS was effective in reducing performance differences among racial groups in a statistics course. Cheney et al. (2011) also obtained a similar conclusion from an empirical study of the effect of ALEKS in an after-school program for sixth grade students.

Walkington (2013) showed that adaptive learning systems can become more effective by providing learning experiences customized according not only to students' background knowledge levels but also to students' interests, which provides a teaching strategy to increase the effectiveness of adaptive learning systems. Klaverena et al. (2017) found in a large-scale study in Dutch secondary schools, that the effectiveness of adaptive learning systems can be affected by various factors such as study burden, time spent in the adaptive learning system, etc., and can be lower than that of a traditional teaching method.

KDI School's adoption of ALEKS in 2019: Background information

KDI School of Public Policy and Management ('KDI School,' hereafter) is a graduate school in Korea specializing in public and development policy with four master's and two Ph.D. programs. Each year, KDI School admits about 400 students, excluding short-term exchange students, with around 390 master's and 10 doctoral students. On average, 45% of the student body comprises international students, representing about 70 countries, while the school's entire alumni represent 137 countries as of 2020. English is the language of instruction at KDI School.

On average, about 200 students take the Quantitative Methods (QM) course each year. It is an introductory to mid-level methodology course covering major topics in statistics and regression analysis. QM is a required course for students in two master programs. Before the school first adopted ALEKS in the 2019 fall semester, the school offered six QM sections each year on average, with 30~35 students in each section.

KDI School began the process of adopting an adaptive learning system into the curriculum in mid-2017. The school reviewed several options, and arrived at the decision to use ALEKS, and built the bridge between ALEKS and the school's learning management system (LMS) in 2019. In early 2019, the school decided to use ALEKS in the QM course, to run a small-scale simulation course in 2019 summer, and to adopt ALEKS in a QM course in 2019 fall semester as a pilot course.

The simulation course consisted of one instructor, one teaching assistant (TA) and four students, and lasted for two weeks. The simulation course proceeded like a regular course with ALEKS, except that there was no actual lecture. The simulation course was finished successfully, and its experience informed the design of the first actual QM course with ALEKS.

The school offered two QM sections in the 2019 fall semester, and ALEKS was adopted in one of the sections. It was the school's first regular credit-bearing course to adopt an adaptive learning system. In fact, it was the first regular credit-bearing course with an adaptive learning system in Korea. The school decided to adopt ALEKS in all QM sections from 2020.

Results

A brief review of ALEKS system

The ALEKS' statistics course consists of 162 'topics.' A topic is the minimum unit of ALEKS' contents. Each topic consists of a number of questions and supporting materials, such as explanations to the questions, e-textbook, video clips, dictionary of formulae, and so on. ALEKS' questions appear to have been designed in such a way that the highest priority is given to assessing students' knowledge about the basic concepts and main teaching of the topic. It is not easy to find very easy or very difficult questions. Also, most questions are short-answer questions; students need to solve the question and input the final answer in the provided space.

A student is said to have 'learned' a topic when she or he answers certain number of questions correctly, according to ALEKS' scoring system.[3] The instructor classifies the selected topics into a number of groups, and assigns each group to each week. Each group of topics is called an 'objective.' In each week, students are supposed to learn the topics in the week's objective. For this reason, the ideal way to build objectives is to synchronize the objectives and the course schedule as precisely as possible. While there is feasibility in a mild lag between the two, the course is likely to progress more smoothly with their synchronicity.

The instructor can build the schedule of the objectives—the opening and closing times—at her or his disposal. Usually, instructors open an objective at Sunday midnight, and close it at 11:59 pm the following Sunday. For this reason, ALEKS objectives might be regarded as weekly assignments or problem sets from the students' viewpoint.

Knowledge Check is a test with questions from the topics already learned, and is given to students after a certain amount of log-in time. Students take different knowledge checks at different times. If a student fails to answer a question from a topic in a knowledge check, the topic becomes 'unlearned,' and needs to be learned again. A topic is said to have been 'mastered' when a student correctly answers a question from that topic. ALEKS does not ask students to learn mastered topics.

There is one more type of knowledge check, called 'Initial Knowledge Check (IKC),' which is similar to a placement test. Every student must take the IKC, and the questions are from all topics in the course. The one explained above is called 'Progress Knowledge Check.' Topics mastered in the IKC do not have to be learned either, so students with much prior knowledge can advance more quickly. The instructor can inactivate Progress Knowledge Checks, but cannot inactivate IKC.

The first QM course with ALEKS in 2019 Fall

Course setting

The first QM course with ALEKS was offered in the 2020 fall semester. The instructor created an ALEKS class in the system before the beginning of the semester. The instructor selected 73 from 162 topics in the ALEKS' statistics class, and generated nine objectives, which were assigned to weeks two to ten. Progress knowledge checks were inactivated, due to the goals of the course, the average workload of students, evaluation policy, etc. Selecting topics and generating objectives is the most important part of creating a course with ALEKS.

Evaluation of the course was based on students' attendance and participation (5%), performance in ALEKS (25%), two quizzes (20% and 10%), and the final examination (50%). While there are several ways to reflect the performance in ALEKS in an evaluation, the number of topics learned in time (NOTLIT) was used. The NOTLIT is the sum of the NOTLITs in each objective, and henceforth, NOTLIT ÷ 73 × 25% is the formula for a student's score for ALEKS performance.

The first quiz and the final exam were administered in the traditional paper-and-pencil style, while the second was administered as an ALEKS quiz. An ALEKS quiz is similar to an objective, and the questions are from the topics selected by the instructor.

There were 34 students from 20 countries, and two TAs. All students were full-time master's students. During the first class, a staff member from the Teaching and Learning Division helped students open an ALEKS account, then the students took the IKC.

IKC scores

Time spent and the number of topics mastered in the IKC are depicted in Figure 1. During the IKC, 34 students spent 1,567 minutes, with the average and standard deviation being 46.1 and 44.5 minutes, respectively.

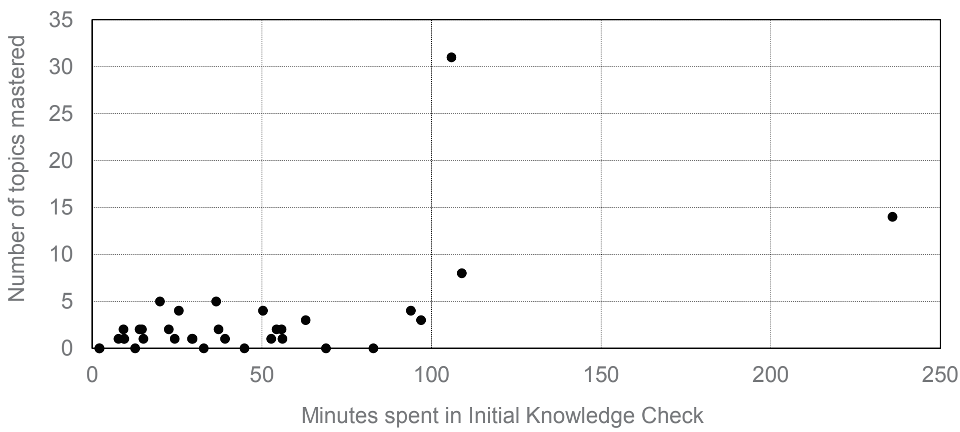


Figure 1. Time spent on the initial knowledge check and topics mastered (2019 Fall QM course)

Note that most students, 31 out of 34, mastered only five or fewer topics out of 73 in the IKC. The average and standard deviation were 3.1 and 5.6, respectively, and were 1.7 and 1.5, respectively, when excluding three outliers who mastered more than 5 topics. This result implies that students' prior knowledge in the course contents was very low. A student's IKC score can be regarded as the ALEKS pretest score.

Log-in time, NOTLIT, and NOTL

Students' work on objectives in ALEKS is the most important part of a course with ALEKS. It is what makes students' efforts and the system itself meaningful. The *raison d'etre* of an adaptive system is to assess students' knowledge level in various topics as quickly and precisely as possible, and to provide a customized learning path to maximize learning and minimize learning time, thus enabling 'learning at her or his own pace.'

Total log-in time might be a reasonable measure of students' engagement. The data from ALEKS reports that 34 students spent 2,752 hours in ALEKS, with a minimum of 29 hours, a maximum of 175 hours, an average of 81 hours, and a standard deviation of 37 hours. About 98% of total log-in time was spent on learning topics, with the remaining 2% on IKC, navigating inside the system, etc.

Students spent an impressive amount of time inside ALEKS, most of which was on 'learning' topics, that is, solving questions, reviewing lecture notes and the textbook, and studying supporting materials. Hence, the NOTLIT could be regarded as a meaningful measure of students' performance in ALEKS. In fact, NOTLIT is the ALEKS posttest score. The NOTLIT as a percent of total number of topics (NOT) in each objective is given in Figure 2. For example, the NOT in the first objective was 11, while the average NOTLIT was 10.7, which was 97.6% of total NOT. It rose to 99.8% in the second objective, and finally reached 100% in the fifth objective. The overall average NOTLIT was 71.8, which was 98.4% of 73 topics.

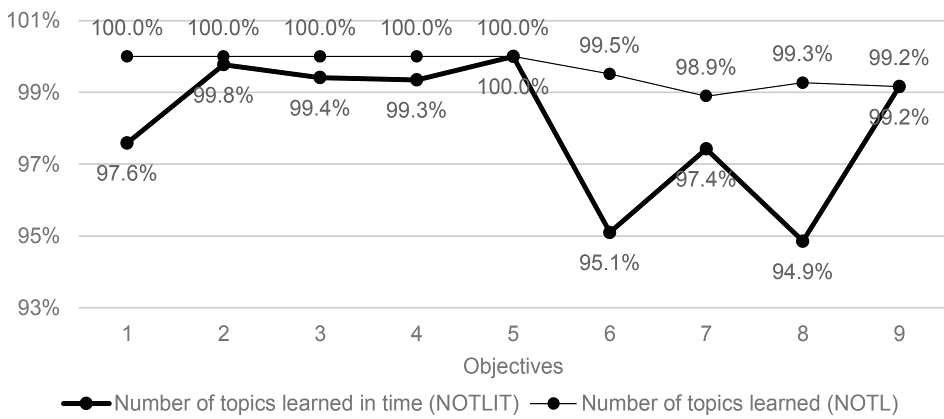


Figure 2. NOTLIT and NOTL (as % of total number of topics, 2019 Fall QM course)

Number of topics learned (NOTL) is also a meaningful measure. It is the number of topics learned by the end of the semester. It can be interpreted as the measure of 'total learning' during the semester. The overall average NOTL was 99.7%. This implies that the students understood and solved questions from 99.7% of the topics in the course.

It is likely that the above result was caused partly by the evaluation scheme. When a topic is not learned in time, about 0.34 point ($= 1 \div 73 \times 25\%$) is lost in evaluation, and, furthermore, students observe the loss in real time. Also, students have sufficient time to work on objectives, unlike a quiz or exam. This appears to have strongly motivated the students to log into ALEKS and 'stay by the computer' until all topics are learned.

Of 34 students in the course, 28 students learned all topics in time, that is, 28 students' NOTLIT was 73. The NOTLITs of the other six students were 71, 70, 69, 68, 64, and 56. These data imply that NOTLIT might not be a proper 'performance' measure since its distribution is extremely concentrated on the full score. Instead, NOTLIT would be regarded as a 'means' or a 'tool' to motivate students to 'work hard' on the course.

When a student finishes learning all topics in an objective in time, the next objective is opened and the student can begin learning the topics in the next objective. When a student has considerable prior knowledge of the topics in an objective, or when a student's schedule allows more time in a week than average, the student can speed up the work in the objective and finish it early. This feature also allows students to follow the course 'at her or his own pace.' In fact, a few students finished learning more than 90% of the topics at around the middle of the course.

Test scores

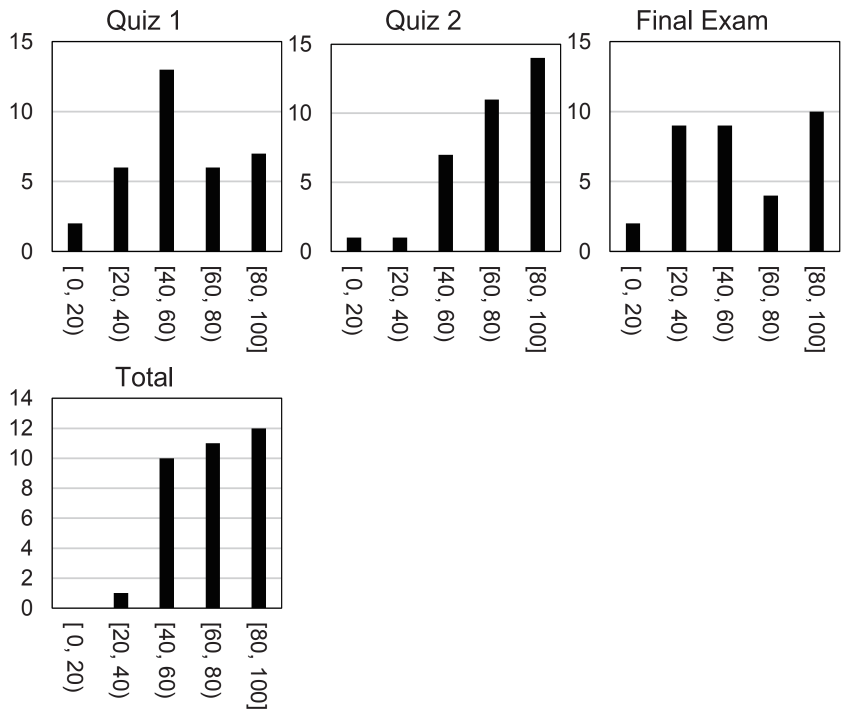
Descriptive statistics of the test scores are summarized in Table 1. Considering the averages and the standard deviations of the two quizzes and the final exam scores, we may conclude that ALEKS quizzes were relatively easier than paper-and-pencil type tests. We also observe that the first quiz and the final exam show highly similar averages and standard deviations, which might reflect the characteristic of the instructor. The high correlation coefficient between the first quiz and the final exam scores, 0.7125, also confirms this conjecture.

Table 1. Descriptive statistics of the test scores (2019 Fall QM course)

	Quiz 1	Quiz 2	Final exam	Total
Weight	10%	10%	50%	100%
Method	Paper & Pencil	ALEKS	Paper & Pencil	-
Average	56.7	72.6	56.9	71.0
Standard Deviation	22.4	20.8	25.9	16.4
Minimum	17.7	8.3	10.5	38.0
Maximum	95.5	100.0	96.0	96.8
Correlation coefficients				
Quiz 2	0.4227			
Final exam	0.7125	0.6248		

Note. Descriptive statistics of attendance (with weigh 5%) and ALEKS progress (with weight 25%) are not included in the table.

A more meaningful comparison can be drawn from the distributions of the test scores. Note from Figure 3 that the distribution of the first quiz scores is almost unimodal and symmetric. The distribution of the final exam scores can also be said to be close to a symmetric and unimodal distribution, although there is a high frequency in the top bracket. On the other hand, the distribution of the ALEKS quiz scores is highly skewed to the left, and we can conclude that it is significantly different from those of the other two. The distribution of the total scores is also left-skewed.



Note. '[40, 60],' for example, represents the interval '40 ≤ score < 60' and [80, 100] represents '80 ≤ score ≤ 100.'

Figure 3. Distributions of test scores (2019 Fall QM course)

Interesting relationships are observed between total log-in times and test scores, as depicted in Figure 4. We found a negative relationship, if not strong, between total log-in times and all test scores. The correlation between the total log-in times and each test score (ρ) is given in each diagram in Figure 4. We also found that most data points lie in a triangle, as depicted in Figure 4, implying that the longer the log-in time, the smaller the variation in test scores.

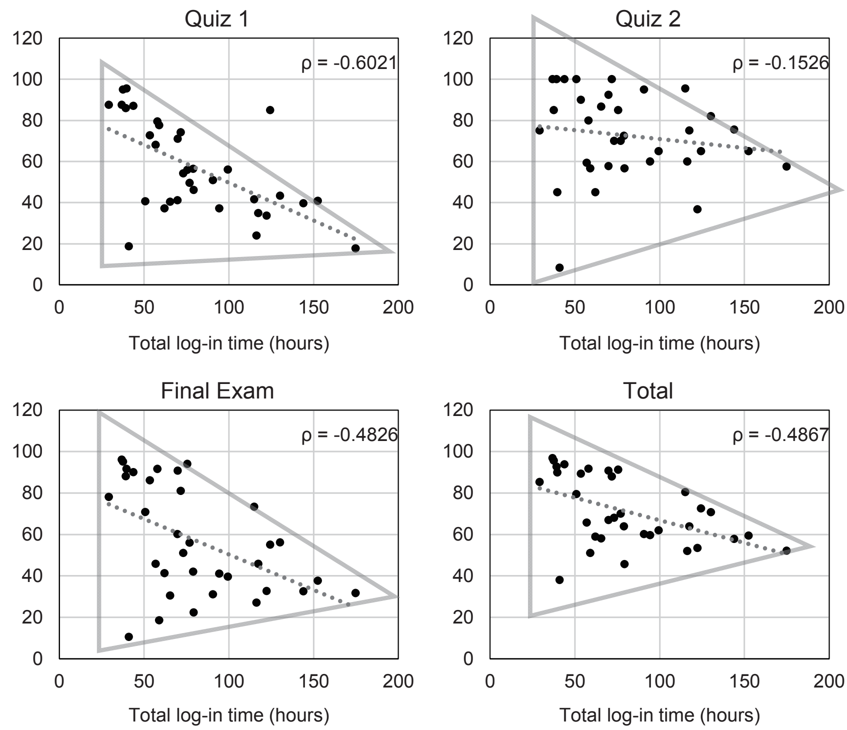


Figure 4. Total log-in time and test scores (2019 Fall QM course)

Hypothesis testing

The following are the hypotheses to be tested in this paper.
[H1] ALEKS has no effect on students' knowledge in statistics and econometrics.
[H2] Test scores follow a normal distribution.
[H3] There is no linear relationship between ALEKS login times and test scores.

Our study is a one-group pretest-posttest quasi-experimental design since we have only one group and the students are not randomly assigned. In this case, IKC scores and the NOTLIT can be used as the pretest and posttest scores, respectively. We ran a paired sample t-test, and the value of the test statistic was $t = 63.6$ with p-value almost zero. We reject [H1], and conclude that ALEKS is effective in improving students' knowledge in the course contents.

Unfortunately, however, we cannot test the effect of ALEKS measured by quizzes or final exam scores because we do not have their pretest scores. At KDI School, constructing a proper control group in such a way that we can test the effectiveness of ALEKS based on the comparison of the performance of treatment and control groups was not possible, as in Nwaogu (2012).

We ran a χ^2 goodness-of-fit test for [H2], that is, the normality of test scores. While

Kolmogorov-Smirnov test is known to be more rigorous, it has the limitation that the distribution parameters in the null hypothesis must not be estimated from the sample (Mood et al., 1974). Since KDI School does not have as long experience of using ALEKS, we cannot assume a well-established distribution of the scores, which is why the Kolmogorov-Smirnov test is not appropriate.

Due to the small sample size, we classified the test scores into the following four ranges: $0 \leq x^2 < \bar{x} - \sigma$, $\bar{x} - \sigma \leq x^2 < \bar{x}$, $\bar{x} \leq x^2 < \bar{x} + \sigma$, and $\bar{x} + \sigma \leq x^2 \leq 100$. Considering that the scores are truncated at 0 and 100, the probabilities were adjusted by dividing by the probability $p\,[0 \leq x^2 \leq 100]$. Table 2 summarizes the test statistics and the p-values for the normality of the scores of the first and second quizzes, final exam and total score. We conclude that when the significance level is 1%, we do not reject the hypothesis [H2] for the first quiz's scores and reject [H2] for other scores.

Table 2. χ^2 Goodness-of-Fit test for the normality of the test scores (2019 Fall QM course)

	Quiz 1	Quiz 2	Final exam	Total
Test statistic	5.6	60.5	21.9	18.2
p-value	1.85×10^{-2}	7.37×10^{-15}	2.88×10^{-6}	1.95×10^{-5}

Note. Degree of freedom of all χ^2 test statistics is 1.

We can test [H3], that is, the significance of the Pearson correlation coefficient, by a t-test. The test statistic is computed by $t = \rho\sqrt{n-2}/\sqrt{1-\rho^2}$ with degree of freedom n-2, where ρ denotes the Pearson correlation coefficient. Test results are given in Table 3. Note that the linear relationship between the total time spent in ALEKS and the second quiz scores is not statistically significant, while those with the first quiz, final exam, and total scores are significant at 5% significance level.

Table 3. t-test for the significance of correlation coefficient (2019 Fall QM course)

	Correlation coefficient between time spent in ALEKS and			
	Quiz 1	Quiz 2	Final exam	Total
ρ	-0.6021	-0.1526	-0.4826	-0.4867
t	-4.2664	-0.8734	-3.1170	-3.1519
p-value	0.0001	0.1945	0.0019	0.0018

Note. Degree of freedom of all t-test statistics is 32.

This result provides an important implication for the instructors who use ALEKS in their classes. It needs to be noted that the time spent in ALEKS has dual meanings. More time spent in ALEKS might imply that the student 'studied more' and it could result in higher test score. On the other hand, the time spent in ALEKS itself might be an indication that the student has less knowledge of the content, so more time in ALEKS might be linked to a lower test score. In fact, the relationship between ALEKS time and test scores might be nonlinear. More specifically, it is possible that the relationship is of inverted-U shape. This might be why we observed highly mixed empirical results.

The second QM course with ALEKS: 2020 Spring

As mentioned earlier, KDI School decided to adopt the ALEKS system in all QM courses from 2020. Two QM courses were offered in the spring semester, one of which was taught by the instructor who taught the QM course with the ALEKS system in the previous semester.

The instructor created a new ALEKS class inside the system. In this semester, the instructor significantly reduced the number of topics, from 73 down to 50, for two reasons. In the previous semester, many students appealed that the workload was too great. The average log-in time was about ten hours per week. Also, the faculty members who teach QM courses agreed with the students. Second, the instructor determined that some topics could be safely deleted from the class for various reasons, for example, in terms of importance, compliance with the purpose of the course, etc. Finally, the instructor created eight objectives with 50 topics.

The evaluation scheme also changed significantly. The weight of the ALEKS progress, that is, the NOTLIT, increased from 25% to 40%, based on experience from the previous semester. Quizzes were replaced with problem sets. Considering the fact that a large portion of students are working in the public sector, problem sets were designed so that students can apply course material to actual situations and can have a 'hands on' experience. The final evaluation scheme was based on attendance and participation (5%), performance in ALEKS (40%), five problem sets (10%), and the final examination (45%), which was administered in the traditional setting. Other ALEKS options were determined almost identically as in the previous course. There were 66 students from 27 countries, all of whom were full-time master's students.

The distribution of the number of mastered topics in the IKC, the pretest ALEKS score, was highly similar to that in the previous semester, but the distribution of time spent was significantly different, with a smaller average and standard deviation, because a shorter time was given.

The total log-in time of 66 students inside ALEKS was 2,821 hours, with the minimum being 14 hours, the maximum 127 hours, the average 43 hours, and the standard deviation 22 hours. The NOTLITs and the NOTLs were higher than 95% in all objectives. The overall average NOTLIT was 49.1, which was 98.2% of 50 topics, and the overall average NOTL was 49.5, which was 98.9% of 50 topics.

Table 4 summarizes the descriptive statistics of the scores of the second course with ALEKS. The average of the final exam, 73.9, is significantly higher than those of the first quiz and the final exam in the first course with ALEKS, 56.7 and 56.9, respectively, and very close to that of the second quiz, 72.6. In fact, the intent of the instructor for the final exam was not simply to present easy questions, but to pose questions that measure students' understanding of basic concepts and the capability to apply the knowledge to a setting that is common in actual decision-making situations. This can be confirmed from the distribution of the final exam scores, which is highly similar to that of the ALEKS quiz in the previous semester. Finally, the Pearson correlation coefficient between ALEKS time and final exam scores was -0.6730 (Figure 5).

Table 4. Descriptive statistics of the test scores (2020 Spring QM course)

	Problem sets	Final exam	Total
Weight	10%	45%	100%
Average	96.1	73.9	87.1
Standard deviation	9.4	20.6	10.2
Minimum	54.0	26.0	55.7
Maximum	100.0	100.0	100.0

Note. (1) Descriptive statistics of attendance (with weigh 5%) and ALEKS progress (with weight 40%) are not included in the table. (2) The correlation coefficient between the problem sets and the final exam scores was 0.1608.

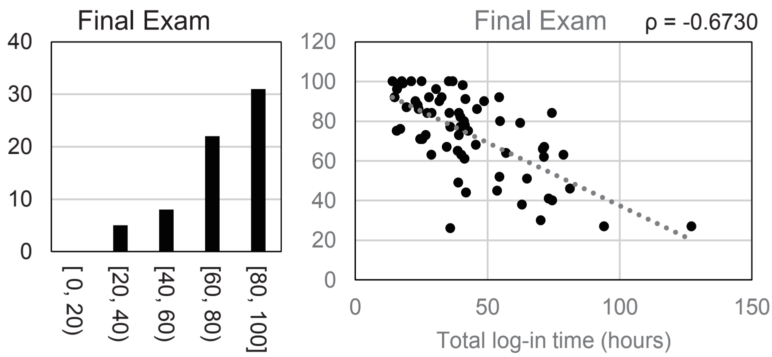


Figure 5. Distribution of the test scores (2020 Spring QM course)

The same set of hypotheses, [H1]~[H3], were tested using same methodology, and the results were similar to those in the previous semester. First, we ran a paired sample *t*-test using the ALEKS' IKC scores and NOTLIT as the pretest and posttest scores, respectively. The value of the test statistic was $t = 93.9$ with a *p*-value of almost zero. We reject [H1], and conclude that ALEKS is effective in improving students' knowledge in the course contents.

Second, we ran a χ^2 goodness-of-fit test to test the normality of the final exam and total scores. The values of the test statistics were 188.8 and 607.4, respectively, with *p*-values of almost zero. We reject the null hypothesis that the scores follow the normal distribution, which is consistent with the first graph in Figure 4.

Third, we ran a *t*-test for the significance of the linear relationship between ALEKS time and the final exam scores. The value of the test statistic was -7.28, with a *p*-value of almost zero. We reject the null hypothesis, and we conclude that the linear relationship is statistically significant.

Indirect indicators of the effectiveness of ALEKS

Academic achievement would be the most important criterion in assessing teaching methods. Unfortunately, however, we cannot directly evaluate the degree to which

ALEKS enhances academic achievement from the experience of KDI School because the adoption of ALEKS was not accompanied by a control group. In this subsection, we will review 'indirect' evidence that might be helpful or meaningful in assessing its effect.

First, the NOTLIT and the NOTL are the strongest indirect indicators, since they signify the amount of knowledge students acquired. Their importance can be discounted because these results are obtained in an uncontrolled setting, that is, students answer questions with access to every resource. However, the experience at KDI school implies that the effect is at least significant. In the first semester, 28 students learned all 73 topics in time, and the NOTLITs of the other six students were 56~71. In the second semester, 51 students learned all 50 topics in time, 12 learned 47~49 topics, and the NOTLITs of other three students were 45, 42 and 33. The NOTL statistics are more impressive. In conclusion, we can expect a considerably high level of knowledge acquisition with the ALEKS system, accompanied by a properly designed course, such as number of topics, evaluation scheme, etc.

Second, students' log-in times can be used as important information regarding the effect of ALEKS. Above all, most students spend considerable amount of time inside ALEKS, and they study, read materials, and try to understand and solve questions in the system. The instructor, however, needs to pay attention to the students' log-in time since too much log-in time can reduce effectiveness.

Third, total log-in time can also be interpreted as a measure of academic deficit, as explained previously. For this reason, average learning time per topic or average number of topics learned per hour (ANOTH) can be a more meaningful indicator. Table 5 summarizes the ANOTH records in both semesters.

Table 5. Average number of topics learned per hour (ANOTH)

	2019 Fall	2020 Spring
Total learning time (hours)	2,697	2,780
Total number of topics learned (except mastered topics)	2,370	3,055
Average number of topics learned per hour (ANOTH)	0.88	1.10

The ANOTH data need to be interpreted carefully since they depend on many factors, such as students' characteristics, the instructor's teaching method, course setup, total number of topics in the semester, that is, students' workload, and the topics' difficulty levels. The ANOTH data of multiple classes can be comparable when most of these factors are controlled.

According to Table 5, the ANOTH in the 2020 spring semester was 1.10, which was 25% greater than that of the previous semester, 0.88. This difference could be the result of many factors: a smaller number of topics, and accumulated know-how in the instructor and in the school's support system, to name a few. In addition, the instructor revised the lecture notes and the teaching style to strengthen the harmony between lectures and ALEKS contents. For example, the instructor (i) provided a matching table among lecture notes and lecture schedule, chapter/sections in the textbook, and ALEKS topics, (ii) used the same mathematical notations in the lecture notes as in ALEKS, (iii) tried to synchronize the lecture schedule with ALEKS objectives, etc.

Fourth, ALEKS is computer-based and web-based, which enables flexible learning speed and real-time monitoring. Students can progress at their own pace, so they can allocate their time more efficiently. At the same time, instructors can monitor students' activity in real time, and provide customized help just in time. Also, TAs can cover customized content based on observed progress of students when there are multiple TA sessions. In other words, ALEKS can increase total learning in a 'semester.'

Fifth, students' satisfaction level can also be used as an indirect measure, or, in assessing the effect of ALEKS. KDI School's Teaching and Learning Division conducted two identical surveys in 0-5 Likert scales after both courses were finished, and the results are summarized in Table 6. Note that the satisfaction levels were considerably high in all aspects. Higher satisfaction level in the second semester seems to be based on more streamlined management of the course and more accumulated knowhow in the second course.

Table 6. Survey results

	Technical	Contents	Integration with face-to-face learning	Others
2019 Fall	4.66	4.47	4.23	4.52
2020 Spring	4.80	4.61	4.61	4.67

It is worth discussing in which courses an adaptive learning system can be useful. Most of all, adaptive learning systems are unsuitable in courses where *excellence* is given high priority because the current technology level in AI algorithms cannot meet this need. The current adaptive learning system tries to assess the knowledge level of students by the way students respond to the system's questions, using AI algorithms and accumulated data. For this reason, adaptive learning systems can be applied where the contents can be decomposed into a group of standardized units. Also, adaptive learning systems can be highly effective in helping students acquire basic knowledge, or 'concept mastery.' This is why current commercial adaptive learning classes are concentrated in introductory mathematics, statistics, and a few basic science courses.

The goal of these classes is not to train the top 5% students, but, instead, to help *most* students acquire basic knowledge and the capacity to apply the learning from the course to actual situations. Therefore, the evaluation scheme of these classes should be designed in such a way that the distribution of scores is not of the traditional bell-shaped form, but is skewed to the left.

Conclusion and recommendations

Adoption of commercial adaptive learning systems is expected to spread rapidly in many countries in the near future. In particular, web-based learning systems can be greatly helpful in overcoming the difficulty caused by the COVID-19 pandemic and other potential obstacles. In Korea, for example, many universities have begun reviewing available options, and a few have already launched courses with adaptive learning systems. It is possible that adaptive learning systems might penetrate secondary and elementary education, but at a lower speed, due to language challenges.

The aim of this paper is to test the effectiveness of ALEKS adopted in a statistics course in a graduate school in Korea, which was the first adoption of ALEKS in a regular credit-bearing course in Korea. We found that ALEKS is effective in enhancing academic achievement, defined as progress within the ALEKS system, based on statistical inference, and we found additional indirect evidence. We also found that the distribution of scores is significantly different from traditional bell-shaped distribution, which has not been reported in the literature. Finally, a significant negative linear relationship was found between the time spent in ALEKS and various test scores.

The study in this paper has several limitations. First, this paper employed a one-group pretest-posttest experimental design because preparing a control group was not possible. Second, the study is based on relatively small samples, 34 and 66 students in two courses. Third, the paper studied only a few aspects of the ALEKS system; more comprehensive studies focusing on many other aspects are necessary. These limitations are likely to limit the reliability and applicability of the results. Further rigorous studies based on larger samples in longer periods are strongly desired, not only for KDI School but also for many educational institutions.

Finally, we derive, based on the literature and the experience of KDI School, the recommendations for instructors, universities, and governments to maximize the effect of the newly arriving technology.

The recommendations for instructors can be summarized as follows. First, it is necessary to establish the rationale for the adoption of the system into the course, that is, the reason for and the goal of adopting an adaptive learning system. Various criteria, such as effect on academic achievement, passing rate, concept mastery, etc., must be thoroughly reviewed.

Second, many factors affect the effectiveness and efficiency of the system, and it is necessary to understand those factors and to take proper precautionary measures to maximize effectiveness and efficiency. For example, it is important to construct objectives carefully since a change in the schedule of an objective can result in the change of all subsequent objectives' schedules, and can affect the efficiency of the course and the trust of the instructor. It should be noted that the adoption of an adaptive learning system could lower the flexibility of course management from the instructor's viewpoint.

Third, the harmonization among lecture notes, textbook, and the system's contents is important in maximizing students' academic achievement. The synchronization of the course progress and the progress inside the system is similarly important. This implies that the instructor needs to invest considerable resources in revising and streamlining course content.

Fourth, the instructor needs to explain the features of the system to students as early as possible. For example, ALEKS is highly impatient with mistakes in rounding numbers. Time wasting and student loss in interest can be prevented by explaining the features as early as possible.

Universities need to set up policies regarding adaptive learning. First, we suggest that universities provide school-wide guideline in adopting an adaptive learning system when the adoption is on a large scale. In Korea, universities are suggested to consider the fact that current commercial solutions are in English only. Second, if there exist university-wide regulations regarding grading on a bell-shaped curve, the courses with adaptive learning need to be allowed a more flexible scheme. The courses using an adaptive learning system might be those in which an absolute grading scheme is desired. Third, the university is responsible for hiring and training staff with expertise in adaptive learning and LMS, as well as various resources and services. For example, universities can reduce instructors' burden by providing the students with detailed information and guidelines about the system.

Finally, the government's role, as in many other areas, is to provide public goods and institutional elements at the proper time. Related laws and regulations need to be reviewed so that those that would make universities or private companies hesitate or become reluctant to adopt an adaptive learning system can be amended in advance. The government also needs to provide universities with information, establish technical standards, and construct platforms where information and best practices can be efficiently shared among universities and private companies. Providing tax incentives in purchasing commercial adaptive learning systems could also be considered.

Address for correspondence

Dongseok Kim
KDI School of Public Policy and Management
263 Namsejong-ro, Sejong-si, 30149, Korea
Email: dongseok@kdischool.ac.kr

References

ALEKS Corporation. (2016a). *Higher education math instructor module reference guide.* ALEKS Corporation. https://www.aleks.com/highered/New_Instructor_Module_for_HE.pdf

ALEKS Corporation. (2016b). *New ALEKS student module reference guide.* ALEKS Corporation. https://www.aleks.com/k12/New_ALEKS_Student_Module_Reference_Guide.pdf

ALEKS Corporation (2017). *Instructor's manual for business statistics version 3.18.* ALEKS Corporation. https://www.aleks.com/manual/pdf/educators-bstat.pdf

Carpenter, J. P., & Hanna, R. E. (2006, June 18-21). *Predicting student preparedness in calculus* [Conference session]. 2006 Annual Conference & Exposition, Chicago, Illinois, United States. https://peer.asee.org/1428

Cheney, K. R., Craig, S. D., Anderson, C., Bargagliotti, A., Graesser, A. C., Sterbinsky, A., Okwumabua, T., & Hu, X. (2011). Closing the knowledge gap in mathematics among sixth grade students using ALEKS. In M. Koehler & P. Mishra (Eds.), *Proceedings of SITE 2011--society for information technology & teacher education international conference* (pp. 1425-1427). Association for the Advancement of Computing in Education (AACE).

Franklin City Schools. (2016). *A parent's guide to MAP: Measures of academic progress.* https://www.franklincityschools.com/media/curriculum/MAPparentbrochure201516.pdf

Hagerty, G., Smith, S., & Goodwin, D. (2010). Redesigning college algebra: Combining educational theory and web-based learning to improve student attitudes and performance. *PRIMUS, 20*(5), 418-437. https://doi.org/10.1080/10511970802354527

Hampikian, J., Guarino, J., Chyung, S. Y., Gardner, J., Moll, A. J., Pyke, P. A., & Schrader, C. B. (2007, June 24-27). *Benefits of a tutorial mathematics program for engineering students enrolled in precalculus: A template for assessment* [Conference session]. 2007 Annual Conference & Exposition, Honolulu, Hawaii, United States. https://peer.asee.org/2584

Hu, X., Craig, S. D., Bargagliotti, A. E., Graesser, A. C., Okwumabua, T., Anderson, C., Cheney, K. R., & Sterbinsky, A. (2012). The effects of a traditional and technology-based after-school setting on 6th grade student's mathematics skills. *Journal of Computers in Mathematics and Science Teaching, 31*(1), 17-38.

Hu, X., Luellen, J. K., Okwumabua, T. M., Xu, Y., & Mo, L. (2008). Observational findings from a web-based intelligent tutoring system: Elimination of racial disparities in an undergraduate behavioral statistics course, mimeo. https://www.academia.edu/22504612

Kerr, P. (2016). Adaptive learning. *ELT Journal, 70*(1), 88-93. https://doi.org/10.1093/elt/ccv055

Klaverena, C., Vonk, S., & Cornelisz, I. (2017). The effect of adaptive versus static practicing on student learning-evidence from a randomized field experiment. *Economics of Education Review, 58*, 175-187. https://doi.org/10.1016/j.econedurev.2017.04.003

LaVergne, V. (2007). *The effect of the ALEKS: Web-based learning system on standardized math scores* [Unpublished research project]. Shawnee Mission South.

Mood, A. M., Graybill, F. A., & Boes, D. C. (1974). *Introduction to the theory of statistics* (2nd ed.). McGraw-Hill.

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review, 109*(4), 1426-1460. https://doi.org/10.1257/aer.20171112

Nwaogu, E. (2012). *The effect of ALEKS on students' mathematics achievement in an online learning environment and the cognitive complexity of the initial and final assessments* [Unpublished doctoral dissertation]. Georgia State University.

Oxman, S., & Wong, W. (2014). *White Paper: Adaptive learning systems*. Integrated Education Solutions. https://kenanaonline.com/files/0100/100321/DVx_Adaptive_Learning_White_Paper.pdf

Padilla-Oviedo, A., Mundy, M., & Kupczynski, L. (2016). The effects of instructional strategies, college division, and gender on students' performance in college algebra at a university in South Texas. *Journal of Advance Research in Mathematics and Mathematical Sciences, 1*(8), 22-34. https://nnpub.org/index.php/MS/article/view/550/499

Privitera, G. J., & Ahlgrim-Delzell, L. (2019). *Research methods for education*. SAGE Publications.

Sabo, K. E., Atkinson, R. K., Barrus, A. L., Joseph, S. S., & Perez, R. S. (2013). Searching for the two sigma advantage: Evaluating algebra intelligent tutors. *Computers in Human Behavior, 29*(4), 1833-1840. https://doi.org/10.1016/j.chb.2013.03.001

Taylor, J. M. (2006). *The effects of a computerized-algebra program on mathematics achievement of college and university freshmen enrolled in a developmental mathematics course* [Unpublished doctoral dissertation]. Texas A&M University, College Station.

UNESCO Institute for Information Technologies in Education. (2012). *Personalized learning: A new ICT-enabled education approach* [Policy brief]. https://iite.unesco.org/pics/publications/en/files/3214716.pdf

US Department of Education Office of Educational Technology. (2010). *Transforming American education: Learning powered by technology*. US Department of Education. https://www.ed.gov/sites/default/files/netp2010.pdf

US Department of Education Office of Educational Technology. (2013). *Expanding evidence approaches for learning in a digital world*. US Department of Education. https://tech.ed.gov/wp-content/uploads/2014/11/Expanding-Evidence.pdf

Walkington, C. A. (2013). Using adaptive learning technologies to personalize instruction to student interest: The impact of relevant contexts on performance and learning outcomes. *Journal of Educational Psychology, 105*(4), 932-945. https://doi.org/10.1037/a0031882

Xu, Y., Meyer, K. A., & Morgan, D. D. (2009). A mixed-methods assessment of using an online commercial tutoring system to teach introductory statistics. *Journal of Statistics Education, 17*(2), 1-17. https://doi.org/10.1080/10691898.2009.11889524

Yilmaz, B. (2017). *Effects of adaptive learning technologies on math achievements: A quantitative study of ALEKS math software* [Unpublished doctoral dissertation]. University of Missouri-Kansas City.

Footnotes

1. Detailed explanation of one-group pretest-posttest experimental design or single-case experimental design can be found in Privitera and Ahlgrim-Delzell (2019).
2. See Franklin City Schools (2016), for example, for explanation of RIT.
3. Refer to ALEKS Corporation (2016a, 2016b, 2017) for various features of ALEKS, including its scoring system.