An Artificial Intelligence-Enhanced Phenomenon-Based Learning Approach for Interdisciplinary Understanding and Speaking Skills

Surattana Adipat
Ramkhamhaeng University, Thailand, *surattana.a@rumail.ru.ac.th*

Teacher education programs are critical for creating a new generation of skilled English teachers in Thailand. The interdisciplinary learning process is an important measure for developing expertise in English, enhancing competency-based learning, and improving the quality of learning and teaching skills. This study aimed to 1) determine whether AI-enhanced phenomenon-based learning instruction facilitates the development of preservice teachers' interdisciplinary understanding and English-speaking skills, and 2) compare interdisciplinary understanding and speaking skills development between the experimental group and the control group. Participants were monitored through subjective and speaking tests conducted before and after learning instruction. Descriptive statistics (mean scores, standard deviations, and percentage) and inferential statistics through multivariate analysis of variance (normality test, Levene's test, Bartlett's test, and Box's test applied for assumption test and Wilks' lambda applied for hypothesis testing) were employed to analyze quantitative data collected in this study. Relative gain scores were also analyzed. The findings demonstrated that 1) AI-enhanced phenomenon-based learning instruction improved the interdisciplinary understanding and speaking skills of preservice teachers, and 2) the experimental group participants gained higher interdisciplinary understanding and speaking skills than the control group participants, as indicated by the increase in the performance mean scores across the pretest and posttest.

Keywords: interdisciplinary understanding, speaking skills, artificial intelligence, phenomenon-based learning, preservice teachers

INTRODUCTION

Since English is the universal language of the current age of globalization, English competence has become a critical skill. The significance of the language has also been acknowledged in Thailand, which has adopted English to communicate with people from different cultural, national, and linguistic backgrounds. Although Thai is the country's official language, English is considered its first international language Thai citizens should learn and practice English while traveling abroad as fluency in English has become necessary for meaningful interactions in various situations (Chinwonno, 2001). Since 2015, when Thailand became a member of the Association of Southeast

Asian Nations Economic Community, the English language has been prioritized. The communication abilities of preservice teachers have improved using new paradigms and standards (Loan, 2019), which were deployed in anticipation of the role that graduates would play as teachers in Thailand's development. The demand on schools and colleges to enhance students' communication abilities has increased in recent years, especially after all members of the ASEAN Economic Community decided to adopt English as the official language for international communication and business (Oeamoum & Sriwichai, 2020). Assisting students in improving their speaking abilities is the goal of teachers and administrative authorities. Good English communication skills are essential because it allows learners and people in general to share information, insights, opinions, and convictions with others while preparing them for the professional world. People estimate the language skills of others based on their ability to communicate verbally, whether this is fair or not (Oeamoum & Sriwichai, 2020). Speaking abilities are productive skills and pertain to "the capacity to express sound using particular language in oral communication," as defined by the National Institutes of Health (Chinwonno, 2001).

A significant contribution of Chinwonno's (2001) study is that it investigates why Thai preservice teachers avoid interacting with their colleagues in English both inside and outside the classroom and the factors that prevent them from honing their spoken English communication skills. One critical measure for developing skilled teachers was the introduction of the five-year preservice teaching program, which will help attain the ultimate goal of the English curriculum. According to Chinwonno (2001), this program is being implemented at Thai institutions to produce instructors who will be successful agents of development in the future. The establishment of the teacher education program endeavors to increase the English competence of Thai learners. The major difficulties in English teaching and learning for preservice teachers in Thailand relate to linguistic skills, thinking abilities, English teaching materials and course content, teaching methods, teacher's performance, and institutional support (Dincer & Yesilyurt, 2013). However, preservice teachers have expressed concerns about difficulties with curriculum, reading materials, and other elements that inhibit teaching success (Oeamoum & Sriwichai, 2020). Furthermore, they contend that their lack of tenacity in practicing or exploring additional avenues to communicate in English has contributed to their low confidence in using English for conversation and teaching.

A teacher qualification survey was conducted among 400 Thai English preservice teachers, which revealed that their qualifications were below the syllabus level at which they were teaching and that they had limited English proficiency (Oeamoum & Sriwichai, 2020). The top 40% of participants had a fair degree of fluency; only 20% were teaching students at a level for which they were qualified. In addition, according to Oeamoum and Sriwichai (2020), 65% of primary school English teachers did not major in English during their undergraduate studies. Dincer and Yesilyurt (2013) note that there are few English teachers at the secondary level. A lack of properly trained instructors hampers the tasks expected of teachers, which are currently acknowledged to be difficult and multi-faceted.

Apart from English-speaking skills, preservice teachers must have the ability to teach effectively in difficult circumstances, appropriately use information and communication technology infrastructure, and adjust to a wide range of teaching methods (Chinwonno, 2001). These measures are adopted to ensure that the quality of their teaching continues to improve, and they can successfully manage the enormous administrative workload, all of which is required to fulfill the standards for accreditation (Dincer & Yesilyurt, 2013). In the 21st century, educating preservice teachers requires a coherent plan based on current skills. Collaboration, communication, and critical thinking skills are the main aspects of interdisciplinary learning for preservice teachers. Their interdisciplinary understanding is significant for enhancing the role of competency-based learning in order to develop teachers' skills, especially understanding other subjects, in Thailand's education system (Anurit, 2014). Unfortunately, the researcher finds that there is very limited research on the development of Thai preservice teachers' necessary interdisciplinary understanding in English as a foreign language (EFL) education in Thailand. The skill development of a subject teacher, especially concerning EFL, is fairly difficult in Thailand. Therefore, the interdisciplinary learning process is significant for developing competency-based learning and improving the quality of learning and teaching skills among preservice teachers. Interdisciplinary learning also encourages collaboration between teachers and improves communication skills as these are not individual perceptions but encompass continuous collaboration and exchange between the various agents participating in the EFL education process (Santaolalla et al., 2020). Despite the significant impact of interdisciplinary learning on preservice teachers, the educational system in Thailand faces a significant challenge because interdisciplinary learning projects have not been implemented in teacher education programs. Moreover, Thailand lacks practical learning systems that analyze the impacts of interdisciplinary learning proposals on preservice teachers.

Several researchers have highlighted the lack of empirical studies of teachers' interdisciplinary learning, and thus, further studies exploring the impact of integrated approaches on the results of preservice teacher learning need to be conducted. Therefore, some researchers have asserted the need for focusing on interdisciplinary understanding in EFL education as it is an unchartered territory in the current curricula. Without interdisciplinary understanding, they fail to achieve the high order thinking skills of the cognitive domain—applying, understanding, analyzing, and creating language. Interdisciplinary understanding encourages preservice teachers to teach their students critical thinking skills to improve high order thinking skills, which includes self-directed learning, intellectual curiosity, self-correcting, problem-solving, and applying knowledge (Santaolalla et al., 2020).

Therefore, given its significant role in the innovation projects of educational systems, interdisciplinary learning needs to be incorporated into Thai education. Contemporary cultural, technological, social, and production aspects impact the need for interdisciplinary approaches and the need for preservice teachers to approach education using critical thinking. In the Thai education system, the importance of competence is reflected in the incorporation of fundamental knowledge into basic skills. Thus, connecting the interdisciplinary approach with basic skills is needed as these educational

improvement aspects stimulate learning and teaching (Santaolalla et al., 2020). The acquisition of teaching skills that link with interdisciplinary learning in teacher education enhances educational quality for preservice teachers, imparting to them various teaching methods that can be subsequently applied in their lessons. Innovation in education based on an interdisciplinary approach is challenging as preservice teachers typically receive monodisciplinary instructions because of the compartmentalization of the educational system into several unrelated subjects. This is a significant barrier, hampering the focus on interdisciplinary understanding in classes. Managing the challenges faced by educators in the educational system regarding the interdisciplinary approach necessitates the development of an innovative program that can serve as an example of or approach to interdisciplinary understanding for preservice teachers in the education sector.

To achieve the goals of improving the interdisciplinary understanding and speaking skills of 40 preservice teachers in a university in Thailand, the researcher proposed an AI-enhanced phenomenon-based learning (PhenoBL) instruction for two main reasons. First, PhenoBL is transdisciplinary, collaborative, exploratory, and learner-centered and occurs in a problem-solving environment, where the instructor poses problems or asks questions about the selected world phenomenon, and when preservice teachers work as a team to find answers to the problems or questions posed about a phenomenon that interests them. In phenomenon-based education, studying and understanding a phenomenon together begins with asking questions or raising challenges (Silander, 2015a). Several features of PhenoBL distinguish it from other methods as it combines project-, inquiry-, and problem-based learning, that involve many disciplines in studying one topic. The phenomenon-based approach supports learning in accord with some of the most elegant, sophisticated, and innovative educational ideas, such as project and portfolio learning (Silander, 2015b). Consequently, it provides or encourages the creation of a structure and place for their practical execution involving uses of inquiry to increase students' performance (Shaibu, 2017). It combines the best components of these ideologies and places them in a larger perspective that is frequently lacking in education. The learner is in charge of learning, not the teacher. PhenoBL is designed to ensure that learners are not only the epicenter of the entire learning process (making it relevant to their lives) but also in charge of their learning. As a result, learners' overall learning experience becomes immensely important to them, encouraging them to continue learning.

Second, advancement in digital technology has led to the development of AI, which has been significant for EFL classrooms and English language teaching (ELT) learners with limited opportunities to practice the language communicatively and purposefully. The continuous growth of AI has significantly fostered the learning experience in ELT classrooms and also a high level in the development of teaching skills (Abdelghany Elgohary & Al-Dossary, 2022). Teachers' guidance and creativity, integrated with AI technology, are powerful tools for engaging and encouraging students' development in their language learning process (Al Mukhallafi, 2020). In addition, studies in the current use of AI in ELT illustrate that AI technologies have improved the speaking and listening skills of EFL students. Singh et al. (2021) contend that AI technologies have

simplified setting goals for teachers and students in Thailand to improve their English skills. Most importantly, AI has enabled students to analyze their performance in various English learning areas, allowing them to understand and overcome their weaknesses. AI technologies have also empowered teachers to integrate applications in their practice in order to help evaluate their students' grammar comprehension and performance in other areas of learning (Singh et al., 2021). In addition, AI can be incorporated in many tasks such as AI social media tools, AI used in restaurant and other service sectors, pronunciation practice etc., which can benefit students in their real-world situation.

Therefore, AI-enhanced PhenoBL instruction was carried out in this study with the primary purpose to simultaneously facilitate the development of preservice teachers' interdisciplinary understanding and speaking skills necessary for effective teaching. It was hypothesized that after participating in the study the experimental group participants taught using the AI-enhanced PhenoBL instruction would have improved interdisciplinary understanding and speaking skills. In addition, after participating in the study, the experimental group participants would have higher interdisciplinary understanding and speaking skills than the control group participants.

METHOD

The researcher created a quasi-experimental research design. The participants were a group of preservice teachers in a university in Thailand. They were tested before and after the implementation of the AI-enhanced PhenoBL over a period of four months to assess the development of interdisciplinary understanding and speaking skills.

Participants

Preservice teachers attending the English for Teachers II course were recruited from a university in Thailand using a multistage-random sampling approach, which were divided into three stages. In stage 1, the researcher applied a stratified random sampling technique in which the strata were formed based on participants' majors. From the population of 257 preservice teachers enrolled in the course, a random sample of 80 preservice teachers was obtained based on the proportion of the population in four majors including social studies, English, Thai, and computer. In stage 2, random assignment was employed to divide the sample into two groups including 40 preservice teachers in one group and the other 40 in another group. In stage 3, random treatment was carried out using the lottery method to assign one group to an experiment group and the other one to a control group. Both groups of participants were then provided with the pertinent information on the study's objective—to investigate whether AI-enhanced PhenoBL instruction could enhance interdisciplinary understanding and speaking skills—and of the processes that would enable them to make an informed decision about participating in the study. All participants signed the consent forms for voluntary participation in the study.

Research Instruments

The instruments used in the experiment included eleven lesson plans, two sets of tests, and two analytic rubrics.

Lesson Plans

The lesson plans developed for preservice teachers aimed to provide information on topics such as pollution, the AI tools employed in lessons, instructional approaches/methods/techniques, and course evaluation. Eleven lesson plans were created, with each lesson lasting five hours, amounting to a total course duration of 55 hours.

In one of the lessons, for instance, the topic of AI was introduced to preservice teachers via Zoom videoconferencing. During this lesson, the class began with warm-up activities like creating a word web about the term "artificial intelligence," designing mind maps about familiar AI concepts, and asking questions, including "What is AI?" "Could AI replace human?" "What kinds of service can AI provide?" and "What are the pros and cons of AI?" After these activities, the class was shown a YouTube video about Google Assistant Intelligence (AI) making a haircut appointment and a restaurant reservation with a real human. The participants discovered other activities that Google AI could perform by searching online resources as well as creating and having real-life conversations using Google AI. They then introspected on their experiences in a written report. This was followed by the next activity of discussing recommendation systems on social media services (e.g., Facebook, Instagram), video streaming services (e.g., YouTube), and online shopping services (e.g., Shoppee, Lazada, PowerBuy, Amazon). After discussion, the participants were encouraged to pose some questions about the recommendation system and were prompted to answer their own questions or those of their peers. The activity ended with having them create a small business based on the information they found in the previous activity. For example, they found some trendy clothes, medicine, or furniture recommended by AI; they therefore create a business plan including the items they would sell, and AI Sales Assistant used for product sales such as People.ai, Scratchpad, Veloxy etc. They were also encouraged to investigate other AI Sales Assistant tools for further sale improvement.

Five experts being versed in English teaching, other disciplines and/or course design were requested to respond to what extent they agreed or disagreed with the statements concerning the aspects included in the lesson plans. The form consisted of seven main sections: rationale, learning outcomes, course content, instructional activities, artificial intelligence tools and technology, assessment and evaluation, and other suggestions or comments. A total of 14 items included: clear concepts of the lessons, clear intended outcomes of the instruction, observable and measurable objectives, up-to-date content, relevance of content to predetermined learning objectives, relevance of instructional activities to artificial intelligence-enhanced phenomenon-based learning (AI-enhanced PhenoBL), relevance of instructional activities to research objectives, instructional activities enabling learning achievement, practical instructional activities, appropriate allocated time, AI tools and technology relevant to contents, AI tools and technology enhancing students' interdisciplinary understanding and speaking skills, interesting AI tools and technology that can keep students motivated and engaged in the lessons, relevance of assessment and evaluation to learning objectives. The four-point scoring

scales, which were exemplary (4), proficient (3), developing (2), and limited (1), were applied for each of the aforementioned items.

Tests and Rubrics

The researcher developed two sets of pretests and posttests: 1) a subjective test in which participants were required to create a three-hour lesson plan on an assigned topic (e.g., justice, planet, textiles) to demonstrate their interdisciplinary understanding and 2) a speaking test on an assigned topic (10–15 minutes) to display their speaking skills. These tests were conducted at the beginning and upon completion of the AI-enhanced PhenoBL instructions.

To assess the pretest and posttest results, the researcher developed two rubrics, one for evaluating a three-hour lesson plan that demonstrated their interdisciplinary understanding and one for assessing an oral presentation that illustrated their speaking skills. The first rubric had four criteria: interdisciplinary tasks/activities, insightful questioning, gathering useful information from multiple disciplines, and adaptability and creativity. The second had seven criteria: content, organization, vocabulary, error, pronunciation, fluency, and confidence. The four-point scoring scales, which were exemplary (4), proficient (3), developing (2), and limited (1), were applied for each of the aforementioned rubrics. Under each criterion, there was a statement which clearly stated expected performance for each of the four scores. With the sum of the scores received from each criterion on the rubrics, the scores ranged from 1–16 for the interdisciplinary understanding, and from 1–28 for the speaking skills.

The index of item-object congruence was examined to evaluate content validity of the test items and rubrics. The same five experts who evaluated the lesson plans determined the extent to which the test items and rubrics conformed to the specified composites of skills they were to measure. The results indicated good content validity with the item-objective congruence value was 1.0. In addition, the difficulty index (p) and discrimination index (r) were calculated, and reliability was evaluated by conducting a pilot study in which one-item tests were given to 40 students who shared characteristics with those students participating in the main study; two assessors assessed these tests with the process of interrater reliability to establish the reliability, and further discussion was held in case any disagreements regarding scores arose.

For the subjective pre- and posttests assessed using the interdisciplinary understanding rubric, the results showed a difficulty level of $p = 0.341$ and 0.330, respectively, and good discrimination ($r = 0.963$ and 0.963, respectively). Cronbach's alpha was.978 and .980, respectively, indicating excellent reliability. For the speaking pre- and posttests evaluated using the speaking skills rubric, the results revealed a difficulty level $p = 0.312$ and 0.313, respectively, and good discrimination ($r = 0.993$ and 0.993, respectively). Cronbach's alpha was.993 and .991, respectively, indicating excellent reliability.

Data Collection and Analysis

Data were collected over a period of three months at a university in Thailand to answer the two research questions pertaining to what extent the AI-enhanced PhenoBL

instruction could simultaneously improve the preservice teachers' interdisciplinary understanding and speaking skills (research objective one), and to what extent the experimental group participants' interdisciplinary understanding and speaking skills would improve compared to the control group participants after participating in the study (research objective two).

Before the implementation of the AI-enhanced PhenoBL instruction, participants performed pretests on the first day of the program. During the study, classroom activities were video recorded with the unstructured observation approach. Upon the completion of the study, they performed posttests on the thirteenth week. Different analytical rubrics were used to score each of the two aforementioned items (i.e., interdisciplinary lesson plan and oral presentation) in both pretests and posttests. The total score achieved for each test item was calculated from the number of criteria and the four-point scoring scales of the rubrics. Two assessors, the researcher and an additional assessor, evaluated the test using identical rubrics. The process of interrater reliability took place to establish the reliability of the grading, and further discussion was held in case any disagreements regarding scores arose.

Descriptive statistics including mean scores, standard deviations, and percentage were employed to analyze quantitative data collected in this study. In addition, the researcher performed inferential statistics using multivariate analysis of variance (multivariate ANOVA) using Levene's test for equality of error variances, tests for normality, Bartlett's test for sphericity, and Box's test for equality of covariance matrices for assumption test. In addition, multivariate tests using Wilks' lambda were applied for hypothesis testing, and relative gain (RG) scores were used to assess the improvement of interdisciplinary understanding and speaking skills.

Ethical Considerations

Ethical aspects were carefully considered while conducting this study. The researcher followed a specific code of conduct during the data collection process to ensure that the rights of the study participants were protected, to enhance the validity of the experiment, and to maintain scientific integrity. In compliance with ethical procedures, the research proposal was submitted to an institutional review board to verify the ethical appropriateness of the aims and design of the study as well as whether these adhered to the code of conduct of the institution. Additionally, the procedures and materials employed in the study were verified to ensure that they satisfied the code of conduct. After receipt of ethical approval to conduct the study, the researcher commenced the data collection process in accordance with the approved procedures.

Before beginning the study, participants were informed about the study's objective of evaluating the progress of their interdisciplinary comprehension and speaking abilities, in addition to the processes that would be applied as the program continued. A consent form containing details, such as the study title, author's background, process through which the research would be conducted, potential risks, and gains anticipated from the study, was provided to each participant. Furthermore, the form provided assurances to the participants regarding the confidentiality of their data, their right to decline or

withdraw from participation in research at any time without penalty or any negative consequences, and their ability to ask the author questions about details of the study.

The participants were offered sufficient time before the commencement of the program to contemplate their decision regarding whether to participate and offer their consent. They were assured that refusing to participate would have no effect on their grades in the future and their relations with the author would not be impacted. In particular, maintaining the confidentiality of all the artifacts they submitted and the data that was recorded was emphasized. They were informed that the files storing their data was password-protected and accessible to only the researcher and the other assessor, and use of the data would be restricted to purely educational research. Furthermore, the participants were advised that their identities would not be made publicly available once the findings of the study were released.

FINDINGS

This section presents the experimental results from the pretests and posttests conducted, respectively, during the first and last sessions of the program. Data are presented in two primary sections on 1) the simultaneous development of experimental group participants' interdisciplinary understanding and speaking skills, and 2) a comparison of the experimental group participants' and the control group participants' interdisciplinary understanding and speaking skills development, with three subsections based on the means of analysis: 1) means and standard deviation, 2) multivariate ANOVA, and 3) RG scores.

Development of Experimental Group Participants' Interdisciplinary Understanding and Speaking Skills

Means and Standard Deviation

Forty preservice teachers were required to take pretests and posttests, which were evaluated using analytic rubrics to determine the extent to which their interdisciplinary understanding and speaking skills had improved after they participated in the AI-supplemented PhenoBL instruction. Figures 1 and 2, and Tables 1 and 2 present the increase in mean scores for interdisciplinary understanding (16 points total) and speaking skills (28 points total) across the two tests.
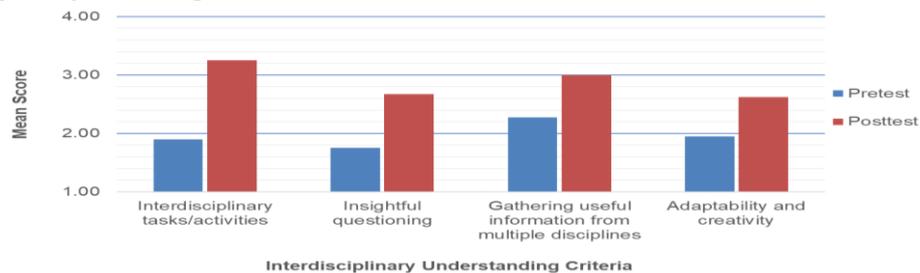
Figure 1
Comparison of mean scores for the four interdisciplinary understanding criteria across the two tests

Table 1

Pretest and posttest means and standard deviations of interdisciplinary understanding (n = 40)

Interdisciplinary Understanding Criteria	Pretest		Posttest	
	M	SD	M	SD
Interdisciplinary tasks/activities	1.9	0.67	3.25	0.74
Insightful questioning	1.75	0.54	2.68	0.66
Gathering useful information from multiple disciplines	2.28	0.78	3.00	0.75
Adaptability and creativity	1.95	0.81	2.63	0.70
Overall	7.88	2.40	11.55	2.56

The data presented in Table 1 show that the mean score and standard deviation increased reasonably from pretest (M = 7.88, SD = 2.40) to posttest (M = 11.55, SD = 2.56). When considering each performance criterion of interdisciplinary understanding and speaking skills, the results obtained from the subjective test of the assigned topic revealed that preservice teachers performed better in the posttest in terms of all four criteria.
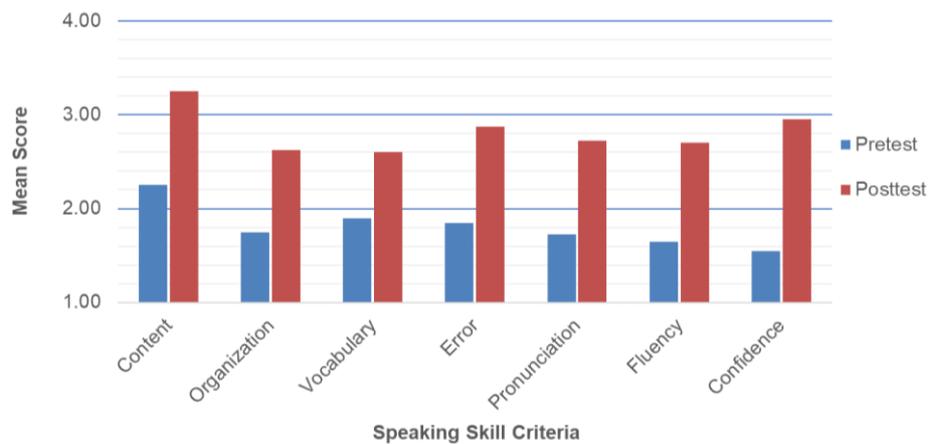
Figure 2

Comparison of mean scores for the seven speaking skills criteria across the two tests

Table 2
Pretest and posttest means and standard deviations of speaking skills (n = 40)

Speaking Skill Criteria	Pretest		Posttest	
	M	SD	M	SD
Content	2.25	0.67	3.25	0.67
Organization	1.75	0.54	2.63	0.74
Vocabulary	1.90	0.59	2.60	0.63
Error	1.85	0.58	2.88	0.76
Pronunciation	1.73	0.82	2.73	0.68
Fluency	1.65	0.48	2.70	0.97
Confidence	1.55	0.50	2.95	0.99
Overall	12.68	3.36	19.73	4.69

The data presented in Table 2 show that the mean score and standard deviation increased reasonably from pretest (M = 12.68, SD = 3.36) to posttest 2 (M = 19.73, SD = 4.69). Considering each performance criterion of speaking skills, the results obtained from the oral presentation of the assigned topic reveal that preservice teachers performed better in the posttest for all seven criteria.

Multivariate ANOVA

In the multivariate ANOVA procedure, the linear combination of the dependent variables, interdisciplinary understanding and speaking skills, was employed for analysis. To test the assumptions, the following four tests were conducted: 1) Tests for normality, 2) Levene's test for equality of error variances, 3) Bartlett's test for sphericity, and 4) Box's test for equality of covariance matrices. In addition, multivariate tests using Wilks' lambda were conducted to test the hypothesis.

1) Assumption Testing: Normality Test, Levene's Test, Bartlett's Test, Box's Test

The results of the tests for normality as presented in Table 3 demonstrate that both interdisciplinary understanding and speaking skills variables were normally distributed in a normal curve, with the value of Sk/Std. Error and Ku/Std. Error less than 2.000.

Table 3
Results of normality tests

Tests for Normality			
		Sk/Std. Error	Ku/Std.Error
Interdisciplinary understanding	Pretest	-.197/.374 = .526	-1.437/.733 = 1.960
	Posttest	-.557/.374 = 1.489	-1.004/.733 = 1.369
Speaking skills	Pretest	-.325/.374 = .868	-1.362/.733 = 1.858
	Posttest	-.408/.374 = 1.090	-.405/.733 = .552

The results of Levene's test for equality of error variances reveal equal variances and no significant difference between the variances in the pretest and posttest mean scores for interdisciplinary understanding and speaking skills (F = .005, df1 = 1, df2 = 78, p = .944) and (F = 3.091, df1 = 1, df2 = 78, p = .083), respectively.

The results of Bartlett's test for sphericity reveal that the two dependent variables, interdisciplinary understanding and speaking skills, were correlated and were not an identity matrix at the .05 significance level (chi-square = 19.837, df = 2, p = .000).

The results of Box's M test for the homogeneity of the variance-covariance matrices reveal homogeneous covariance matrices for the mean scores for the interdisciplinary understanding and speaking skills obtained from the pretest and the posttest (Box's M = 5.362, F = 1.738, df1 = 3, df2 = 1095120.00, p = .157), indicating that the assumptions are valid.

2) Hypothesis Testing: Wilks' lambda

As illustrated in Table 4, the multivariate test results reveal that the mean vector of interdisciplinary understanding and speaking skills between the pretest and the posttest differed statistically at the .05 significance level (Wilks' lambda = .026, F = 1427.126, df = 2, 77, p-value = .000).

Table 4
Results of multivariate tests using Hotelling's T-Squared

Effect		Value	F	Hypothesis df	Error df	Sig.
AI-enhanced PhenoBL instruction	Pillai's trace	.974	1427.126*	2.000	77.000	.000
	Wilks' lambda	.026	1427.126*	2.000	77.000	.000
	Hotelling's trace	37.068	1427.126*	2.000	77.000	.000
	Roy's largest root	37.068	1427.126*	2.000	77.000	.000

*p-value ≤ .05

RG Scores

Kanjanawasee's (2013) RG score was employed to determine the positive gain in preservice teachers' interdisciplinary understanding and speaking skills after participating in the AI-enhanced PhenoBL instruction. The following intervals were used to interpret the levels of improvement in the study: 100 to 76 (very high), 75 to 50 (high), 50 to 26 (medium), and 25 to 0 (low) (Kanjanawasee, 2013). The analysis of the improvement of interdisciplinary understanding and speaking skills is elaborated on below.

1) Interdisciplinary Understanding

The data presented in Figure 3 show the positive RG scores for each of the four performance criteria of interdisciplinary understanding, which were categorized under the medium level of improvement. The highest RG (64.29%) score was awarded for the "interdisciplinary tasks/activities" criterion, followed by the "gathering useful information from multiple disciplines," "adaptability and creativity," and "insightful questioning" criteria, with RG scores of 42.03%, 41.11%, and 32.93%, respectively.
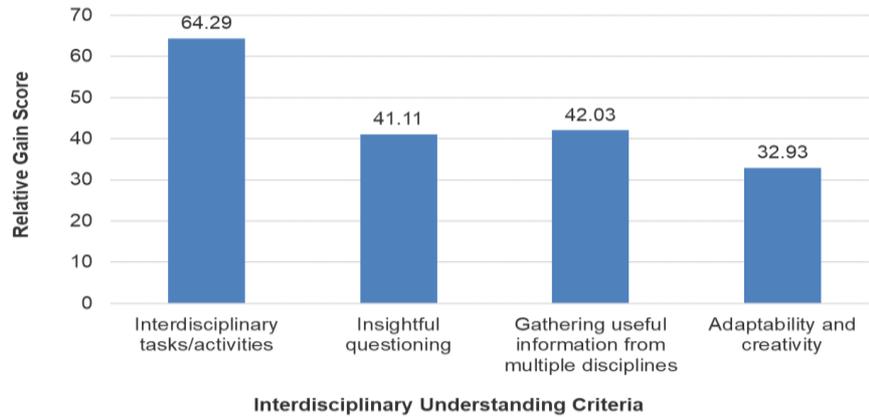
Figure 3

Relative gain scores of interdisciplinary understanding criteria across the two tests

2) Speaking Skills

The data presented in Figure 4 show the RG scores for each of the seven performance criteria of speaking skills, which were categorized under the medium level of improvement. The highest RG score (57.14%) was awarded for both "confidence" and "content" criteria. The "error" criterion received an RG score of 47.67%, followed by the "fluency," "pronunciation," "organization," and "vocabulary" criteria with RG scores of 44.68%, 43.96%, 38.89%, and 33.33%, respectively.
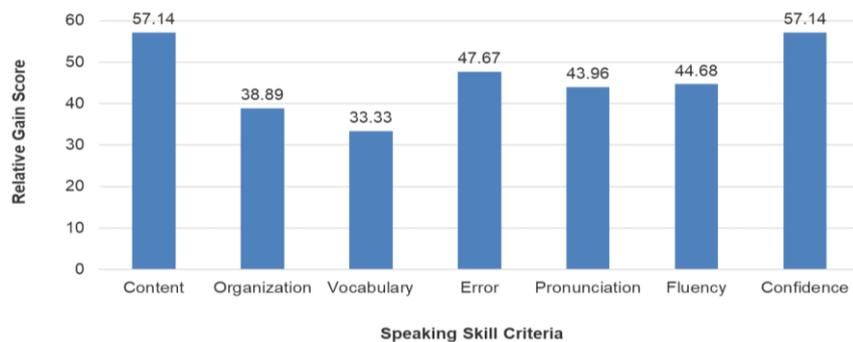
Figure 4
Relative gain scores of speaking skill criteria across the two tests

A Comparison of the Experimental Group Participants' and the Control Group Participants' Interdisciplinary Understanding and Speaking Skills Development

Means and Standard Deviation

Eighty preservice teachers (40 in the experimental group and 40 in the control group) were required to take the same pretests and posttests, which were evaluated using

analytic rubrics to determine the extent to which their interdisciplinary understanding and speaking skills had improved after they participated in the AI-supplemented PhenoBL instruction. Figure 5 and Table 5 present the increase in mean scores for interdisciplinary understanding (16 points total) and speaking skills (28 points total) between the two groups.
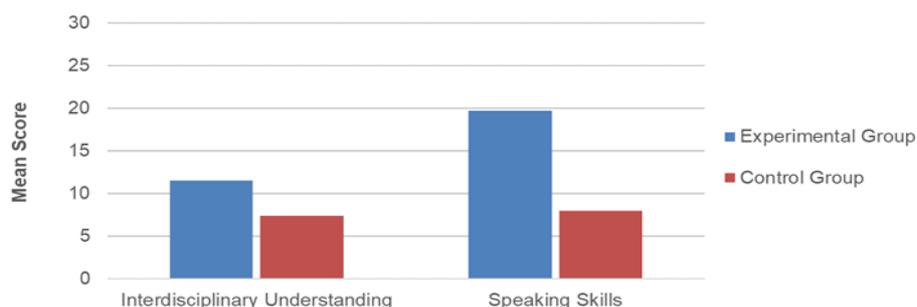
Figure 5

Comparison of mean scores for the interdisciplinary understanding and speaking skills between the experimental and control groups (n = 80)

Table 5

Means and standard deviations of interdisciplinary understanding and speaking variables between the experimental and control groups (n = 80)

Dependent Variables	Experimental Group)n = 40(		Control Group)n = 40(	
	M	SD	M	SD
Interdisciplinary understanding	11.55	2.56	7.43	2.51
Speaking skills	19.73	4.69	8.03	3.22

The data presented in Table 5 show that higher scores for both interdisciplinary understanding and speaking skills were attained by the experimental group than the control group. Considering the interdisciplinary understanding, experimental group (M = 11.55, SD = 2.56) scored significantly higher than the control group (M = 7.43, SD = 2.51). Considering the speaking skills, experimental group (M = 19.73, SD = 4.69) scored significantly higher than the control group (M = 8.03, SD = 3.22).

Multivariate ANOVA

In the multivariate ANOVA procedure, the linear combination of the dependent variables, interdisciplinary understanding and speaking skills, was employed for analysis. To test the assumptions, the following four tests were conducted: 1) Tests for normality, 2) Levene's test for equality of error variances, 3) Bartlett's test for sphericity, and 4) Box's test for equality of covariance matrices.

1) Assumption Testing: Normality Test, Levene's Test, Bartlett's Test, Box's Test

According to the analysis of common assumptions of the data obtained from participants in the experimental group engaging in the AI-enhanced PhenoBL instruction and

participants in the control group engaging in the conventional teaching method, the following results were found:

To assess the distribution of two dependent variables using Sk/Std. Error and Ku/Std.Error, the yielded values are expected to not be greater than 2.00. The distribution assessment of the two dependent variables in the experimental and control groups showed that the yielded values were not greater than 2.00. Therefore, the obtained data were normally distributed as presented in Table 6.

Table 6
Results of normality tests

Tests for Normality			
		Sk/Std. Error	Ku/Std.Error
Interdisciplinary understanding	Experimental	-.557/.374 = 1.489	-1.004/.733 = 1.369
	Control	.011/.374 = .029	-1.322/.733 = 1.803
Speaking skills	Experimental	-.408/.374 = 1.090	-.405/.733 = .552
	Control	.107/.374 = .286	-1.403/.733 = 1.914

The results of Levene's test for equality of error variances reveal equal variances and no significant difference between the variances of interdisciplinary understanding and speaking skills derived from the experimental group engaging in the AI-enhanced PhenoBL instruction and from the control group engaging in the conventional teaching method (p-value = .995 and .055 respectively).

The results of Bartlett's test for sphericity reveal that the two dependent variables, interdisciplinary understanding and speaking skills, were significantly correlated and were not an identity matrix at the .05 significance level (chi-square = 15.876, df=2, p = .000).

The result of Box's M test for the homogeneity of the variance-covariance matrices reveal homogeneous covariance matrices for the mean scores for the interdisciplinary understanding and speaking skills obtained from the pretest and the posttest (Box's M = 5.817, F = 1.885, df1 = 3, df2 = 1095120.00, p = .130), indicating that the assumptions are valid.

In conclusion, the analysis of the data obtained from the experimental group engaging in AI-enhanced PhenoBL instruction and the control group engaging in the conventional teaching method show that it satisfied common assumptions for statistical testing and was adequate for Multivariate Analysis of Variance (MANOVA) conducted to test the research hypotheses.

DISCUSSION AND CONCLUSION

This study aimed to ascertain whether AI-enhanced PhenoBL instruction could improve the interdisciplinary understanding and speaking skills of preservice teachers. The results demonstrated an increase in the mean scores for both interdisciplinary understanding and speaking skills from the initial to the final test, suggesting that the AI-enhanced PhenoBL instruction indeed benefits preservice teachers. In particular, the preservice teachers had an overall mean of 7.88 in interdisciplinary skills and 12.68 in

English-speaking skills before taking AI-enhanced PhenoBL instruction, which increased remarkably to 11.55 and 19.73, respectively after pariticipating in AI-enhanced PhenoBL teaching. In addition, after participating in the study, the experimental group participants had higher interdisciplinary understanding and speaking skills than the control group participants. With regards to interdisciplinary understanding, experimental group (M = 11.55, SD = 2.56) scored significantly higher than the control group (M = 7.43, SD = 2.51). Considering the speaking skills, experimental group (M = 19.73, SD = 4.69) scored significantly higher than the control group (M = 8.03, SD = 3.22).

These results are similar to those in previous studies. Lonka (2015) contended that PhenoBL improved critical areas of learning, such as school engagement, academic interest, and social-emotional learning. Wakil et al. (2019) similarly ascertained that students' test performance increased by more than 10% as a result of implementing PhenoBL. Furthermore, Lonka (2015), Nguyen (2018), Wakil et al. (2019), and Naik (2019) unanimously agreed that the fundamental features of PhenoBL lie in its interdisciplinary approach. In addition, Wakil et al. (2019) asserted that the interdisciplinary approach of PhenoBL motivates students' learning and enhances their skills and abilities, including both interdisciplinary and speaking skills.

The results of the study are also aligned with those of Naik (2019), who stated that Finland developed the PhenoBL as a strategy to promote interdisciplinary and language teaching, among many other skills, rendering the nation's education system apt and practical. Lonka (2015) also argued that nations should use technology to enhance students' education in classrooms. Therefore, given the inefficient use of technology in teaching, most governments should reform their educational policies to enable students to utilize AI to enhance their learning and comprehension of a range of skills and abilities. Given the emphasis placed on these fundamental talents, Finland has created groundbreaking educational methods that aid in developing a broad knowledge base and a thorough understanding of abilities that are essential for tackling a wide range of challenges in the 21st century (Lonka, 2015).

This study focused on aspects of two sets of skills—interdisciplinary understanding and English-speaking skills altogether—and it utilized a range of criteria to illustrate improvement in each skill via AI-enhanced PhenoBL. The set of interdisciplinary skills included performing interdisciplinary tasks and activities, insightful questioning, gathering valuable information from multiple disciplines, and adaptability and creativity under interdisciplinary skills. The set of English-speaking skills included content, organization, vocabulary, error, pronunciation, fluency, and confidence. The preservice teachers exhibited remarkable improvement in all these criteria. Mølstad and Karseth (2016) asserted that PhenoBL differs from traditional learning methods, which are passive and only involve memorization of the curricula, as in conventional teaching where teachers were standing in front of the room instructing, and students were seated at their desks, listening attentively and taking notes (Schug, 2013).

A critical component of PhenoBL is student participation in hands-on tasks focused on responding to queries and solving problems (Grusche, 2019). Wakil et al. (2019) argued

that a greater focus on practical education and student independence facilitates critical learning. In addition to boosting students' critical thinking and creativity, PhenoBL also provides opportunities to polish communication, cooperation, and game-based learning abilities through several learning activities (Nguyen, 2018). Naik (2019) contended that PhenoBL realized positive outcomes, particularly developing skills, critical analysis, and student engagement.

The results also showed that integarting AI and PhenoBL into teaching notably improved each criterion of the students' interdisciplinary skills, particularly interdisciplinary tasks, insightful questioning, adaptability, and creativity. It also improved students' English-speaking abilities, vocabulary, pronunciation, and fluency. In addition, the findings revealed that the students need more advanced training in English, as reflected in their progress after receiving AI-enhanced PhenoBL. One probable reason for this could be that although the students developed their English competence in class, they were unfamiliar with the premise of applying what they had learned in class to practice in front of others. Therefore, the implementation plan must be modified to enable students to suitably apply the skills learned by including English-speaking activities. The results of this study are consistent with those in Saeheng (2017) and Wongsuriya (2020), who documented that project-based, interdisciplinary activities improved students' language skills (such as their English-speaking abilities). The reading, writing, and speaking abilities of the study participants and their vocabulary and translation proficiency improved by utilizing English-speaking skills to collect, analyze, and synthesize information while working on their projects. Furthermore, according to Alamri (2021), students who integrated their language abilities and assimilated information from various sources were better able to communicate in real-world situations.

This study discovered that AI-enhanced PhenoBL has two significant practical implications. First, incorporating an interdisciplinary initiative of this kind into English lessons is needed, particularly in countries where avenues for speaking English publicly are limited. Second, as PhenoBL enables students to employ language skills they have acquired in class while also building their communication skills in English, including a initiative of this kind as part of a school's grading system seems to be a suitable option. Poonpon (2017) outlined a framework for teaching PhenoBL in a classroom setting. After outlining the project's theme, the instructor asked a group of students to analyze a theme and articulate how they would present their perspectives on it. Each group was provided a period of one or two weeks to prepare their speech before presenting their project in class. Following their presentation, the instructor noted any areas for development in each group and interacted with each group individually concerning the adjustments required. Additionally, Ostergaard et al. (2010) stated that phenomenon-based education provides learners with a framework to engage with the outer world. A direct effect of such a framework is that students are exposed to real-world circumstances that demand critical thinking and the application of their knowledge to deliver meaningful answers.

LIMITATIONS

Firstly, in this study, most non-English major preservice teachers' inadequate prior knowledge and interest in the English language possibly led to exhaustion and frustration. Therefore, they needed motivation and additional support compared to English-major preservice teachers who majored in English, who had better language proficiency and could meet the academic standards of the course. Second, as this study explored how to improve students' interdisciplinary and public speaking abilities, the scope was restricted in terms of studying the general advantages of AI-enhanced PhenoBL.

REFERENCES

Abdelghany Elgohary, H. K., & Al-Dossary, H. K. (2022). The effectiveness of an educational environment based on artificial intelligence techniques using virtual classrooms on training development. *International Journal of Instruction, 15*(4), 1133–1150. https://doi.org/10.29333/iji.2022.15460a

Al Mukhallafi, T. R. A. (2020). Using artificial intelligence for developing English language teaching/learning: An analytical study from university students' perspective. *International Journal of English Linguistics, 10*(6), 40–53. https://doi.org/10.5539/ijel.v10n6p40

Alamri, M. M. (2021). Using blended project-based learning for students' behavioral intention to use and academic achievement in higher education. *Education Sciences, 11*(5), 207–219. https://doi.org/10.3390/educsci11050207

Anurit, T. (2014). Interdisciplinary learning to promote critical thinking skills in Thai EFL students. *Journal of Humanities, 1*(11), 1–21.

Chinwonno, A. (2001), *A comparison of Thai and English reading comprehension strategies of pre-service teachers in Thailand*. Ohio University Press.

Dincer, A., & Yesilyurt, S. (2013). Pre-service English teachers' beliefs on speaking skill based on motivational orientations. *English Language Teaching, 6*(7), 88–95. https://doi.org/10.5539/elt.v6n7p88

Grusche, S. (2019). Phenomenon-based learning and model-based teaching: Do they match? *Journal of Physics: Conference Series*. IOP Publishing, *1287*(1), 012066. https://doi.org/10.1088/1742-6596/1287/1/012066

Kanjanawasee, S. (2013). *Classical Test Theory* (7th ed.). Chulalongkorn University Publishing.

Loan, N. T. T. (2019). Reflective teaching in an EFL writing instruction course for Thai pre-service teachers. *Journal of Asia TEFL, 16*(2), 561-575.

Lonka, K. (2015), *Innovative schools: Teaching and learning in the digital era*. European Parliament. European Union.

Mølstad, C. E., & Karseth, B. (2016). National curricula in Norway and Finland: The role of learning outcomes. *European Educational Research Journal, 15*(3), 329-344. https://doi.org/10.1177/1474904116639311

Naik, R. P. (2019), *Phenomenon-based learning in Finland.* Jyväskylä: University of Jyväskylä.

Nguyen, H. P. (2018), *Phenomenon-based learning in Finnish and Vietnamese upper secondary school curriculum for English as a foreign language*. University of Jyväskylä.

Oeamoum, N., & Sriwichai, C. (2020). Problems and needs in English language teaching from the viewpoints of pre-service English teachers in Thailand. *Asian Journal of Education and Training, 6*(4), 592–601. https://doi.org/10.20448/journal.522.2020.64.592.601

Østergaard, E., Lieblein, G., Breland, T. A., & Francis, C. (2010). Students learning agroecology: Phenomenon-based education for responsible action. *Journal of Agricultural Education and Extension, 16*(1), 23–37. https://doi.org/10.1080/13892240903533053

Owolade, A. O., Salami, M. O., Kareem, A. O., Oladipupo, P. O. (2022). Effectiveness of guided inquiry and open inquiry instructional strategies in improving biology students' achievement. *Anatolian Journal of Education, 7*(2), 19–30. https://doi.org/10.29333/aje.2022.723a

Poonpon, K. (2017). Enhancing English skills through project-based learning. *English Teacher, 40*, 1–10.

Saeheng, P. (2017). A study of e-learning, blended learning, and traditional teaching methods to motivate autonomous learning in English reading comprehension of Thais learners. *IJELTAL, 2*(1), 1–20. https://doi.org/10.21093/ijeltal.v2i1.36

Santaolalla, E., Urosa, B., Martín, O., Verde, A., & Díaz, T. (2020). Interdisciplinarity in teacher education: Evaluation of the effectiveness of an educational innovation project. *Sustainability, 12*(17), 1–23. https://doi.org/10.3390/su12176748

Silander, P. (2015a). Digital pedagogy. In P. Mattila & P. Silander (Eds.), *How to create the school of the future: Revolutionary thinking and design from Finland* (pp. 9–26). University of Oulu, Center for Internet Excellence.

Silander, P. (2015b). Rubric for phenomenon-based learning. Retrieved from http://www.phenomenaleducation.info/phenomenon-based-learning

Singh, G., Kaur, G., & Thaemphan, L. (2021). Artificial intelligence in ESL-a survey on school students and teachers in Ubon Ratchathani Thailand. *The Eleventh Annual International Symposium of Foreign Language Learning (11th AISOFOLL)* (pp. 344–349).

Wakil, K., Rahman, R., Hasan, D., Mahmood, P., & Jalal, T. (2019). Phenomenon-based learning for teaching ICT subject through other subjects in primary schools. *Journal of Computer and Education Research, 7*(13), 205–212. https://doi.org/10.18009/jcer.553507

Wongsuriya, P. (2020). Improving the Thai students' ability in English pronunciation through mobile application. *Educational Research and Reviews, 15*(4), 175–185.