https://doi.org/10.23095/ETI.2024.25.1.067
Educational Technology International
2024, Vol. 25, No 1, 67-92

A Study on the Effectiveness of LMS for Improving College Student's Mathematics Performance using a Propensity Score Matching Method

Heejoo PARK
Yonsei University
Korea

Sunyoung BU
Hongik University, Sejong
Korea

Jihoon RYOO*
Yonsei University
Korea

This study aims to verify the practical effectiveness of learning management system (LMS) by introducing a LMS enhancing digital assessment utilizing automatic item generation in order to strengthen college student's mathematics performance. Teaching assisted with digital assessment in the LMS was applied to college mathematics classes, and the research question is whether or not students in the classes utilizing the LMS perform better than the regular classes. In particular, a calculus course, which is the foundation of important artificial intelligence technology in the future, was utilized in this study. The participants of this study were 248 freshmen in science and engineering who were taking calculus courses at a small to mid-size university. A total of 156 freshmen were selected after applying a propensity score matching method (PSMM), 78 from classes utilizing the LMS and 78 from regular classes without the LMS assisted with the digital assessment. As a result, it was found that there was a statistically significant difference in the math academic growth of students who used the LMS and those who did not. In other words, when LMS was used in calculus, students' academic growth was greater. The results of this study are meaningful in that they observed students' academic growth and confirmed that LMS enables a positive role in students' academic growth. In addition, if digital assessment is strengthened and LMS that enables individualized learning analysis is introduced and implemented in educational institutions, it is expected to play a major role in strengthening students' academic performance.

Keywords : Automatic Item Generation, College Student's Mathematics Performance, Digital Assessment, Learning Management System, Propensity Score Matching Method

---

* Dept. of Education, Yonsei University, Seoul, ryoox001@yonsei.ac.kr

Introduction

As generative AI (artificial intelligence) such as ChatGPT has emerged and developed rapidly, the generative AI is expected to take a lead in many activities in education and changes the overall industrial structure in EduTech. In particular, the generative AI actively generates results according to the user's needs, enabling it to learn large amounts of data (hyper-scale data) and even reach the realm of creativity (Yang & Yoon, 2023). There are various research activities related to the generative AI such as research on the analysis of learner types (Lee & Heo, 2023), research on the use of ChatGPT in an education field (Cha & Im, 2023), and research on writing education using intelligence (Chang, 2023), research on the direction and policy analysis of AI education in domestic and foreign countries (Kim, 2021), a study on ways to improve school mathematics curriculum for AI and software education (Park & Hong, 2022). Also, a number of studies have been conducted on education in the era of artificial intelligence, and a study on the use of AI in school education (Hong, Kim, & Park, 2021).

In sum, if generative AI such as ChatGPT is used appropriately for the purpose in the education curriculum, it can help learners with their learning tools and learning methods and can play a role as a learning assistant by providing learning materials or ideas. This means that the scope of use in various subjects can gradually expand in that feedback can be received in real time regardless of time and place. By checking more materials and information about the curriculum materials, instructors can also provide learners with a wider range of better resources selected by students' level and utilize them in actual classes. In addition to acting as a learning aid for students, instructors can also prepare their teaching with the aid of the generative AI. That is, it can even serve as an assistant to instructors for designing classes, preparing daily teaching materials, and providing feedback that cannot be given immediately to students. In other words, if the positive functions of artificial intelligence are appropriately utilized in individual subjects, and students also use it for assignments and studies according to their individual interests and levels, the advantages of AI integrated into the curriculum by grade, level, and subject can be maximized.

Hong et al. (2021) classifies the use of AI in education into two types: in the one is a form that performs teaching and learning functions in an independent form without an instructor, and the other form is that it can exist as a tool that plays an auxiliary role to help the instructor. Utilizing AI and its assistive tools, more instructors are designing their courses and supporting individual students' self-directed learning. In addition, in the future, AI will become more common throughout our society, learning through AI and obtaining necessary information, and the ability to use AI is expected to emerge as an important competitive advantage (Kim, 2023; Yoon & Yang, 2021).

In this era of the 4th Industrial Revolution, where many areas of society, including education, are transitioning from an offline environment to a digital environment, AI and big data technologies are recognized as key elements of national competitiveness. Additionally, since 2019, the government has been supporting the opening of AI graduate programs in universities, emphasizing AI-related research needed for the future society, such as computer vision, natural language processing, and big data analysis. In order to take a lead on this development in AI-era, the importance of mathematical literacy courses in higher education such as calculus, statistics, and linear algebra should be being emphasized, and those courses are needed to strength their backgrounds for their majors as a general education (Ryoo & Kwon, 2021).

Based on the study on the connection between high school mathematics and university mathematics curricula (Sin, 2006), the problem of declining academic ability in basic university mathematics leads to difficulties in later in-depth major studies for students in engineering or science departments. There are many studies to diagnose the need for liberal arts mathematics, improve the current basic mathematics curriculum and operation, and enable students to study their major curriculum systematically (Kim, 2017; Lee et al., 2011; Park & Pyo, 2013; Pyo & Park, 2011).

In addition, many parts of our society have been converted to an online environment, such as changes in the educational environment brought about due to COVID-19, the rapid expansion of AI, and technological advancements. These changes in the educational field have caused various confusion by school and region, but in this situation, Learning Management System (LMS) plays the most core and leading role in online education that began after the COVID-19 pandemic was performed (Kim & Ryu, 2019; Park et al., 2022). In addition, when not only the educational environment changes due to COVID-19, but also the online education paradigm along with the development of AI is also changing significantly, the LMS must play a role in helping learners' academic growth in terms of educational equity (Edsurge, 2016).

Therefore, in line with the trend of the great transition to the digital age and at a time when strengthening college mathematics competency is becoming important as mentioned above. This study aims to verify the practical effectiveness of LMS by introducing a LMS, CLASS-Analytics (Sa et al., 2021) enhancing digital assessment utilizing automatic item generation (Gierl & Haladyna, 2012), to strengthen college student's mathematics performance. On the LMS, students checked the course, performed assignments, solved problems, and received real-time feedback. In addition, receiving feedback does not mean the end of their assignments but the AI tutor provides solutions to the items that students answered incorrectly, and also provides additional equivalent problems to those items. The research question is whether or not students in the classes utilizing the LMS perform better than the regular classes. To examine the effectiveness of the utilization of CLASS-Analytics, the Propensity Score Matching Method (PSMM; Rosenbaum & Rubin, 1983; Rubin, 2007) was applied. The main reason of using PSMM is to draw an approximation of causality on group effectiveness between the treatment group and the control group.

Theoretical Framework

In this study, we sought to verify the effectiveness of the LMS by confirming the difference in academic growth between the group that used the LMS and the group that did not. Specifically, the purpose is to determine whether students ultimately have the learning capabilities for later major subjects through the basic college mathematics courses such as calculus, provided to 1st and 2nd undergraduate students.

College Mathematics

There is a large gap in the knowledge of college mathematics among 1st and 2nd undergraduate students in science and engineering majors and engineering colleges, and the absence of basic liberal arts math education leads to difficulties in studying majors. Although the high school curriculum is being changed as advanced mathematics courses are no longer required in Korea, there are no changes in the college math curriculum. It was said that the problem of lack of basic academic ability in college mathematics and the large differences by school and region, especially majoring in the STEM area, which can directly lead to the problem of educational equity in college basic math classes, are acting as obstacles to nurturing talent for the 4th industry (Ryoo & Kwon, 2021). Therefore, since there is a significant difference in the basic academic ability of college 1st and 2nd undergraduate students in mathematics, promotion to major studies makes it difficult to achieve a systematic curriculum (Kim, 2017; Pyo & Park, 2010). In the same context, Lee et al. (2022) pointed out that differences in students' levels or capabilities are not considered when operating the basic mathematics curriculum in universities. In other words, it was said that basic math skills should be measured in advance, and the content or programs to strengthen math skills should be provided to students who need it.

Before talking about the large gap in basic mathematics competency, mathematics is classified as the subject with the highest level of basic academic deficiencies among middle school and high school students even before college mathematics. Looking at the numbers specifically, the country conducts a national-level academic achievement evaluation every year to determine the percentage of achievement levels and the percentage of students lacking basic academic skills for each subject in Korea. As of 2021, the percentage of students falling short of basic academic skills in math subjects is 30% for third-year middle school students. It can be seen that, in Figure 1, 11.6% for Mathematics is twice as high as the 6.0% for Korean language subjects and 5.9% for English subjects. Similarly, while the number of second-year high school students was as high as 14.2%, the Korean language subject was 7.1% and the English subject was 9.8%, showing that the math subject showed a figure of more than twice as high as lacking basic academic ability (Korean Ministry of Education, 2022; see <Table 1> and [Figure 1]). Due to the highly hierarchical nature of mathematics as a subject, it is natural that students who do not meet the level of the basic academic ability in mathematics in middle and high school will face difficulties in college mathematics. In particular, in order to cultivate talented people who will lead the AI era, the importance of mathematics subject increases. This also suggests that education that can strengthen mathematical capabilities should be a priority task. From this perspective, we would like to conduct a study using CLASS-Analytics, an LMS that embeds digital assessment in class, as a way to maximize college mathematics performance.

Table 1.
Proportion of the level of below basic academic ability by subject (%)

Subject / Year	3rd grade middle school			2nd year of high school		
	Korean	Mathematics	English	Korean	Mathematics	English
2019	4.1	11.8	3.3	4.0	9.0	3.6
2020	6.4	13.4	7.1	6.8	13.5	8.6
2021	6.0	11.6	5.9	7.1	14.2	9.8

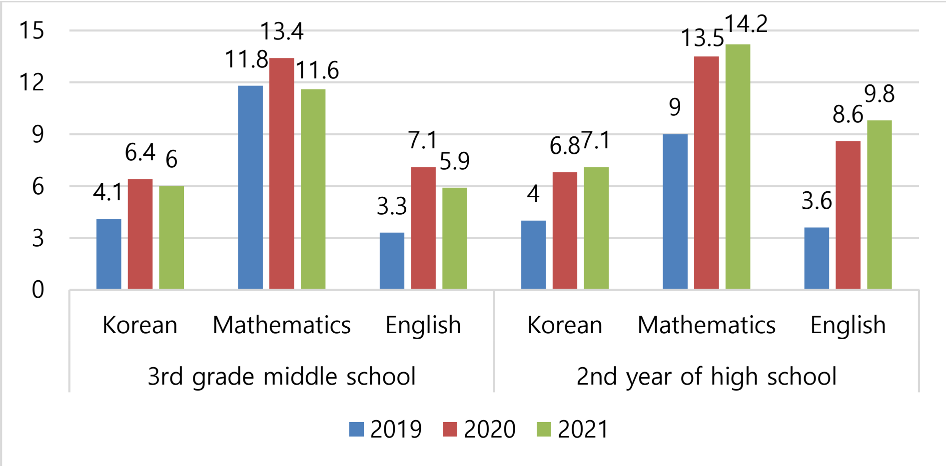

Figure 1. National level academic achievement below basic academic ability

Learning Management System (LMS)

Higher education institutions such as Minerva University have been conducting fully online education in the most of curriculum even before COVID-19, and have implemented LMS into online education accordingly. Online education has rapidly accelerated in Korea since COVID-19, and the development of AI has also served as an opportunity to expand the importance of online education. As a result, most educational systems are preparing for education in the post-COVID era, and various changes are needed to implement online education more efficiently and effectively. At that time, LMS became a key element in online education, and it can be seen through previous research that through LMS, learners are performing a role beyond directly participating in online classes. Kim and Cho (2018) showed that LMS plays the role of providing continuous feedback to learners individually throughout the learning process. Ultimately, it was said that students can make up for their shortcomings and improve their learning performance, and that instructors can check the learning path of these students. Kim (2021) also conducted an evaluation study showing that learners' learning performance can be quickly evaluated, instructors can provide feedback on the results, and teaching and learning methods can be improved through this process. In the similar context, Park and Gil (2020) reported that the ultimate goal of online classes conducted through LMS should be conducted entirely in the form of learner-centered classes equivalent to offline classes. Furthermore, Edsurge (2016) suggested that individual profiles should be provided to students through the LMS, and that this ultimately plays a role in predicting learners' individual learning models, which is an element that should be included in the LMS.

In addition to changing the delivery format in classes, online assessment within the LMS has been emphasized (Park & Hong, 2022). Eventually, through online education via LMS, it is necessary to play a role as a tool to not only achieve various academic achievements in higher education, but also to resolve inequities in educational opportunities that may exist in local universities or underprivileged classes. Such an effort in educational equity would resolve various educational problems and educational gaps. In summary, LMS goes beyond its role as a mediator between learners and instructors in online education, designing and managing individual learner-centered curriculum, recording learners' academic progress in data, and managing academic achievements. It can be confirmed that it is performing its efficient, effective, and core role.

Methods

In this study, a survey was conducted on 248 1st and 2nd undergraduate students in the science and engineering field who were taking calculus courses at a small to mid-size university in Korea, and the 156 sample subjects examined was selected through the propensity score matching method. Accordingly, we would like to verify whether the use of an AI-based LMS with digital assessment, CLASS-Analytics, can have any effect on students' academic growth in mathematics.

Data

Sample data consist of students enrolling in 10 calculus courses at a midsize university of about 7,000 students in Korea. A total of 248 students responded to the survey, and the matched sample through the propensity score matching method consisted of 156 students. The subjects of the survey were 79 students from three classes, which used LMS, and 169 students from seven classes, which did not use LMS. To ensure objectivity in selecting the control and treatment groups, individual groups were selected from the same college. In particular, all classes were selected from groups that took the same test at the same time, ensuring that there were no defects in sample selection. The university provides basic college mathematics courses in common to science and engineering departments for 1st and 2nd undergraduate students, and the case of the university, which has introduced and implemented an AI-based LMS, was set as the subject of this study.

CLASS-Analytics is an LMS equipped with AI functions such as AI-based teaching assistant and AI-based tutor, and is currently being used in classes at higher education institutions. The AI-based tutor helps students solve mathematical terms or concepts that they have difficulty solving, and check and record the student's learning status by providing additional problems of a similar type for incorrect problems. Additionally, it provides information based on ChatGPT on the fly. CLASS-Analytics is also equipped with functions such as loading lecture materials, real-time online assignments, providing learning analysis profiles for each student, and providing on-the-fly feedback, so that students can preview of classes with concept questions and review classes sufficient calculus questions generated by AIG function in the subject and check the status of individual student's learning. This study examines its effectiveness by distinguishing between three classes that use CLASS-Analytics and seven classes that do not.

In order to collect factors that influence academic growth, students were surveyed with 31 questions consisting of 13 educational performance scale questions, 11 educational achievement questions, and 7 educational satisfaction questions. The specific survey questions used EPI (Educational Performance Indicators), a scale used in the study by Kang et al (2010), to measure educational performance, including Cognitive, Affective, and Sociocultural. Among the educational performance scales consisting of improvement and sociability), this study used 13 questions about information of composition ability in cognition, which is an indicator related to students' academic growth. EPI is a scale developed by identifying the cognitive, emotional, and sociocultural domains as essential factors in the learning process in which university students' learning capabilities and subjects are developed with a focus on competency. In particular, it is significant in that learning competency does not stop at measuring individual learning achievements, but rather improves the overall quality of university education based on this. In addition, EPI is a scale created by verifying the validity of measurement tools in previous studies that studied many existing educational performance scales and reconstructing and supplementing sub-factors to suit the capabilities of Korean college students (Kang et al, 2010).

Next, along with the rapid development of technology, various technological and environmental changes, including AI technology, have occurred in the field of education, leading to the emergence of 21st century core competencies such as problem-solving competency, creativity, critical thinking ability, and cooperation competency in universities. Accordingly, there has been a change in the form of classes in education, such as PBL (Problem-Based Learning) classes that aim to develop problem-solving capabilities, and accordingly, many studies have been conducted on which type of class shows the highest learning effectiveness. From this perspective, the teaching effect measurement tool used was the scale used in the study by Lee et al. (2023). Among the teaching effect measurement tools consisting of instructor class design and operation, learner problem-solving activities, educational achievement, and educational satisfaction, this study used LMS including those tools. To measure the teaching effect that can be achieved by using in class, the educational achievement and educational satisfaction questions were also used. In addition to the information of cognitive, educational achievement, and educational satisfaction constructed in this way, students' major and gender were additionally investigated to identify factors affecting academic growth, and five items were used as covariates in the analysis.

Propensity Score Matching Method (PSMM)

In this study, we intend to verify the differences between the treatment group and the control group through the Propensity Score Matching Method (PSMM; Rosenbaum & Rubin, 1983; Rubin, 2007). The propensity score matching method was proposed by Rosenbaum and Rubin (1983) and is a quasi-experimental research design that can be used when conducting a randomized experiment is not possible. This is a method of analyzing the effectiveness of a program or experiment by forming a comparison group that is most suitable for the treatment group (Guo & Fraser, 2015).

The propensity score matching method first sets the dependent variable in a binary form with the treatment group as 1 and the control group as 0, and then performs a logistic regression analysis that sets the covariates as independent variables. In this study, the propensity score of the treatment group is estimated, and samples having similar the propensity score are extracted from the control group, to match the two groups. Data from the unmatched treatment group and control group are not used in the analysis. By forming a comparison group with the most similar characteristics to the treatment group, it is possible to prevent selection bias that distorts statistical analysis when the sample is incorrectly selected. Selection bias refers to an increase in the probability of a research subject receiving a specific treatment due to a specific covariate. In this case, the covariate acts as a confounding variable in identifying the difference in results between the treatment group and the control group. The propensity score matching method has the advantage of reducing this bias.

First, the PSMM is premised on two assumptions (Rosenbaum & Rubin, 1983;

Rubin, 2007). First, it is based on the conditional independence assumption, which assumes that participation in the experiment and the final dependent variable are independent. Therefore, the difference in performance that appears in the dependent variable is controllable by the observed predictor variable, and characteristics other than the predicted variable are unrelated to performance. As a result, unobserved characteristics do not affect performance, and observed predictor variables determine the performance of the dependent variable. In this way, the assumption of conditional independence can be summarized as Eq. (1). In other words, the assumption is that whether a treatment group z is assigned to a covariate x is independent of the dependent variable r_i.

$$(r_1, r_0) \perp z|x \tag{1}$$

Second, a common support assumption is required, which assumes that whether or not subjects participating in the analysis participated in the experiment is in a common area. In other words, it is assumed that the experiment participation probability of the control group and the treatment group is within a common area, which can be summarized as Eq. (2). In other words, the assumption is that the probability assigned to the treatment group z of the target to be analyzed is within a common area.

$$0 < \Pr(z = 1|x) < 1 \tag{2}$$

Therefore, through these two assumptions, the performance of the dependent variable can be estimated without selection bias of the students participating in the experiment. In addition, the most common propensity score matching method is nearest-neighbor matching, which involves pairing a participating group with a non-participating group. The difference in performance is calculated from the sample of each group, and the average value of the calculated difference is used to calculate the effect (Rosenbaum & Rubin, 1983). Therefore, the propensity score of the study subjects $i(i = 1, \ldots, N)$ corresponding to the control group and the treatment group is calculated as the conditional probability of the treatment group $(z = 1)$ relative to the control group given the observed covariate x_i, and can be summarized as Eq. (3), (4).

$$e(x) = pr(z = 1|x) \tag{3}$$

$$pr(z_1, \dots, z_n|x_1, \dots, x_n) = \prod_{i=1}^{N} e(x_i)^{z_i}\{1 - e(x_i)\}^{1-z_i} \tag{4}$$

The propensity score calculated in this way is obtained, and the average difference between the propensity scores of the control group and the treatment group can be calculated as the average treatment effect. Specifically, in this study, when using the propensity score matching method, if the class that used CLASS-Analytics in class and the class that did not have the same propensity score, the covariate was calculated to come from the same distribution. This process reduces potential bias that may arise from the unbalanced distribution of covariates between the treatment and control groups and extracts unbiased samples when analyzing the effectiveness of the experiment (Mitra & Reiter, 2012).

Analysis was conducted using the *MatchIt* package based on R (Ho, Imai, King & Stuart, 2007; 2011), utilizing the one-to-one nearest-neighbor matching function that matches the optimal value of the control group in the treatment group. This is because the most appropriate and valid sample is extracted when sampling for propensity score matching (Gu & Rosenbaum, 1993; Guo & Fraser, 2015; Rosenbaum & Rubin, 1983; Ho et al., 2007; 2011). Therefore, in this analysis, the same sample size was extracted between the group that used CLASS-Analytics in class and the group that did not, and the resulting difference in academic growth in mathematics was confirmed. The variables used were students' major, gender, information of cognitive, educational achievement, and educational satisfaction. Academic growth was calculated using test scores commonly administered to all classes. The T value score of the standardized midterm exam was subtracted from the T value score of the standardized final exam, and the final difference between the standardized scores was used for analysis.

Results

In this study, before verifying the effectiveness of students' academic growth in mathematics confirmatory factor analysis (CFA) was conducted on the analysis tool to confirm three factors measured reliably and validly (<Table 2>).

Table 2.
Descriptive statistics of three factors

Factor	Item	Mean	SD	Kurtosis	Skewness	VIF
Cognitive	Q1	3.750	0.857	-0.441	-0.075	1.811
	Q2	3.712	0.923	-0.662	0.193	1.618
	Q3	3.897	0.878	-0.924	1.088	1.693
	Q4	3.995	0.750	-0.698	1.118	2.071
	Q5	3.429	0.927	-0.166	-0.352	2.923
	Q6	3.359	1.014	-0.165	-0.489	2.709
	Q7	3.935	0.751	-0.518	0.695	2.547
	Q8	3.810	0.912	-0.573	0.028	1.630
	Q9	3.304	0.978	-0.076	-0.449	1.823
	Q10	3.332	0.960	-0.072	-0.296	2.094
	Q11	3.043	0.963	0.135	-0.236	2.188
	Q12	3.342	0.879	0.052	-0.286	2.341
	Q13	3.543	0.848	-0.219	-0.025	1.882
Educational Achievement	Q14	3.408	1.031	-0.400	-0.232	2.683
	Q15	3.875	0.776	-0.560	0.614	2.280
	Q16	3.457	0.996	-.0314	-0.278	3.020
	Q17	3.837	0.820	-0.470	0.164	2.404
	Q18	3.190	1.184	-0.175	-0.750	6.118
	Q19	3.098	1.169	-0.088	-0.744	5.714
	Q20	3.690	0.873	-0.649	0.479	3.232
	Q21	3.212	1.047	-0.060	-0.417	3.381
	Q22	3.495	1.003	-0.495	0.137	3.826
	Q23	3.484	0.923	-0.183	-.0278	3.040
	Q24	3.255	1.089	-0.241	-0.364	3.163
Educational Satisfaction	Q25	3.435	1.129	-0.573	-0.280	3.075
	Q26	3.875	0.875	-0.792	1.089	3.416
	Q27	3.565	0.984	-0.393	-0.192	2.306
	Q28	3.766	0.944	-0.700	0.679	3.222
	Q29	3.696	1.048	-0.631	0.176	4.342
	Q30	3.826	0.942	-0.636	0.472	4.254
	Q31	4.022	0.868	-0.752	0.543	3.670

Table 3.
Factor loadings for three factors

Factor	Item	Cognitive	Achievement	Satisfaction
Cognitive	Q1	1.000		
	Q2	0.861		
	Q3	0.728		
	Q4	0.949		
	Q5	1.421		
	Q6	1.490		
	Q7	1.021		
	Q8	0.885		
	Q9	1.203		
	Q10	1.102		
	Q11	1.136		
	Q12	1.109		
	Q13	1.038		
	Reliability	0.854		
Educational Achievement	Q14		1.000	
	Q15		0.630	
	Q16		0.971	
	Q17		0.697	
	Q18		1.089	
	Q19		1.064	
	Q20		0.932	
	Q21		1.136	
	Q22		1.124	
	Q23		0.965	
	Q24		1.076	
	Reliability		0.913	
Educational Satisfaction	Q25			1.000
	Q26			0.866
	Q27			0.892
	Q28			0.982
	Q29			1.020
	Q30			0.891
	Q31			0.799
	Reliability			0.895

In addition, individual loading values of the three factors were calculated through CFA, and the results are summarized in <Table 3>. Since all loading values were above 0.4, it was confirmed that convergent validity was satisfied. Next, as a result of examining the reliability as an internal validity, the value was 0.854 for cognitive, 0.913 for educational achievement, and 0.895 for educational satisfaction. It was confirmed that all three individual scales showed high reliability.

As for the goodness-of-fit index of the CFA model, the χ^2-value was 793.994 (df=478), the RMSEA value was 0.058, the CFI value was 0.896, and the SRMR value was 0.069, respectively. These are summarized in <Table 4>. Accordingly, the model fit between the established CFA model and the data satisfies RMSEA = 0.058 (required < 0.06), and although it falls short of CFI = 0.896 (required > 0.90), the CFI value is close to the goodness-of-fit index and satisfies SRMR = 0.069 (required < 0.08) (Hu & Bentler, 1999), the suitability of the established research model was confirmed to be appropriate.

Table 4.
Model evaluation of the CFA model

χ^2	df	RMSEA (90% CI)	CFI	SRMR
793.994	478	0.058 (0.051-0.065)	0.896	0.069

Next, as a result of selecting samples of the treatment group and control group using the one-to-one nearest-neighbor matching function within the propensity score matching methods, the descriptive statistics according to the data distribution are shown in <Table 1>. As can be seen in <Table 5>, the distribution between the treatment group and the control group in the data after propensity score matching compared to the original data was selected by reviewing the similarity of the mean and variance in the matching variables, cognitive, educational achievement, and educational satisfaction. Through this, a total of 156 people were selected, including 78 in the treatment group and 78 in the control group, who were most suitable for this analysis.

Table 5.
Distribution of raw data and propensity score matching data

	Raw data		Propensity score matching data	
	Treatment group	Control group	Treatment group	Control group
Academic growth	50.82	50.26	50.75	48.91
Midterm	50.36	47.69	50.52	51.19
Final exam	50.99	48.13	51.08	50.19
Cognitive	3.59	3.49	3.57	3.61
Educational Achievement	3.60	3.27	3.59	3.44
Educational Satisfaction	3.80	3.63	3.79	3.66
Number of students	79	159	78	78

Next, the distribution of the samples selected as the treatment group and control group using the propensity score matching method is shown in [Figure 2]. Matched Treatment Units and Matched Control Units are data selected through the propensity score matching method, and it can be seen that the treatment group and control

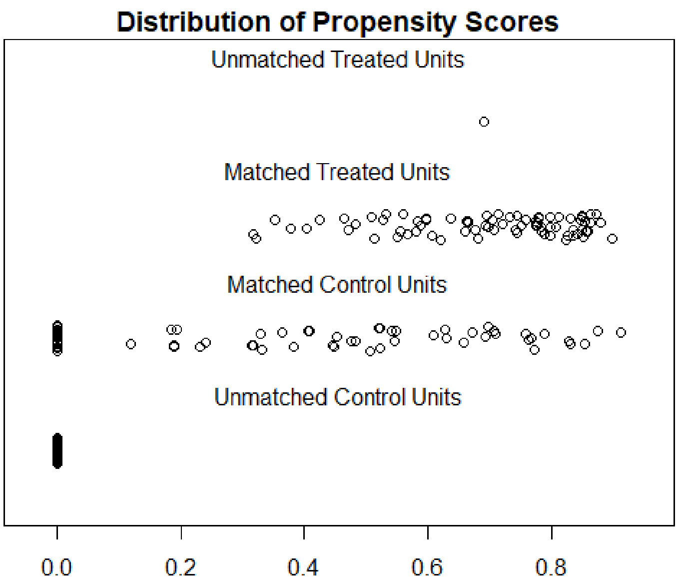

Figure 2. Sample selection by group through propensity score distribution

group show similar distributions. In the treatment group, the sample selected through the propensity score matching method consisted of 78 people, excluding 1, from the total of 79 people in the treatment group, and similarly, the data of 78 out of 169 people in the control group were selected as the sample.

As a result of the study, analysis of variance was conducted to examine changes in academic growth in students when LMS was used in college mathematics courses. Before performing analysis of variance, the independence of the samples was confirmed, normality was verified, and finally, homoscedasticity was confirmed to be satisfied. As a result of conducting t-test after confirming the assumptions of t-test, it was found that there was a statistically significant difference in academic growth. The results are summarized in <Table 6>.

Table 6.
Results from t-test over both treatment and control groups

Dependent variable	Group	N	Mean	SD	F	Pr(>F)
Academic growth	Treatment group	78	50.359	9.009	4.459*	0.036
	Control group	78	47.693	11.568		

As a result of checking the effect on students' academic growth when CLASS-Analytics, an LMS enhancing digital assessment via automatic item generation was used in the calculus subject to strengthen college student’s mathematics performance, which was investigated in this study, it was confirmed that there was a statistically significant difference. In other words, when CLASS-Analytics was used in the calculus subject to strengthen college student’s mathematics performance, it was found to be more effective in students' academic growth.

Discussion and Conclusion

This study aims to verify the effectiveness of students' academic achievement capabilities by introducing CLASS-Analytics to higher education institutions and implementing it in college mathematics courses. In particular, the CLASS-Analytics with enhanced digital assessment was used in actual college courses as a medium for learning, including loading lecture materials, learning tools, implementation of assignments, analysis of each student's profile, and provision of feedback. We confirmed the resulting differences in academic growth and ultimately attempted to verify the effectiveness of CLASS-Analytics as a tool for strengthening college student's mathematics performance.

The results of this study are similar to those of Kim (2022), who found that learning through LMS is effective for students. In particular, it was confirmed that sufficient and necessary information could be selectively obtained through LMS. Woo et al. (2012) found that learning participation activities through LMS were related to grades, and specifically confirmed that LMS learning participation activities predicted more than 32% of students' grades. In addition, the use of LMS plays a positive role in the interaction between instructors and learners beyond the teaching and learning mechanism in that the class content, output from class activities, and other students' output can be clearly utilized as learning materials. In the same context, the use of LMS records and monitors learners' activities by supporting instructional design. It also enables a systematic role to sufficiently discuss cognitive processes between instructors and learners, which ultimately makes it effective for learners. Thees results were similar to Choi's (2019) findings. A similar conclusion was drawn to the research results of Lee et al. (2023), who found that it promotes learning outcomes.

In addition, with regard to general education in college mathematics, which this study sought to identify as a core subject, colleges provide basic mathematics courses such as calculus, statistics, and linear algebra to students in order to improve basic university mathematics skills. Likewise, many previous studies have shown that it is necessary to diagnose basic mathematics academic ability for freshmen, and that it is necessary at the university level to provide college mathematics courses to improve basic mathematics competency (Lee et al., 2010; Pyo & Park, 2011). In the study by Lee et al. (2022), they also point out that there is a large gap in the basic mathematics academic ability of college entrance students. To this end, it is necessary to measure students' abilities and strengthen students' mathematics abilities to better understand college students. By providing a curriculum for the students, there is a need to guarantee students' right to future learning, and there is a great need to review and change the curriculum itself for students' successful college education in various major fields.

In this study, the one-to-one nearest-neighbor matching function in the PSMM was used to match the treatment group's optimal value with the control group's value, and a quasi-experimental design method was selected to extract samples from the two most comparable groups. Therefore, the significance of the study is that it selected students with the most similar conditions without selection bias and verified the effectiveness of academic growth in mathematics. The results of this study also confirmed that the use of LMS has a positive effect on strengthening students' college mathematics performance. Thus, the use of CLASS-Analytics would be recommended.

In addition, we would like to conduct follow-up research to verify the effectiveness of AI tutor functions and feedback. Also, there is a need for discussion about which method and time to provide feedback can be most effective in mathematics studies. In particular, there should be a great need to assist students' basic mathematics academic ability by enabling preview and review of basic mathematics, allowing large discrepancies among students' academic ability, and by enabling sufficient individualized learning through a LMS equipped with digital assessment. In addition, this study is meaningful in that it observed students' academic growth rather than their academic achievement, confirming that the use of LMS in one-semester classes can play a positive role in students' academic growth.

Also, this study utilized only one of educational performance index scale, Cognitive, in the PSMM due to the time constraint in the class setting. As a future research, we would like to explore more covariates that may be associated with student's learning.

Finally, through the results of this study, we studied the feasibility of LMS that enables individualized learning analysis, and it is expected that these research results will be introduced and stably implemented in educational institutions and play a significant role in improving students' academic capabilities. Lastly, there should be a need to investigate the impact of LMS use on students' academic growth in mathematics courses longitudinally by collecting annual data rather than just one semester. This also needs to be seen in larger sample data for 1st and 2nd college students entering science and engineering fields. I suggest conducting the same study.

References

Cha, M., & Im, H. (2023). A Study on University Professors' Perception on Educational Applicability of ChatGPT in English Classes. *Culture and Convergence, 45*(5), 109-118.

Chang, S. (2023). ChatGPT has Changed the Future of Writing Education - Focusing on the response of writing education in the era of artificial intelligence -. *Research on Writing, 56*, 7-34.

Cho, S., Pyo, Y., Park, J., & Seo, J. (2017). Error types and guidance measures for college students with poor basic academic ability in mathematics. *Studies in Mathematical Education, 1*, 91-93.

Choi, M. (2019). "A Study on the Interactivity of Smart LMS in a University Class", *Journal of Digital Convergence, 17*(3), 395-404.

Edsurge. (2016). *Decording adaptive*. London: Pearson. Retrieved July 12, 2018, http://d3e7x39d4i7wbe.cloudfront.net/static_assets/PearsonDecodingAdaptiveWeb.pdf.

Gierl, M. J., & Haladyna, T. M. (Eds.). (2012). *Automatic item generation: Theory and practice*. Routledge.

Gu, X., & Rosenbaum, P. R. (1993). Comparison of multivariate matching methods: Structures, distances, and algorithms. *Journal of Computational and Graphical Statistics, 2*, 405–420.

Guo, S., & Fraser, M. W. (2015). *Propensity score analysis: Statistical methods and applications* (2nd ed.). ThousandOaks, CA: Sage.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). "MatchIt: Nonparametric preprocessing for parametric causal inference,". *Journal of Statistical Software, 15*, 199-236.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software, 42*(8). 1-28.

Hong, S., Kim, H., & Park, Y. (2021). Exploring the Potentials of AI Integration into K-12 Education. *Journal of Korean Association for Educational Information and Media, 27*(3), 875-898.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional fit criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Kang, M., Heo, H., Jo, I., Shin, J., & Seo, J. (2010). Developing an educational performance indicator for new millennium learners. *Journal of Research on Technology in Education, 43*(2), 157-170.

Kim, A. (2021). Analysis of Learner Behavior and Learning Performance using LMS Big Data in the COVID-19: Focused on J-University. *The Journal of Learner-Centered Curriculum and Instruction, 21*(6), 565-579.

Kim, E. (2022). A Study on the Relationship Between Digital Literacy and Major Satisfaction of Nursing Students: Intermediary Role of LMS Learning Effects. *The Journal of the Convergence on Culture Technology, 8*(4), 27-35.

Kim, K. (2017). *A study on mathematics attitude and mathematics achievement of low achievement students in university mathematics*. Master's dissertation, Pukyung National University, Busan, Korea.

Kim, K., & Ryu, B. (2019). Expectations for and concerns about the use of Fourth Industrial Revolution technology in education. *The journal of Educational Studies, 50*(3), 55-79.

Kim, S. (2021). *The proposal of AI education direction reflecting cases in major countries and review of Korean textbooks*. Master's dissertation, Jeonbuk National University, Jeonbuk, Korea.

Kim, S., & Cho, Y. (2018). Proliferation of Online Learning and the Implications for Teaching and Learning in Future University Education. *Journal of Lifelong Learning Society, 14*(4), 51-78.

Kim, T. (2023). Can chatGPT Become a Tool for Innovation?: ChatGPT Use Cases and Prospects. *The AI Report, 1*, National Information Society Agency. Daegu.

Korean Ministry of Education. (2022). *2021 National Level Academic Achievement Assessment Results and Response Strategies Announced*. Sejong.

Lee, H., Park, H., & Yang, J. (2023). Development and Validation of Measuring Problem-Based Learning Effectiveness. *Journal of Educational Technology, 39*(1), 35-63.

Lee, J., Lee. S., Kwon, H., & Lee, G. (2010). A Study on the relation between Mathematical Scholastic Ability and Scholastic Aptitude Test. *Communications of Mathematical Education, 25*(4), 629-639.

Lee, J., Lee, S., Kwon, H., & Lee, G. (2011). The Effect of Basic Mathematical Ability Improvement Program on Scholastic Achievement and Scholastic Motivation: A Case Study on Engineering Freshmen in D University. *Communications of Mathematical Education, 25*(1), 167-184.

Lee, J., Park, H., Lee, J., & Park, M. (2023). A Comparative Analysis on LMS Features for Supporting Instructors. *Journal of Digital Contents Society, 24*(2), 399-410.

Lee, S., & Heo Y. (2023). Types of Learned helplessness Learners and Factors Influencing on Changes in the Era of Future Education based on ChatGPT. *The Journal of General Education, 20*, 87-119.

Lee, Y., Hwang, Y., Yun, J., Yoon, J., & Hwang, I. (2022). Development of Basic Mathematical Ability Test for University and Application Effect of Extra-curricular Program. *The Journal of General Education, 20*, 87-119.

Mitra, R., & Reiter, J. P. (2012). A comparison of two methods of estimating propensity scores after multiple imputation. *Statistical Methods in Medical Research, 25*(1), 188–204.

Park, H., Ryoo, H., Kwon, J., & Ryoo, J. (2022). Change of Paradigm on LMS for Online Education: LMS Implementing Learning Analytics and Online Assessment. *The Educational Research for Tomorrow, 35*(2), 49-72.

Park, J., & Pyo Y. (2013). Improvement strategies of teaching methods for university basic mathematics education courses by ability grouping. *Communications of Mathematical Education, 27*(1), 19-37.

Park, J., & Gil, J. (2020). Edutech in the Era of the 4th Industrial Revolution. *KIPS Transactions on Software and Data Engineering, 9*(11), 329-331.

Park, S., & Hong, H. (2022). A Study on the Improvement of School Mathematics Curriculum for SW · AI Education. *The Journal of Korean Association of Computer Education, 25*(2), 59-77.

Pyo, Y., & Park, J. (2010). Effective management strategies of basic mathematics for Low achievement students in university general mathematics. *Communications of Mathematical Education, 24*(3), 525-541.

Pyo, Y., & Park, J. (2011). Effective Management Strategies of University Basic Mathematics by Ability Grouping. *Journal of educational research in mathematics, 21*(1), 87-103.

Rubin, D. B. (2007). The design versus the analysis of observational studies for causal effects: Parallels with the design of randomized trial. *Statistics in Medicine, 26*, 20–36.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55.

Ryoo, J., & Kwon, J. (2021). *Equity of Educational Opportunity in Higher Education focusing on Fundamental Liberal Education: Cultivating of People on the 4th Industrial Revolution*. Korean Council for University Education. Seoul.

Sa, S., Ryoo, H. S., & Ryoo, J. H. (2022). Application of AIG implemented within CLASS Software for Generating Cognitive Test Item Models. *Educational Technology International, 23*(2), 157-181. https://doi.org/10.23095/ETI.2022.23.2.157

Sin, S. (2006). *A Study on the Connection between the 7th High School Mathematics Curriculum and University Mathematical Education*. Master's dissertation, Ulsan University, Ulsan, Korea.

Woo, J., Kim, B., & Bu, K. (2012). A Study on the Verification of Influential Power over Learning Effect by the Learning Participation Activities: the LMS of the O Cyber University. *The Journal of Korean Institute of Information Technology, 10*(4), 97-103.

Yang, J., & Yoon, S. (2023). Beyond ChatGPT to the Generative AI Era: Cases of Media · Content Generation Type AI Services and Ways to Securing Competitiveness. *Media Issues & Trends, 55*, 62-70.

Yoon, S., & Yang, J. (2021). *Fundamentals of AI and Data Analysis: Digital Business Survival Strategies*. Pakyoungsa.

Hee Joo PARK
Graduate Student, Dept. of Education, Yonsei University, Seoul.
Interests: Educational Statistics & Data Science
E-mail: heejoo514@ yonsei.ac.kr

Sunyoung BU
Associate Professor, Dept. of Liberal Arts, Hongik University, Sejong.
Interests: Mathematics
E-mail: syboo@hongik.ac.kr

Ji Hoon RYOO
Professor, Dept. of Education, Yonsei University, Seoul.
Interests: Educational Statistics & Data Science, Learning Analytics and Engineering, Learning Management System, Psychometrics
E-mail: ryoox001@yonsei.ac.kr

Received: March 10, 2024 / Peer review completed: April 6, 2024 / Accepted: April 18, 2024