Jurnal Pendidikan Teknik Mesin
Volume 10 nomor 1 (Mei 2023)
Available Online at https://ejournal.unsri.ac.id/index.php/ptm/index
p-ISSN: 2355-7389, e-ISSN: 2656-5153

ENHANCEMENT OF CRITICAL THINKING SKILLS VOCATIONAL STUDENTS IN INDONESIA USING PROBLEM-BASED LEARNING-STEM BY E-LEARNING

Yadi, Farhan[1*], Herminarto Sofyan[2], Herman Dwi Surjono[2]
[1]Universitas Sriwijaya
[2]Universitas Negeri Yogyakarta
[*]Corresponding Author: farhan@unsri.ac.id

Info Artikel

Sejarah Artikel:
Diterima : April 2023
Disetujui : Mei 2023
Dipublikasikan : mei 2023

Keywords:
Vocational Education, Problem-Based learning, STEM, E-learning, Critical Thinking

Abstract

This study aims to develop a problem-based learning model with the Science, Technology, Engineering, and Mathematics (STEM) approach through e-learning in vocational high schools (SMK) to enhance learning outcomes and critical thinking skills of the students of light vehicle electrical maintenance, which is valid, practical, and effective. The design phase began with determining the fundamental philosophy, theoretical background, hypothetical model, and evaluation aspects, while exhaustive tests were conducted to validate the developed system during the development phase. Based on practical testing in four vocational high schools in Indonesia, the developed PBL-STEM e-learning system is considered valid and effective for enhancing the quality of vocational education. The critical thinking skills of experimental class students have increased, both in groups and individually, as indicated by the progressive percentage increase of 38.32%. The application of this model has proven to be effective in enhancing critical thinking, skills.

INTRODUCTION

In Indonesia, technical skills are primarily taught in vocational high schools. Nonetheless, regrettably, education enhancements supported by research for technological development are still conducted at the university level. However, technical and vocational education and training (TVET) at the high school level must be extended as economies and communities, including industrial communities, are growing (Catts, Falk, and Wallace, 2011). Therefore, vocational education enhancement at the high school level must be based on technical practice. For field practice-based learning to provide students with the knowledge necessary to solve problems, it must be based on real-world problems (Problem-based Learning/ PBL) and incorporate knowledge of science, technology, engineering, and mathematics (STEM). In addition, STEM provides the theoretical knowledge and industry-related competencies that aim to prepare STEM graduates to be leaders capable of meeting the needs of the 21st century (Zizka, McGunagle, and Clark, 2021).

In Vocational Education, besides technical skills, thinking skills, a cognitive process broken down into concrete steps and then used as thinking guidelines, is mandatory. For example, inference, the ability to connect clues, facts, or information with existing knowledge to predict an outcome, is required for vocational graduates to be met with the industrial revolution. One of the thinking skills that must be developed to be part of a global industrial citizen is critical thinking skill (Rodríguez-Sabiote, Olmedo-Moreno, and Exposito-Lopez, 2022). The development must be carried out in teamwork since collaboration and partnership is one of the keywords in the market competition in the future.

Problem-based learning (PBL) methods have received much attention from educational researchers in response to the challenges of vocational education, which should provide case-based, real-world learning experiences. Besides, PBL can also sharpen the student's critical thinking. Lonergan, Cumming, and O'neill (2022) have studied the effectiveness of PBL in various secondary school settings. By utilizing statistical methods, the significant differences in student comprehension before and after PBL implementation have been identified.

Wu, Hsieh, and Wu (2022) have highlighted two specific issues. There are technology use and constructivist learning. Its proposal further defines three key determinants, technology affordances, e-content interaction, and e-social interaction, as the determinants of e-learning use and academic performance. One of the benefits of an e-learning system in vocational education is the flexibility of content based on the learner's interest. For that purpose, Alatrash, Priyadarshini, Ezaldeen, and Alhinnawi (2022) have developed a system based on the Artificial Intelligence technique to conduct sentiment analysis in determining the content of an e-learning system. During the Covid-19 pandemic, e-learning plays a significant role in education (Bao, 2020).

Li, Zheng, Huang, and Xie (2022) investigated STEM learning in a simulation environment. In that study, self-regulated learning was considered a complex dynamical system and used to analyze the behaviour of high and low performers when learning about engineering. The result of the study shows that low performers require high repetition of a self-regulated learning system than the high performers.

STEM does not concentrate only on the education process because it is designed to ensure its graduates meet industry standards. Study about alumni of the STEM learning method has been investigated by Lavi, Tal, and Dori (2021). The study identifies and matches form to skill by analyzing 1500 alumni. The study's findings indicate that three critical factors in the twenty-first century are domain-general skills, which include complex problem-solving, critical thinking, individual learning, and question-posing. The soft skills include creativity, collaboration, intercultural communication, entrepreneurship, and oral and written communication, and STEM-specific skills, including engineering design, experimenting and testing, and STEM knowledge application.

Based on the literature mentioned above review, it can be concluded that a problem-based e-learning system for vocational education within a STEM framework has great potential to develop. To the authors' best knowledge, that system has not been developed by previous researchers in the education field and will be developed in this study.

RESEARCH METHOD

The teaching-learning system developed in this study is a PBL-STEM in an e-learning environment. In STEM, the teaching-learning system design must prioritize three essential factors: the subject that learns, which is the students, the subject who teaches, who are the teachers; and the object to learn, which is the education subjects. Therefore, the students must be considered first when developing the e-learning system. An effective PBL assumes that the learning process must be constructive, self-directed, collaborative, and contextual (Dolmans, De Grave, Wolfhagen, and van der Vleuten, 2005). Besides, the learning process must be done in a group to encourage students to co-construct the necessary knowledge (Yew and Goh, 2016).

Therefore, in this study, there are five steps implemented in the PBL that involve the students: group formation, problem understanding by the group, individual reporting, discussing new knowledge, and continuing the process until the group can solve the problem. The teacher primarily functions as a tutor who stimulates elaboration, integration of knowledge, and student interaction by asking questions, requesting clarifications, and applying knowledge. Students and group registration and discussion forum with a subject would then be the modules to be developed for the e-learning system based on these steps.

The next step is considering the subject to teach, which is the teachers. In STEM, the focus of the teaching process is learner-centred teaching which makes the teacher facilitate and encourage the students as learners to be more active in the learning process. The roles of the teacher are to register the subjects, determine learning strategy, including synchronous, asynchronous, and blended, define the learning media based on learner-centred teaching and evaluate the students' performance.

The last step is understanding the required process to register the teaching and learning materials that meet the STEM requirements. In the beginning, the teachers will determine the subjects to lecture and their objectives. Every subject will be evaluated to ensure the content has physical word phenomenon (science), application of science (technology), design or development or building (engineering), and computation (mathematics). The evaluation for every subject will be based on thinking skills, creativity, and communication, with the parameters to evaluate those items shown in Table 1.

Table 1. Parameters for every evaluation item

Evaluation Item	Evaluation Parameters
Critical thinking skill	a. ability to clarify the case study b. ability to predict and anticipate the problem c. ability to make decision d. ability to conclude

The explanation above will then be used to develop a Use Case Diagram (UCD) and Data Flow Diagram (DFD) as the basis for the e-learning system development. Figure 1 shows the UCD, while Figure 2 shows the DFD level-1 of the developed e-learning system.

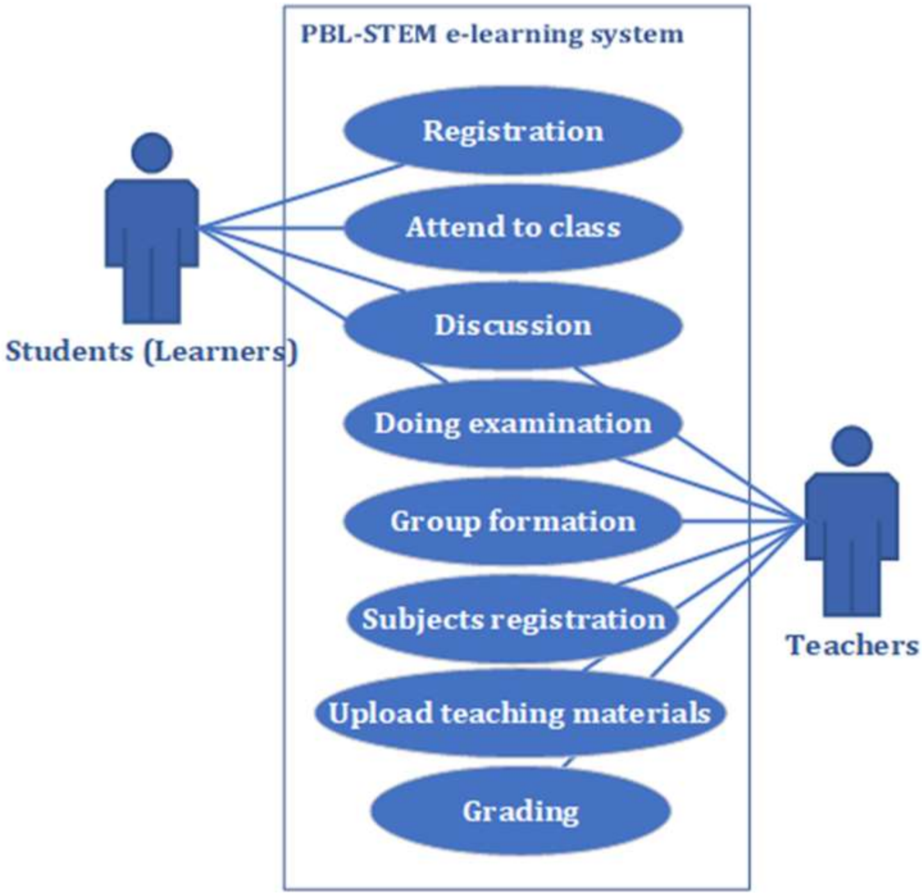

Figure 1. The USD of the PBL-STEM e-learning system

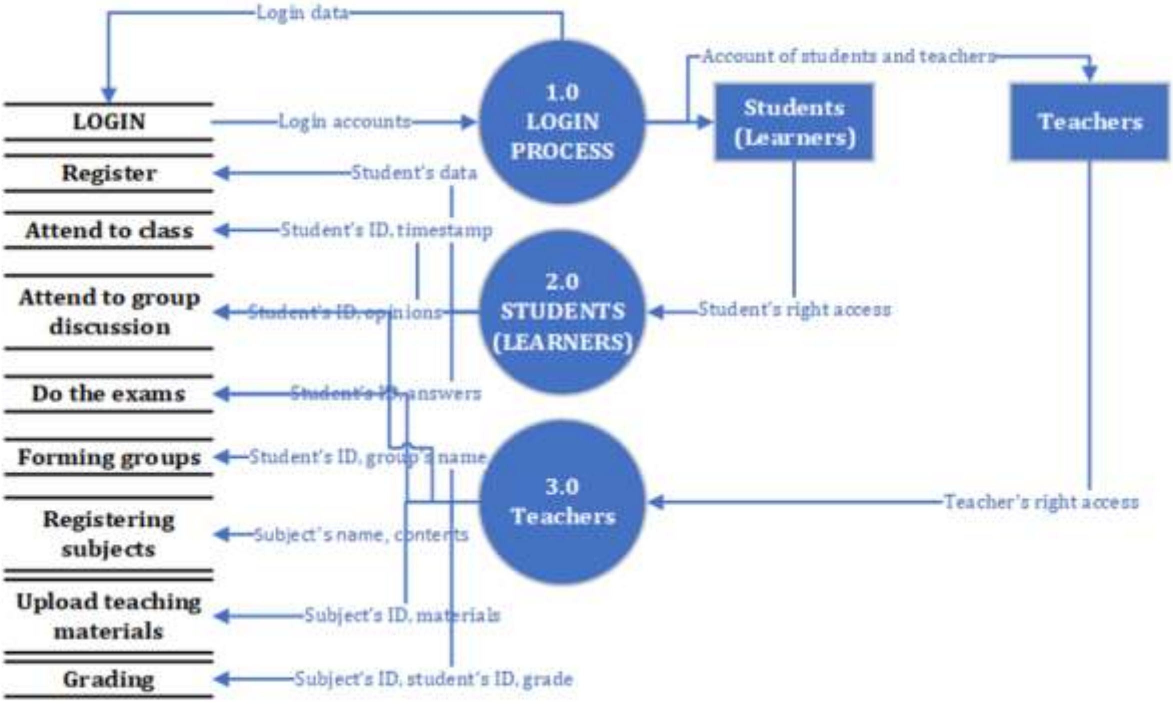

Figure 2. DFD level-1 of the developed e-learning system

From the DFD level-1, it can be understood that the proposed PBL-STEM e-learning system enables the teacher to upload the subject's content, divide it into sections, and include a learning strategy. An asynchronous learning strategy is implemented for theories that must be repeated independently by students. Then, for the theory portion that requires evaluating student comprehension, the learning system is blended, where real-world case studies will be uploaded, and students must comprehend and discuss in a group. Then, students are instructed to provide their perspectives and exchange viewpoints within their groups until a solution is agreed upon and deemed the best. The instructor will evaluate student group discussions and propose solutions. In the final phase, the teacher will hold a synchronous session to listen to each group's presentation and facilitate discussions to ensure the students achieve the learning objectives.

Model design trials are carried out in two forms, namely limited trials (small group) and expanded trials. In the limited trial, the number of participants involved in this activity were 4 subject teachers and 6 students who were selected purposively on the considerations and suggestions of the teachers due to the conditions of the COVID-19 pandemic. The design of this trial was made simulative where the role of the developer/ the researcher is very dominant, assisted by the teacher in carrying out this activity. Researcher's observation of the process of using the initial draft of the model by users is very important, where this is useful for determining the objectivity and feasibility of the model and efforts to improve it, as well as assessment and input from teachers and students, need to be facilitated.

The population of this study were SMK students who were carrying out learning in PKKR subjects and teachers teaching PKKR subjects in Palembang City. The size of the study sample was determined by using the "Puposive Sample" technique, where this technique was carried out in consideration of the large number of public and private vocational schools in the automotive sector in the city of Palembang. The number of automotive vocational schools in Palembang is 20, of which there are 5 state vocational schools and 15 private vocational schools.

Data collection instruments include interview guidelines, documentation, input and suggestion sheets, expert assessment questionnaires, model questionnaires and learning achievement tests. Data analysis used followed the research process at each stage including Qualitative descriptive analysis, content validity, instrument reliability, quantitative descriptive analysis, normality, homogeneity, comparative analysis and effectiveness test.

RESULT AND DISCUSSION

Content Validation

In this study, to test the feasibility of the PBL-STEM e-learning system model, its content is validated using expert opinions on four factors: skill explanation, instrument indicators, grading criteria, and language. Table 2 displays the result of the five expert's judgments.

Based on Table 2 below, Aiken's V score of the proposed instruments is 0.91, meaning that the proposed teaching-learning content is adequate. The agreement index of the validators is an essential factor to measure when developing an instrument. Table 3 shows the ICC calculation result for the agreement index of the validators, and based on Table 3, the ICC is considered very reliable

Table 1. Result of the validator judgment for the proposed teaching-learning content

No	Aspects	Validators					Average	Aiken's V
		1	2	3	4	5		
1	Explaining skills Components in the instruments are the explanation of every skill to be obtained	4	4	4	4	4	4	1
2	Instrument indicators							
	a. The indicators are measurable	3	3	3	3	3	3	0.87
	b. The indicators are formulated from the skills to be obtained	4	4	3	4	3	3.6	0.87
3	Grading criterions							
	a. The criteria are clearly explained	4	4	4	3	4	3.8	0.93
	b. Score of every criterion is clearly stated.	4	3	4	4	4	3.8	0.93
4	Language aspects							
	a. Using standard language	4	4	4	4	4	4	0.87
	b. Using easy-to-understand sentences.	4	4	4	4	4	4	0.93
	c. Using sentences that do not cause multiple interpretations	4	4	4	3	3	3.6	0.87
	Average						3.725	0.91
5	General validation						Status	
	Rubric for assessment and scoring of student skills	A	A	A	A	A	Very good	

Table 2. The ICC Calculation Result

Intraclass Correlation Coefficient

	Intraclass Correlation	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.412[a]	.111	.789	4.356	7	28	.002
Average Measures	.778[c]	.383	.949	4.356	7	28	.002

Validity and Reliability Test the feasibility and practicality of the model

Validity and reliability of the feasibility and practicality test for the PBL-STEM e-learning system and the manual books used by teachers and students were also conducted; the result is shown in Table 4.

Table 3. Validity and Reliability Test the feasibility and practicality of the model

No	Evaluation Aspects	Validators					Average	Aiken's V	ICC	
		1	2	3	4	5			Average	Individual
1	Feasibility of model									
a	Instructions	4	4	4	4	4	4.00			
b	Model	3.67	3.67	3.83	3.67	3.83	3.73			
c	Content	4	4	4	4	4	4.00			
d	Implementation mechanism	4	4	4	4	4	4.00	0.940	0.649	0.270
e	Language	3	4	3.50	3.50	3.50	3.50			
f	Assessment rubric	4	3	4	4	4	3.80			
g	General validation	A	A	A	A	A	Feasible, with minor correction			
2	Practicality of model for teachers									
a	Instructions	4	4	4	4	3.50	3.83			
b	Model	3.67	3.83	3.67	3.67	3.83	3.73			
c	Content	3	4	4	4	4	3.80			
d	Implementation mechanism	3.67	3.67	3.67	3.33	3.33	3.53	0.920	0.826	0.487
e	Language	4	4	4	4	4	4.00			
f	Assessment rubric	4	4	4	4	4	4.00			
g	General validation	A	A	A	A	A	Feasible. with minor correction			
3	The practicality of the model for students									
a	Instructions	4	4	4	4	4	4.00			
b	Model	3.83	3.67	3.67	3.67	3.83	3.73			
c	Content	4	4	4	4	4	4.00			
d	Implementation mechanism	4	4	4	4	4	4.00	0.930	0.649	0.270
e	Language	3.50	4	3	3.50	3.50	3.50			
f	Assessment rubric	4	3	3	4	3	3.60			
g	General validation	A	A	B	A	A	Feasible, with minor correction			

From Table 4, it can be seen that Aiken's V of all of the aspects is more than 0.5, and in table 2 the average ICC for all of the aspects is fallen in the range of reliable to very reliable.

Content validation for the evaluation instruments

To evaluate the students learning, test instruments are also developed. However, before being used, the evaluation instruments have been tested for their content and construction. The content validation emphasizes the validity of the proposed evaluation instruments associated with the factors to be evaluated. Even though the content validation can be carried out based on teacher opinions, five teachers as validators gave their views to get objective results. The validation is based on the Likert scale, and the result is shown in Table 5.

Table 4. Content validation for the evaluation instruments

No	Evaluation Aspects	Validators					Average	Aiken's V	ICC	
		1	2	3	4	5			Average	Individual
1	Multiple choices (35 questions) Question and choices	3.74	3.80	3.71	3.71	3.77	3.75	0.920	0.791	0.431
2	Essay (5 questions) Question description	3.50	3.67	3.67	3.50	3.50	3.57	0.890	0.839	0.510
3	Analysis (9 questions) Description and recency	3.78	4.00	3.78	3.78	3.78	3.82	0.940	0.929	0.724
4	General validation	A	A	A	A	A	Feasible, with minor revisions			

From Table 5, it can be seen that Aiken's V of all of the aspects is more than 0.5 and based on Table 5, the average ICC for all of the aspects is fallen in the range of reliable to very reliable. Minor revisions are still required to improve the proposed evaluation instruments. The construct validity emphasizes how far the evaluation instruments are related theoretically to the factors to be measured. Therefore, construct validity is carried out for every question, and Pearson correlation is used as the correlation measurement. Since the correlation between the questions and the factors to be measured is still unknown, then a two-tailed significance test is also used. The construct validity of the multiple choices questions is carried out by asking 14 students (N = 14) to answer the questions, and the result is shown in Table 6.

Table 5. Construct validation of the evaluation instrument

No.	Pearson Correlation	Sig. 2 tailed	Validation Result $R_{count} > R_{table}$, r_{table} = 0.532	**No.**	**Pearson Correlation**	**Sig. 2 tailed**	Validation Result $R_{count} > R_{table}$, r_{table} = 0.532
Q1	.616*	.019	Valid	Q26	.577*	.031	Valid
Q2	.650*	.012	Valid	Q27	.750**	.002	Valid
Q3	.572*	.033	Valid	Q28	.593*	.025	Valid
Q4	.629*	.016	Valid	Q29	-.292	.311	Invalid
Q5	.580*	.030	Valid	Q30	.671**	.009	Valid
Q6	.171	.559	Invalid	Q31	.583*	.029	Valid
Q7	.550*	.042	Valid	Q32	.831**	.000	Valid
Q8	.626*	.017	Valid	Q33	-.156	.594	Invalid
Q9	-.010	.972	Invalid	Q34	.578*	.030	Valid
Q10	.554*	.040	Valid	Q35	.602*	.023	Valid
Q11	.564*	.036	Valid	Q36	.616*	.019	Valid
Q12	-.297	.302	Invalid	Q37	.831**	.000	Valid
Q13	.564*	.036	Valid	Q38	.172	.558	Invalid
Q14	.656*	.011	Valid	Q39	.061	.837	Invalid
Q15	.621*	.018	Valid	Q40	-.169	.562	Invalid

No.	Pearson Correlation	Sig. 2 tailed	Validation Result $R_{count} > R_{table}$, r_{table} = 0.532	No.	**Pearson Correlation**	**Sig. 2 tailed**	Validation Result $R_{count} > R_{table}$, r_{table} = 0.532
Q16	.831**	.000	Valid	Q41	.593*	.025	Valid
Q17	.556*	.039	Valid	Q42	-.214	.463	Invalid
Q18	.045	.879	Invalid	Q43	.029	.921	Invalid
Q19	.784**	.001	Valid	Q44	.831**	.000	Valid
Q20	.650*	.012	Valid	Q45	-.275	.342	Invalid
Q21	.564*	.036	Valid	Q46	-.378	.182	Invalid
Q22	.650*	.012	Valid	Q47	-.049	.867	Invalid
Q23	.661*	.010	Valid	Q48	.535*	.049	Valid
Q24	-.018	.951	Invalid	Q49	.691**	.006	Valid
Q25	.540*	.046	Valid	Q50	.831**	.000	Valid

According to Table 6, 15 invalid questions must be excluded from the evaluation instrument. As a result, the evaluation instrument contains only 35 questions. The following validation is the reliability test for the 35 questions utilizing the KR-20 formula, as depicted in Equation 1 (Kuder and Richardson, 1937).

$$r = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K} p_i q_i}{\delta_x^2}\right)$$
$$\delta_x^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n} \tag{1}$$

K is the number of question items, *p* is the number of students with correct responses to the test item, *q* is 1-*p*, *X* is the question score, and r is the reliability score. The reliability score for the 35 questions is $r = 0.966$, greater than the r-table for 14 samples with a 95% significance level of 0.532. Therefore, the reliability of the evaluation instruments can be determined.

The following validation is the discriminating power test to determine whether the proposed evaluation instrument can distinguish two successful or unsuccessful conditions (Smalldon and Moffat, 1973). Such a test measures the evaluation instrument's ability to distinguish students with high understanding and students with low understanding. The formula to test the discriminating power of the evaluation instrument is depicted in Equation 4, while the category of the index of distinguishing power is shown in Table 7. *DP* is the index of discriminating power, B_A is students' high understanding, B_B is students' low understanding, J_A is the number of students in the high-understanding group, and J_B is the number of students in the low-understanding group.

$$DP = \frac{B_A}{J_A} - \frac{B_B}{J_B} \tag{2}$$

Table 6. Category of the index of discriminating power

DP	Category
0.00 – 0.19	Bad
0.20 – 0.39	Sufficient
0.40 – 0.69	Good
0.70 – 1.00	Very good
Negative value	Must be excluded

From the 35 questions, there are 3 questions in the "very good" category, 27 questions in the "good" category, and five questions in the "sufficient" category. The same students are used to assess the construct validity of the essay questions as for the multiple-choice questions. Table 8 displays the results of the construct validity test of the essay questions, while Table 9 displays the reliability test results based on Cronbach's alpha value.

Table 7. Result of the construct validity of the essay questions

Question No	Pearson Correlation	Significance two-tailed	Validation Result
E1	.484	.079	Invalid
E2	.830**	.000	Valid
E3	.603*	.022	Valid
E4	.529	.052	Invalid
E5	.362	.203	Invalid
E6	.430	.125	Invalid
E7	.697**	.006	Valid
E8	.626*	.017	Valid
E9	.121	.681	Invalid
E10	.724**	.003	Valid

Table 8. Result of the reliability test of the essay questions

Case Processing Summary

		N	%
Cases	Valid	14	100.0
	Excluded	0	0.0
	Total	14	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.769	5

Practical Implementation of the Developed PBL-STEM e-Learning System

Limited Area Practical test

The purpose of this limited practical test is to evaluate the implementation of the developed PBL-STEM e-learning system in a limited area. Four senior teachers provided the materials, and six students were purposely chosen based on their subject performance. During the teaching-learning process, several evaluators have also given evaluations for the PBL-STEM e-learning system in addition to the four teachers and six students. The results are shown in Tables 10 and Table 11.

Table 9. Result of teachers' feasibility of the model

No.	Evaluation Aspect	Validator (Teachers)				Score (%)	Average
		I	II	III	IV		
1	Content	20	20	18	17	93.75	3.75
2	Learning Acceleration	8	8	7	8	96.88	3.88
3	Feedback and Adaptation	8	8	8	7	96.88	3.88
4	Motivation	4	4	3	4	93.75	3.75
5	Presentation Design	4	4	4	4	100.00	4.00
6	Interaction and Usability	8	8	8	8	100.00	4.00
7	Accessibility	4	4	4	4	100.00	4.00
8	Standards Compliance	4	4	4	4	100.00	4.00
9	Group Formation	8	8	8	8	100.00	4.00
10	Problem Identification	11	12	12	12	97.92	3.92

No.	Evaluation Aspect	Validator (Teachers)				Score (%)	Average
		I	II	III	IV		
11	Problem Determination	7	8	7	8	93.75	3.75
12	Brainstorming	8	8	7	8	96.88	3.88
13	Problem Formulation	4	4	4	4	100.00	4.00
14	Ideas	4	4	4	4	100.00	4.00
15	Build the knowledge	8	8	8	8	100.00	4.00
16	Summarise the solution	8	8	7	8	96.88	3.88
Average						**97.92**	**3.92**

Table 10. Result of student feasibility of the model

No.	Evaluation Aspect	Respondent (Students)						Sum	Score (%)	Average
		R1	R2	R3	R4	R5	R6			
1	Model Usability	16	14	15	16	16	16	93	96.88%	3.88
2	Readability of Guidelines	18	20	18	20	19	20	115	95.83%	3.83
3	Model Implementation	14	16	16	16	15	16	93	96.88%	3.88
4	Ease of test	12	11	12	11	11	12	69	95.83%	3.83
Average									**96.35%**	**3.85**

Based on Tables 10 and Table 15, it can be concluded that the PBL-STEM e-learning system is feasible and applicable for use as a teaching-learning system in vocational senior high schools

Expanded Area Practical test

The next test is conducted by implementing the PBL-STEM e-learning system for four vocational senior high schools in Palembang, Indonesia, that have received an "A" accreditation from the Indonesian Minister of Education. Four classes from each school serve as the study subjects and are divided into the control and experimental groups. A control group is a student group with a teaching-learning process supported by the PBL-STEM system. In contrast, an experimental group is a student group with a teaching-learning process supported by the PBL-STEM e-learning system. Table 12 and 13 shows the number of students and teachers from every school and group.

Table 11. Number of students from every school and group

School	Control Group		Experiment Group	
	Class	Students number	Class	Students number
School A	XII	16	XII	24
School B	XII	21	XII	17
School C	XII	20	XII	20
School D	XII	17	XII	23
Total no. of students	74		83	

Table 12. Number of teachers from every school and group

School	Control Group		Experiment Group	
	Class	Teachers number	Class	Teachers number
School A	XII	1	XII	1
School B	XII	1	XII	1
School C	XII	1	XII	1
School D	XII	1	XII	1
Total no. of teachers	4		4	

Two tests were carried out for all of the students in both groups: pre-test and post-test. *The n-gain* score proposed by Meltzer is used to justify the improvement level of implementing the PBL-STEM with and without an e-learning system. The formula to calculate the *N-gain* score is depicted in Equation 5, while the criteria for the *N-gain* score are shown in Table 14. (Meltzer, 2002).

$$g = \frac{os - es}{ps - es} \tag{5}$$

Where:
g : N-gain; os : post-test score; es : pre-test score; ps : maximum possible score

Table 13. *N-gain* score criteria

N-gain score	Criteria
$g > 0.7$	High
$0.3 \leq g \leq 0.7$	Medium
$G < 0.3$	Low

The average score of average pre-test, post-test and N-gain scores with $ps = 100$ for both groups is shown in Table 15.

Table 14. Average pre-test, post-test and N-gain for all of the groups

No	School	Control Group			Experiment Groups		
		Pre-test	Post-test	*N-gain*	Pre-test	Post-test	*N-gain*
1	School A	41.25	52.50	0.19	35.50	65.40	0.47
2	School B	31.90	46.62	0.22	35.12	64.82	0.46
3	School C	19.50	30.90	0.14	27.85	62.20	0.48
4	School D	28.12	39.76	0.16	30.35	67.30	0.53
Average		30.19	42.45	0.18	32.20	64.97	0.48

Based on the *N-gain* value in Table 18, it can be concluded that the PBL system can enhance students' STEM knowledge. However, the proposed e-learning system can significantly improve students' STEM comprehension.

Critical Thinking Skill Measurement

The critical thinking skill was evaluated through Focus Group Discussions (FGDs). After the teacher gives a subject material, the teacher will ask the students to conduct an FGD and evaluate the student's performance based on the rubric of the subject. In this study, the FGD was carried out three times for every subject, and Table 16 shows the average score of the students for the 3 FGDs.

Table 15. The score of the FGD in the evaluation of the critical thinking skill

No	School Name	Critical Thinking Skill						
		D_1	D_2	D_3	Average	Score Improvement (%)		
						D_2 - D_1	D_3 - D_2	D_3 - D_1
1	School A	18.92	23.04	25.33	22.43	21.81	9.95	33.92
2	School B	16.76	20.94	24.71	20.80	24.91	17.98	47.37
3	School C	17.10	21.70	25.20	21.33	26.90	16.13	47.37
4	School D	17.30	19.43	21.70	19.48	12.28	11.68	25.40
	Average	**17.52**	**21.28**	**24.23**	**21.01**	**21.44**	**13.89**	**38.32**

The improved progress in the critical thinking skill of the students is depicted in Figure 2

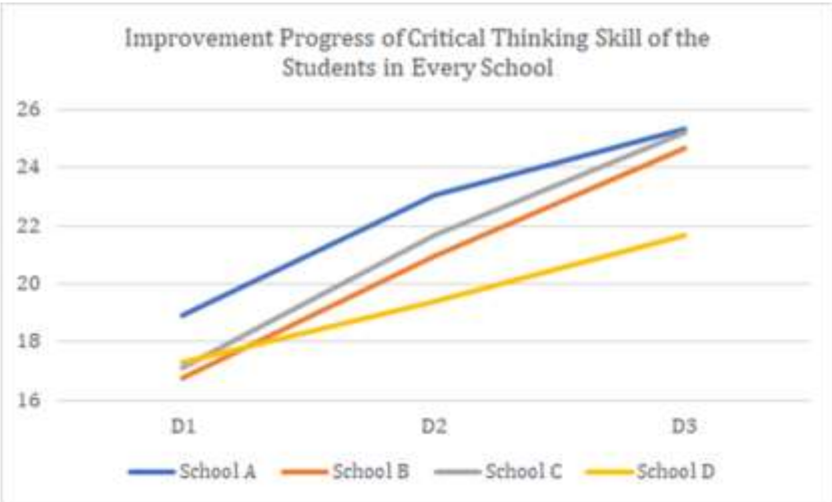

Figure 1. Improvement progress of critical thinking of the students in every school

Based on Figure 2, it can be said that the student's critical thinking skills can be improved linearly up to 38.32 %. Statistical test based on Wilcoxon Sign Rank Test is also carried out to justify whether the proposed PBL-STEM e-learning system can improve students' critical thinking skills. The result of the test is shown in Table 17.

Table 16. Difference test on the critical thinking skill based on Wilcoxon Sign Rank test

Test Statistics

	D2 - D1	D3 - D2	D3 - D1
Z	-7.795^b	-7.157^b	-7.896^b
Asymp. Sig. (2-tailed)	.000	.000	.000

Based on Table 17 above, the significant value of the difference of every test is less than 0.05, it means that, in general, the proposed PBL-STEM e-learning system can be used to improve the critical thinking skill of the students.

CONCLUSION

This paper describes developing a PBL-STEM e-learning system, beginning with conceptual development, testing learning instrumentation, evaluating learning outcomes comprehensively, and testing variables used to evaluate the learning success of vocational high school students in Indonesia. The proposed PBL-STEM e-learning system permits students to repeat learning materials requiring independent study and to engage in discussions to determine the root causes of problems. In addition, the proposed blended learning system permits teachers to evaluate students' critical thinking skills. The results of measuring the progressivity of students' critical thinking skills have increased both individually and in groups, especially after the application of the model reaches the final stage. Seen from the increase in the percentage of the average score of 38.32% and a significance of 0.00. This system's design concept is easily transferable to other learning systems implementing a STEM-based problem-based learning system in an e-learning environment. Therefore, it is recommended to extend this study by implementing it in different education systems with the support of the latest technologies, such as Augmented Reality (AR) and Virtual Reality (VR), to enable the students to interact in an immersive environment.

REFERENCE

Aiken, L.R. 1985. Three coefficients for analyzing the reliability and validity of ratings. *Educational and psychological measurement*, 45(1), pp. 131-142.

Alatrash, R., Priyadarshini, R., Ezaldeen, H., & Alhinnawi, A. 2022. Augmented language model with deep learning adaptation on sentiment analysis for E-learning recommendation. *Cognitive Systems Research*, 75, pp. 53-69.

Bao, W. 2020. COVID-19 and online teaching in higher education: a case study of Peking University. Human Behavior and Emerging Technologies, 2(2), pp. 113-115.

Catts, R., Falk, I., & Wallace, R. 2011. *Vocational Learning Innovative Theory and Practice*. Springer Science+Business Media. New York.

Cronbach, L. J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), pp. 297-334.

Dolmans, D.H.J.M., De Grave, W., Wolfhagen, I.H.A.P., and van der Vleuten, C.P.M. 2005. Problem-based learning: Future challenges for educational practice and research. *Medical Education*, 39(7), pp. 732–741. https://doi.org/10.1111/j.1365-2929.2005.02205.x

Kuder, G.F., Richardson, M.W. 1937. The theory of the estimation of test reliability. *Psychometrika* 2, 151–160.

Lavi, R., Tal, M., & Dori, Y.J. 2021. Perceptions of STEM alumni and students on developing 21st century skills through methods of teaching and learning. *Studies in Educational Evaluation*, 70, 1-11.

Li, S., Zheng, J., Huang, X., & Xie, C. 2022. Self-regulated learning as a complex dynamical system: Examining students' STEM learning in a simulation environment. *Learning and Individual Differences*, 95, 1-9.

Lonergan, R., Cumming, T.M., & O'neill, S.C. 2022. Exploring the efficacy of problem-based learning in diverse secondary school classrooms: Characteristics and goals of problem-based learning. *International Journal of Educational Research*, 112, pp. 1-12.

Meltzer, D.E. 2002. The relationship between mathematics preparation and conceptual learning gains in physics: A possible "hidden variable" in diagnostic pre-test scores. *American Journal of Phys*ics 70 (12), 1259-1268.

Rodríguez-Sabiote, C., Olmedo-Moreno, E.M., and Exposito-Lopez, J. 2022. The effects of teamwork on critical thinking: A serial mediation analysis of the influence of work skills and educational motivation in secondary school students. *Thinking Skills and Creativity*, 45, 1-12.

Smalldon, K.W., and Moffat, A.C. 1973. The calculation of discriminating power for a series of correlated attributes. *Journal of the Forensic Science Society*, 13(4), pp. 291-295.

Wu, I.L., Hsieh, P.J., & Wu, S.M. 2022. Developing effective e-learning environments through e-learning use mediating technology affordance and constructivist learning aspects for performance impacts: Moderator of learner involvement. *The Internet and Higher Education*, 55, pp. 1-16.

Yew, E. H. J., and Goh, K. 2016. Problem-based learning: An overview of its process and impact on learning. *Health Professions Education*, 2(2), 75–79. https://doi.org/10.1016/j.hpe.2016.01.004.

Zizka, L., McGunagle, D.M., & Clark, P.J. 2021. Sustainability in science, technology, engineering and mathematics (STEM) programs: Authentic engagement through a community-based approach. *Journal of Cleaner Production*, 279, 1-15.