

PAPER NAME AUTHOR

13. 3465-Mairing-Similarity Check.docx J.P. Mairing

WORD COUNT CHARACTER COUNT

6846 Words 37730 Characters

PAGE COUNT FILE SIZE

15 Pages 2.5MB

SUBMISSION DATE REPORT DATE

Apr 18, 2024 11:56 AM GMT+7 Apr 18, 2024 12:09 PM GMT+7

## 2% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

• 1% Internet database

Crossref database

• 0% Submitted Works database

- 2% Publications database
- · Crossref Posted Content database

## Excluded from Similarity Report

• Bibliographic material

# Development of Problem, YouTube, and ChatGPT Learning Model to Improve Students' Proving Ability in Real Analysis

#### Abstract

The lack of experience and ability to construct logical, deductive, rigorous, and definite mathematical arguments or proofs is the main difficulty for students in addressing problems of real analysis proof. The difficulties faced by students in solving Real Analysis problems could be overcome by implementing a suitable learning model. The developmental process of problem-solving-oriented learning models with the assistance of tools in solving these problems is an appropriate solution to overcome the difficulties. Students will always be faced with proof problems to enhance their experience, and by utilizing resources such as YouTube videos and ChatGPT, they will increase their knowledge to produce the proofs. This model is closely in line with the principles of Socratic Method learning theory. The research aimed to develop a PYChat (Problem, YouTube, and ChatGPT) learning model in fulfilling the criteria of valid, effective, and practical while enhancing students' ability to solve problems. The model was produced through developmental research. The research works in several stages, including planning the initial product and research instruments, an exploratory study, developing the initial product, validating the initial product, and field testing. The subjects were mathematics education students from a state university in Central Kalimantan who were taking Real Analysis course in the academic year 2023/2024. The research instruments included the expert validation sheet, the observation sheet, the mid-test, the final test, and the initial product of the PYChat model consisted of the learning plan, the worksheets contained Real Analysis problems to prove, and the module contained ChatGPT tasks and YouTube videos links. The research results indicated that the PYChat model met validity criteria based on the average scores of the validity model by two experts was 84. The score means the PYChat model was concluded in the very valid category. The model was effective according to research findings, showing that the students in the PYChat class scored higher than those in the PBL (Problem-Based Learning) class. The PBL model has come to use in the courses of the past three years. In addition, the PYChat model was proven to be practical since its syntaxes could be applied effectively by the lecture, and all groups submitted the problem solutions in each meeting. Therefore, using Artificial Intelligent such as ChatGPT in learning has potential for the future. However, its implementation must be carefully planned and integrated with other problems or learning media.

Keywords: ChatGPT, Problem-Based Learning, Video YouTube, Real Analysis, Proof Problems

#### Abstrak

Kurangnya pengalaman dan kemampuan untuk menyusun argumen atau bukti matematika yang logis, deduktif, teliti, dan jelas merupakan kesulitan utama bagi mahasiswa dalam menangani masalah pembuktian analisis real. Kesulitan yang dihadapi oleh mahasiswa saat menyelesaikan masalah pembuktian Analisis Real dapat diatasi dengan menerapkan model pembelajaran yang sesuai. Pengembangan model pembelajaran yang berorientasi dengan penyelesaian masalah dengan alat bantu dalam penyelesaian masalah tersebut merupakan solusi yang sesuai untuk mengatasi kesulitan tersebut. Mahasiswa akan selalu berhadapan dengan masalah pembuktian untuk menambah pengalaman mereka dan dengan memanfaatkan video YouTube dan ChatGPT akan menambah pengetahuan mereka attuk membuat pembuktian tersebut. Model ini juga sejalan dengan prinsip teori pembelajaran Socratic method. 122<mark>elitian ini</mark> bertujuan <mark>untuk mengembangkan model pembelajaran</mark> PYChat (Problem, YouTube, <mark>dan</mark> ChatGPT) yang valid, efektif, dan praktis untuk meningkatkan kemampuan mahasiswa dalam menyelesaikan masalah pembuktian Analisis Real. Model tersebut dihasilkan melalui penelitian pengembangan. Tahap-tahap penelitiannya adalah perencanaan produk awal dan instrument penelitian, studi eksplorasi, pengembangan produk awal, validasi produk awal dan uji lapangan. Subyek penelitiannya adalah mahasiswa Pendidikan matematika dari salah universitas negeri di Kalimantan Tengah yang mengambil mata kuliah Analisis Real di tahun akademik 2023/2024. Instrumen penelitiannya adalah lembar validasi ahli, lembar pengamatan, ujian tengah semester, ujian akhir semester, produk awal dari model PYChat yang terdiri dari rencana pembelajaran semester, lembar kerja yang memuat masalah pembuktian, dan modul yang memuat tugas ChatGPT dan link video YouTube. Hasil penelitian menunjukkan model PYChat memenuhi kevalidan berdasarkan rata-rata skor validitas model oleh dua ahli adalah 84, yang berarti bahwa model PYChat termasuk dalam kategori sangat valid. Model PYChat efektif berdasarkan temuan penelitian yang menunjukkan bahwa mahasiswa di kelas PYChat memiliki skor lebih dari kelas PBL (Problem Based Learning). Model PBL telah dilaksanakan dalam tiga tahun terakhir dalam mata kuliah tersebut. Selain itu, model PYChat terbukti praktis karena dalam implementasinya seluruh sintaks PYChat dapat diterapkan dengan baik, dan semua kelompok mengumpulkan penyelesaian dari masalah di setiap pertemuan. Dengan demikian, penggunaan kecerdasan buatan seperti ChatGPT dalam pembelajaran memiliki potensi di masa

#### INTRODUCTION

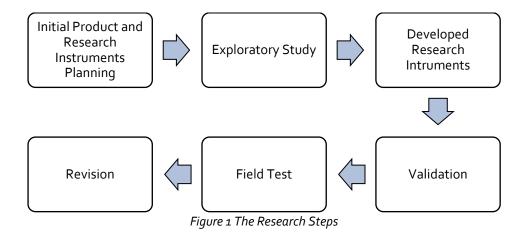
Mathematical proof is one of the crucial aspects of mathematics in higher education especially in Real Analysis course. Students are required to comprehend abstract mathematical concepts dan to solve the problems in the course. The complexity of solving the problems in the course often presents difficulties for students. The difficulties may arise from various factors. One of the factors is students' lack of experience in solving the problems that are isomorphic to the ones they are currently working on (Mairing, 2020).

Mathematical proof is typically argumentation, not derivation (Ashton, 2021). Mathematical arguments or proofs should be logical, deductive, rigorous, and definite. A proof should start with a clear statement of what is to prove, followed by a series of logical steps that lead to the conclusion. Those steps must be sequential and deductive, meaning each step corresponds to the truth obtained or learned previously. The proof should be rigorous, meaning it should be free of errors and gaps in logic way. In addition, the proof should be clear and easy to follow so that anyone with a basic understanding of the subject matter can understand it (David, 2017; Lerman, 2020; Sriraman & Umland, 2020; Umland & Sriraman, 2020). To construct a logical, deductive, rigorous, and definite mathematical argument or proof, students need to be proficient in cognitive, affective, and conative abilities are necessary.

Several researchers have explored the relationship between cognitive and affective factors, such as mathematical anxiety, self-efficacy, learning engagement, and backward reasoning, in the

of provision mathematical proofs (Shimizu, 2022). Additionally, research has investigated the impact of adversity quotient and resilience on mathematical proof problem-solving ability indicating that these factors positively affect the ability of students to solve problems (Hakim & Murtafiah, 2020). Furthermore, some research exploring students' thinking processes in mathematical proofs emphasized the crucial role of analytical thinking, the ability to accurately implement knowledge (definitions, axioms, and theorems), and their understanding and problem-solving skills. (Faizah et al., 2020; Firmasari et al., 2022; Wijayanti & Mulyono, 2021). Another research has indicated a significant difference in students' mathematical proving ability based on levels of university accreditation, suggesting the need to improve students' ability to solve problems through practice, motivation, and appropriate learning materials (Agustyaningrum et al., 2020). These studies collectively emphasized the complex interplay between cognitive, affective, and conative factors in solving problems. They also underscored the importance of fostering analytical thinking and providing adequate support and resources to enhance students' proving ability.

A learning model development is one of the solutions to provide adequate resources to make students understand and solve Real Analysis problems. PBL (Problem-Based Learning) enhances students' critical and creative thinking skills by enabling them to effectively analyze cases, propose diverse solutions, find learning more engaging, and confidently articulate their viewpoints during class



discussions (Kardoyo et al., 2020). Another research showed that using problems could encourage students to acquire meaningful knowledge, be active in their learning, and develop numeracy (Afifah et al., 2023; Mairing & Lorida, 2013). PBL would be more effective on students' learning outcomes or problem-solving ability when being assisted by media (Fidan & Tuncel, 2019). An effective medium captivates attention, particularly in this era where students can demonstrate a high interest in technology (Dewi et al., 2020; Hasanah et al., 2023; Mairing, 2020b). In addition, the PBL model and YouTube videos affected students' ability to solve mathematics problems (Mairing, 2020a, 2021). Therefore, this research proposed a learning model using problems, YouTube videos, and AI (Artificial intelligence) that can be a combination to overcome the difficulties. This research proposes a learning model, that is closely connected with the principles of the Socratic Method of learning theory, namely (1) questioning; (2) dialectical inquiry; (3) reflection; (4) active engagement; (5) critical thinking; and (6) open-ended (Delić & Bećirović, 2016; Rahmawati & Suwarjo, 2021).

One AI that can be used in learning mathematics is ChatGPT (Chat Generative Pretrained Transformer). ChatGPT is a generative language model trained by OpenAI using the transformer architecture, and it was specifically designed for conversation or chat tasks. Several studies were conducted to examine the effect of implementing ChatGPT in the classroom, and its ability to solve mathematics problems. However, a few studies were still conducted in the context of mathematics learning, especially in Real Analysis courses. The role of ChatGPT here is not only to provide solutions to problems but also to serve as an idea for students to solve problems or a benchmark solution. Research showed that ChatGPT's ability to solve advanced mathematical problems, especially in those to prove, was below the average performance of typical students (Frieder et al., 2023; Wardat et al., 2023). ChatGPT had not been able to consistently produce high-quality proofs or calculations (Shakarian et al., 2023). Nevertheless, ChatGPT exhibits higher flexibility compared to models specifically trained for a particular task. This flexibility makes ChatGPT a versatile tool suitable for its use across various mathematical fields. In this research, students criticized ChatGPT's solutions as being inappropriate. The activity could develop students' critical thinking and ability to solve problems (Elsayed, 2023; Rizaldi et al., 2021).

Some research was aimed at examining ChatGPT's capabilities in mathematics and its role in education. Halaweh

(2023) stated that ChatGPT was a beneficial tool for learning. Students were directed to demonstrate more critical thinking in evaluating information, fostering creativity and innovation, as well as developing and presenting new ideas through what was generated by ChatGPT (Guo & Lee, 2023; Marbun, 2023; van den Berg & du Plessis, 2023). Nevertheless, it still needed to be used with several cautions, and more guidelines on how to use it safely. It is because ChatGPT is regarded as a significant opportunity in diverse fields, though it is also perceived as an ethical concern, drawing mixed evaluations particularly in the realm of education (Leiter et al., 2023; Tlili et al., 2023). Thus, this research supported the use of ChatGPT as an educational aid by developing models or learning instruments. ChatGPT might have an essential future impact on self-taught learning to have more research (Firat, 2022).

Therefore, it was evident that addressing students' difficulties in solving problems, especially in the Real Analysis course. The research aimed to develop a learning model called PYChat (Problems, YouTube, and ChatGPT). It was applied to overcome the difficulties and to increase students' ability to solve problems to prove in the course. This research used AI in higher education to complement studies on the impact and learning model of using AI in mathematics learning.

#### **METHOD**

This research produced a learning model called PYChat. Therefore, the researchers used developmental research. The development of the PYChat model follows the following stages (Nieveen & Folmer, 2013). Firstly, initial product and research instruments planning. The instruments were a product of expert validation sheets, a mid-test and a final test that included problems to prove in the Real

Analysis course, and an observation sheet for lecturer activities. In addition, the researchers also determined criteria for validity, effectiveness, and practicality of the model. The model could be valid if two experts confirmed that the model aligned with learning theories and Real Analysis content. It could enhance students' problem-solving ability theoretically. The model could be effective if problem-solving ability using the PYChat in the classroom model were significantly higher than those in the previous learning model, namely PBL (Problem-Based Learning). The model fulfilled practical criteria if all groups uploaded their assignments and at least 80% of the PYChat model's syntax could be applied by the lecturer. Secondly, an exploratory study regarded the learning model implementation integrated with ChatGPT, YouTube videos, and mathematical problems. Thirdly, the researchers developed research instruments, namely the expert validation sheet, the observation sheet, the midtest, the final test, and the initial product of the PYChat model consisted of the learning plan, the worksheets contained Real Analysis problems to prove, and the module contained ChatGPT tasks and YouTube videos links. Fourthly, validation of the initial product by two experts using the sheet. Then, the researchers refined the model based on the validation results and the feedback from the experts. Fifthly, field testing in one class of mathematics education students. Sixthly, revision based on the results of the field testing to produce the PYChat model. The Figure 1 shows the graphic organizer of the research steps.

Data analysis for the validity instrument was conducted by calculating the average scores of (1) suitability to the Socratic Method learning theory, (2) suitability to Real Analysis Content, and (3) the

format, content, and language of the PY-Chat module. Validity instruments have been reviewed by experts and are converted into quantitative criteria referring to Table 1.

Table 1 Validity Criteria				
Average of validity aspects score intervals	Criteria			
$\bar{x} \ge 80$	Very Valid			
$70 \le \bar{x} < 80$	Valid			
$60 \le \bar{x} < 70$	Less Valid			
$\bar{x} < 60$	Very Invalid			

Description:  $\bar{x}$ : average of validity score

The research subjects were mathematics education students from one of the state universities in Central Kalimantan who were taking a Real Analysis course in the academic year 2023/2024. The class as the field testing of the PYChat model was selected randomly. Another class was taught by the previous learning model namely PBL. The result was that PYchat which was implemented in class B, and PBL in class A, which means the PY-Chat class (Class B) was the experimental class, while the PBL class was the control class (Class A). The number of students in the PYChat class was 24, while in the PBL class, it was 32.

The researchers administered the worksheets (link: same https://t.ly/ykm2r), mid-test, and final test to assess proving ability for both classes (the class with PYChat and PBL models). The scores of all instruments were compiled to a proving-ability with the weight of each score respectively was 0.2; 0.3 and 0.5. The final score in the Real Analysis course was influenced by the previous proficiency in the preceding course included mathematical proof, namely Abstract Algebra. The score of Abstract Algebra was the initial ability of the research subjects. The initial ability was a covariate in the research. Thus, the dependent variable was the proving ability. The independent variable was the

learning models (the treatments).

Therefore, there were two research hypotheses. The first hypothesis related to the initial ability is as follows.

$$H_0$$
:  $\mu 1_{PYChat} = \mu 1_{PBL}$   
 $H_1$ :  $\mu 1_{PYChat} \neq \mu 1_{PBL}$ 

where  $\mu 1_{PYChat} =$  average the initial ability of the students in PYChat class, and  $\mu 1_{PBL} =$  average initial ability of the students in PBL class. The second hypothesis related to the proving ability is as follows.

$$H_0$$
:  $\mu 2_{PYChat} = \mu 2_{PBL}$   
 $H_1$ :  $\mu 2_{PYChat} \neq \mu 2_{PBL}$ 

where  $\mu 2_{PYChat}$  = average the proving ability of the students in PYChat class, and  $\mu 2_{PBL}$  = average the final ability of the students in PBL class.

The researchers used ANCOVA (analysis of covariance) to conclude both hypotheses since the research involved the covariate (Mairing, 2017). The assumptions for the analysis were linearity of the model, normality of the residuals, residual independence, and homoscedasticity. If the covariance analysis concluded there were significant differences between the proving ability in PYChat and PBL classes, then the researchers administered advanced testing using Tukey researchers comparison. The Minitab 8 to analyze the data and to draw the conclusions.

#### **RESULTS AND DISCUSSION**

#### Results

**Initial Product Development** 

In this stage, the researchers developed the initial product of PYChat and the instruments. The initial product was the lesson plan, the module (link: https://t.ly/EWmpD), and the worksheets. The worksheets contained some problems to prove. The students solved them by each group before class meetings. The solutions to the problem would be presented in groups at class meetings. The learning kits of PBL were the module (link: https://t.ly/vbSAW), and the worksheets. Students in PBL class could learn proof for each theorem or solutions for some example problems in the module. The worksheets and their usage activities in the PBL class were the same as in the PYChat class.

Furthermore, PYChat's module contained tasks for the students to explore solutions to some problems or proofs of some theorems and a QR-code for accessing YouTube videos at channel jackson mairing (link: https://www.youtube.com/@jacksonmairing4566). Thus, there was no proof for each theorem and solution for each problem in the PYChat module. Figure 2 is an example of one of the tasks in the PYChat module.

(b) (-1)a = -a <b>Prelimenery Proof</b>	
ChatGPT Proof	
Final Proof	Study the following video
(c) -(-a) = a Prelimenery Proof  ChatGPT Proof	
Final Proof	Study the following video
	(MSR)
	・ ・

Figure 2 Example of Tasks on the PYChat Module

In Figure 2, the students previously solved the problems by themselves in the initial proof. After that, they asked ChatGPT for solutions' reference. Then, they learn from the available videos on the QR code in the PYChat module. Finally, they proved the problems in the final proof after they found the solutions from ChatGPT and YouTube videos.

The research Instruments were the validation sheet, the mid-test, the final test, and the lecturer's observation sheet. The problems in the mid-test are as follows.

- 1. Let functions  $f: A \to B$  and  $g: B \to C$ . Prove that if  $g \circ f$  is injective, then f is also injective!
- 2. Prove that  $2n-3 \le 2^{n-2}$  for each  $\in N, n \ge 5!$
- 3. If  $c \neq 0$  and  $d \neq 0$ , Prove that  $\frac{(-a)}{(-c)} \cdot \frac{(-b)}{(-d)} = \frac{ab}{cd}!$ 4. If a < b and c < d, Prove that ad + bc < ac + d
- bd!
- 5. If  $0 \le a < b$ , Prove that  $a^2 \le ab < b^2$ !

The final test is as follows.

- 1. Let  $A = \{-3, -1, 2, 5, 7, 11\}$ . (a) Determine  $\sup A$  and prove it! (b) Is  $\inf A = -1$ ? Explain your answer!
- 2. Let  $A = \{x \in R | x^2 < 4\}$ . Determine sup A and prove it!
- 3. Let  $S \subseteq \mathbb{R}$  is no-empty set and bounded, k > 10, and  $kS = \{ks: s \in S\}$ , then prove  $\inf(kS) =$ k inf S.
- 4. Let  $S \subseteq \mathbb{R}$  is no empty set and bounded. Prove that if  $u = \sup S$ , then for each  $n \in \mathbb{N}$ ,  $u - \frac{1}{n}$  is not upper bound of S, but  $u + \frac{1}{n}$  is upper bound of S!

Both tests were administered to PYChat and PBL classes.

#### Validation Test (Validity Criteria)

The PYChat model received positive assessments from two experts. They considered that the model was highly adequate in several influential aspects. Both experts stated that the model aligned with Socratic Method learning theories and Real Analysis content. The PYChat model was aligned with the principle of questioning, as noted by two experts who asserted that presenting problems at each session encourages students to challenge assumptions, provokes thoughtful responses, and quides participants toward a deeper understanding of the topic under consideration. Moreover, the PYChat model also corresponded to the principles of dialectical inquiry and active engagement. They argued that tackling challenging problems in groups enabled students to examine and refine their beliefs through discussion, fostering active discourse within their groups. Furthermore, the PYChat model was consistent with the principles of reflection and critical thinking. It claimed that receiving answers from YouTube videos and ChatGPT can form new perspectives for students, prompting them to reflect on their beliefs and reasoning to analyze questions, and articulate responses from YouTube videos and ChatGPT to develop their answers. Finally, the PYChat model also aligned with the principle of open-ended. Two experts assert that the real analysis proof problems in the PYChat model offer numerous varied proof methods. Furthermore, the learning materials in the lesson plan and the provided problems were already following the specified learning outcomes. Each subtopic also contains correct and relevant content.

Furthermore, the two experts claimed that the PYChat module is valid in format, content, and language. In terms of format, the module was excellent in clarity of writing, appearance, and ease of access via QR code. In terms of content, the module was highly rated for the quality, clarity, and presentation sequence of the materials. In terms of language, the PYChat module did not create significant ambiguity and made it easy for students to understand them. The overall validity results can be seen in Table 2.

Table 2 Validation Results of the PYChat Model

Valida- tors	Learn- ing The- ories	Real Analysis Content	Mod- ule	Aver- age score	Criteria
Validator 1	80	92	72	81,34	Very Valid
Validator 2	84	92	84	86,67	Very Valid
Com- bined Valida- tors	82	92	78	84	Very Valid

Based on Table 2, the average score of the validity model by two experts was 84. The score means that the PYChat model was included in the very valid category. Therefore, the model did not require major revisions. Therefore, it could be administered in the field test. Then, it enhanced students' problem-solving ability theoretically. Thus, the PYChat model fulfilled very valid criteria.

## Field Test (Effectiveness and Practicality Criteria)

The PYChat model was conducted in two sessions namely the outside and the inside classroom. Outside the classroom, the students learned to understand the materials, to prove each Real Analysis theorem, and to solve problems to prove in groups. There were no descriptions or explanations in the module for all proofs of the theorems. Therefore, they sought assistance from ChatGPT, the module, and YouTube videos on the channel Jackson Mairing. Inside the classroom, each group presented and evaluated the solutions. The integration of YouTube videos and ChatGPT in real-analysis learning aimed to assist students in solving problems. The learning activities in PBL were like the PYChat model, but students did not have access to the aiding tools. They learn the material and the proof of each theorem from the PBL model. There were detailed explanations for all proofs of the theorems. The PBL model has been implemented in the Real Analysis course over the last three years. Therefore, both models had the same learning syntax. The difference was in the learning sources.

The students in PYChat and PBL classes solved the same worksheets, midtest, and final test. The average score of tasks in the worksheets was used as the assignment score. Scores of assignment, mid-test, and final test were compiled into scores representing the proving ability of the students with weights of 20%, 30%, and 50%, respectively. In addition, the researchers took the initial ability as the covariate from the students' final scores in the Abstract Algebra course. The course contained problems to prove as well as Real Analysis.

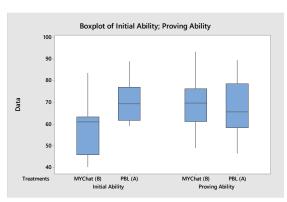


Figure 3. Boxplot of Initial and Proving Ability of the PYChat and PBL classes.

The results indicated that the initial ability of PYChat and PBL classes seemed to be different descriptively. The initial ability of the PYChat class seemed to be lower than those in the PBL class, as seen from the third quartile value in the PYChat class, which was not significantly different from the first quartile value in the PBL

class. However, there was no significant descriptive difference observed in proving ability scores. The range of the scores in the boxplot for the mid-test and final test in both classes appeared almost the same (Figure 3).

Similar results were indicated by the statistical summaries in both classes, where the average initial ability in the PY-Chat class (58.44) and PBL class (70.69) showed a significant difference (See Table 3). Additionally, the distribution of the initial ability in PYChat and PBL classes, respectively, could be observed in the following order: minimum values (40.25 and 59.00); Q1 (45.88 and 61.75); median (61.13 and 69.50); Q3 (63.38 and 77.10); and maximum values (83.75 and 89.00), each category demonstrating a considerable descriptive difference. However, the average score of the proving ability in PY-Chat (60.48) and PBL (68.89) classes showed a relatively small difference. Similarly, the distribution of the proving ability score in PYChat and PBL classes, respectively, could be seen in the following order: minimum values (49.00 and 46.50); Q1 (61.25 and 58.38); median (69.75 and 65.75); Q3 (76.38 and 78.36); and maximum values (93.50 and 89.50), with each category exhibiting only slight descriptive differences.

The difference in the proving ability between PYChat and PBL classes was analyzed using ANCOVA, with the proving ability as the dependent variable, the initial ability as a covariate, and the treatments as the independent variable. The analysis conducted to determine was to

Table 3 Statistical Summaries of the Initial and Proving Abilities in the PYChat and PBL Classes

Variable	Treatments	13	Mean	StDev	Min	Q1	Median	Q <sub>3</sub>	Max
Initial Ability	PYChat (B)	24	58,44	12,25	40,25	45 <b>,</b> 88	61,13	63,38	83,75
	PBL (A)	32	70 <b>,</b> 69	9,65	59,00	61,75	69,50	77,10	89,00
<b>Proving Ability</b>	PYChat (B)	24	69,48	10,86	49,00	61,25	69,75	76 <b>,</b> 38	93,50
	PBL (A)	32	68 <b>,</b> 89	12,71	46,50	58,38	65,75	78 <b>,</b> 63	89,50

prove whether there was a significant influence of the PYChat model on the proving ability. The presence of the influence served as an indicator of the model's effectiveness.

Before determining the difference, the normality test of the residuals would be examined. The Kolmogorov-Smirnov normality test yielded a, indicating that the residual data followed the normal distribution with a 95% confidence level. The second assumption, the homoscedasticity (equality of variances), was satisfied, as seen in the randomness of points around the fitted value and residual graph. The third assumption, the residual independence, was also met as it was observed in the randomness of points around the observation order and residual graph (Figure 4). The fourth assumption, the linearity, was fulfilled as indicated by the Lack-of-Fit, signifying that the model was linear with a 95% confidence level (Table 4). Therefore, the researchers could use AN-COVA to determine the effectiveness of the PYChat model.

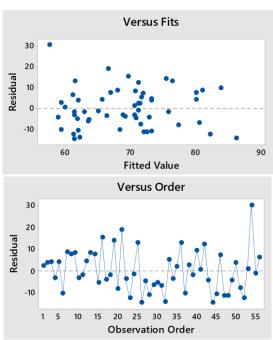


Figure 4 Normality test and Residual of Proving Ability

Table 4 ANCOVA Test for the Proving Ability

7	<b>D</b> E	A 1' CC	A 1' N4C	F-	P-	
Jource	DF	Adj SS	Adj MS	Value	Value	
Initial Ability	1	2759,7	2759,72	29,47	0,000	
Treat- ments	1	778,6	778,62	8,31	0,006	
Error	53	4963,4	93,65			
Lack-of- Fit	47	4650,3	98,94	1,90	0,215	
Pure Er- ror	6	313,1	52,19			
Total	55	7727,9				

The results of the ANCOVA showed that values for the initial ability and the treatments were and, respectively, indicating that the values are less than 0.05. Thus, there was a significant difference between the initial ability and the proving ability in PYChat and PBL classes with a 95% confidence level. Despite the descriptive statistics, such as mean, and distribution of the proving ability were relatively similar in both classes. The difference scores of the initial ability as a covariate in the classes led to a significant difference in the ability between PYChat and PBL classes.

The multiple comparisons test of the proving ability for both classes was further analyzed using the Tukey test. The result showed different letters between the two classes, indicating a significant difference in the proving ability between PYChat and PBL models. The average score of the proving ability in the PYChat class (74.10) was higher than in the PBL class (65.43). The result suggested that the proving ability in the PYChat class was higher than in the PBL class. Thus, the PY-Chat model has proven to be effective in enhancing students' proving ability. Hence, the model fulfilled the effectiveness criteria.

Table 5 Tukey's Multiple Comparisons for the Proving Ability

1 Toving Ability					
Treatments	N	Mean	Grouping		

PYChat (B) 24 74,0986 A
PBL (A) 32 65,4261

Means that do not share a letter are significantly different.

Furthermore, some solution examples from PYChat and PBL classes were compared. The activity aimed to compare different ways of proving from each class. In Figure 5, one of the proofs from the PBL class got the correct conclusion, although its reasoning was less precise. They started the proof with the appropriate definition of function. However, the counter-example was used to prove whether a function was not a member of the correct set. The choice of counter-example should take a member of the set from, not from. It happened because of the lack of experience in proving activity, hence, the students needed help or instructions to make the appropriate proof.

```
1. Misalkan himpunan A=B=\{x\in R\mid -1\leq x\leq 1\} dan himpunan C=\{(x,y)|\ x^2+y^2=1\} yang merupakan subset dari A\times B. Apakah himpunan C merupakan fungsi? Jelaskan! A=B=\{x\in R\mid -1\leq x\leq 1\} C=\{(x,y)|\ x^2+y^2=1\} Akan dibuktikan C merupakan fungsi f dikatakan fungsi dari A ke B, jika f\subseteq A\times B dan \forall\ a\in A ada tepat satu b\in B sehingga (a,b)\in f Karena C=\{(x,y)|x^2+y^2=1\}\subseteq A\times B. maka terbukti C\subseteq A\times B Pilih (0.0)\in A\times B, dimana 0\in A dan 0\in B Karena C\subseteq A\times B Maka, C\subseteq A\times
```

Figure 5. Example 1 of Answers to Exercises from PBL Class

Translating into English:

```
Let A = B = \{x \in \mathbb{R} | -1 \le x \le 1\} and C = \{(x,y)|x^2+y^2=1\} which is subset of A \times B. Is C a function? Explain! A = B = \{x \in \mathbb{R} | -1 \le x \le 1\}. C = \{(x,y)|x^2+y^2=1\}. Will be proven that C is a function. f is said to be a function from A to B if f \subseteq A \times B and \forall a \in A there is only one b \in B therefore (a,b) \in f. Since C = \{(x,y)|x^2+y^2=1\} \subseteq A \times B, then it is proved that C \subseteq A \times B. Choose (0,0) \in A \times B. Then x^2+y^2? 1 0^2+0^2? 1 0 \ne 1
```

Since for  $0 \in A$  and  $0 \in B$ , therefore  $(0,0) \notin f$ , then C is not a function.

Jika kita pilih elemen-elemen dalam himpunan A, yaitu x, untuk setiap nilai x dalam interval [-1, 1], maka kita perhatikan bahwa untuk setiap nilai x tersebut, ada dua pasangan nilai y dalam himpunan C yang memenuhi  $x^2+y^2=1$ . Ini karena lingkaran memiliki dua titik potong dengan garis vertical yang melewati pusat lingkaran. Sebagai contoh:

\*\*Jika x=0, maka terdapat dua nilai y yang memenuhi persamaan tersebut, yaitu y = 1 dan y

- = -1, - Jika x=1, maka terdapat dua nilai y yang memenuhi persamaan tersebut, yaitu y = 0 dan y = -1,
- . Jika x=-1, maka terdapat dua nilai y yang memenuhi persamaan tersebut, yaitu y = 0 dan y = 1.

Karena setiap elemen dalam himpunan A memiliki lebih dari satu pasangan nilai dalam himpunan C, himpunan C tidak memenuhi syarat utama fungsi, yaitu bahwa setiap elemen dalam domain harus memiliki tepat satu pasangan nilai dalam kodomain. Oleh karena itu, himpunan C bukanlah sebuah fungsi.

Dalam konteks ini, himpunan C lebih cocok digambarkan sebagai relasi yang menghubungkan setiap elemen dalam himpunan A dengan lebih dari satu elemen dalam himpunan B, sehingga tidak memenuhi definisi formal dari fungsi.

XXX

#### Translating into English:

If we choose elements from set A, namely x, for every value of x in the interval  $\{-1,1\}$ , then we observe that for each of the x values, there are two pairs of y values in set C that satisfy the equation  $x^2 + y^2 = 1$ . This is because a circle has two intersection points with a vertical line passing through center of the circle.

For example

If x=0, then there are two values of y that satisfy the given equation which are y=1 and y=-1. If x=1, then there are two values of y that satisfy the given equation which are y=0 and y=-1. If x=-1, then there are two values of y that satisfy the given equation which are y=0 and y=1. Since each element in set A has more than one pair of values in set C, it does not satisfy the primer condition of a function, which is that every element in the domain must have exactly one pair of values in the codomain. Therefore, set C is not a function. In this context, set C is more appropriately described as a relation that connects each element in set C with more than one element in set C, thus it is not meet the formal definition of function.

	ektikan apakah himpunan c adalah fungsi.
Diketahui	A=B= {UER -16U613
	C= 1(12,4) 122+ 42=13
	CGAXB
Menurut d	efinisi 1-3 Fungsi.
Relasi f dil	catakan fungsi dari Ake B, jika f C Ax B dan untuk
Setiap a e	A, ada tepat satu b e B schingga (a.b) ef.
Akan dibuk	tikan Jetapa EA, ada tepat satu 6 EB Jehingga (alb) ES
C= 102,4)1	W2+ y2=1 y
Artinya, \$	(a) = 12+42 = 1 2tau f(b) = 12+4==1
Piih DE	le , maka flo) = (0)2+ 12=1
	0+ y'= 1
	42 = 1-0
	7 ° · 1
	y - ± √I
Seri ngga	diferoleh (0,1) 2tau (0,-1) E C
Selanjutri	ya Pilih a=b
maka f	(a) = 1 (an (0,1)
đ	(b) = -1 dai (01-1)
6.1	f(a) + f(b), jadi hinpunan C bukannah fungsi.

Figure 6 Example 1 of Answers to Exercises from PY-Chat Class and ChatGPT

```
Translating Figure 6 into English:
Will be proven C is a function.
Let A = B = \{x \in \mathbb{R} | -1 \le x \le 1\}.
C = \{(x, y)|x^2 + y^2 = 1\}.
C \subseteq A \times B.
According to definition 1.3 function.
The relation f is a function from A to B, for every
a \in A, there is only one b \in B, therefore (a, b) \in f.
Will be proven for every a \in A, there is only one b \in
B therefore (a, b) \in f.
C = \{(x, y)|x^2 + y^2 = 1\}.
It means f(a) = x^2 + y^2 = 1 or f(b) = x^2 + y^2 = 1
y^2 = 1.
Choose 0 \in A, then f(0) = 0^2 + y^2 = 1
                                      v^2 = 1 - 0
                                      y^2 = 1
                                        y = \pm \sqrt{1}
Thus obtained (0,1) or (0,-1) \in C.
```

Furthermore, choose a = b. Then f(a) = 1 from (0,1)f(b) = -1 from (0,-1).

Therefore,  $f(a) \neq f(b)$ , so C is not a function.

In Figure 6, one of the proofs from the PYChat class was managed to prove it correctly. One of the assistants for the students to get the correct solution was from ChatGPT. It produced the correct answer, even though the reasoning or way of the proof was less precise. The ChatGPT's solution helped the students to identify that it was not a function. Then, they could improve it by starting the proof

from the definition of a function, giving a counter-example that supported it, and concluding that the equation was not a function.

```
5) Jika Ośałb. buktikan ażabłb²
karena Ośałb diperoleh affuloj, b-aff dan bep.
karena affuloj dan b-affuloj, b-aff dan bep.
alb-aj fpuloj
ab-a² fpuloj
artinga ażab.
karena b-aff dan bep diperoleh
(b-albep
b²-ab ep
artinya b² > ab
kaiena ażsab dan b²>ab, maka²a² sab < b².
```

Figure 7 Example of Answers to Mid-test from the PBL Class

Translating Figure 7 into English: If  $0 \le a < b$ , prove that  $a^2 \le ab < \frac{1}{4}$ . Since  $0 \le a < b$ , then it follows that  $a \in P \cup \{0\}$ ,  $b - a \in P$  and  $b \in P$ . Since  $a \in P \cup \{0\}$  and  $b - a \in P$ , it follows that  $a(b - a) \in P \cup \{0\}$   $ab - a^2 \in P \cup \{0\}$  It means  $a^2 \le ab$ . Since  $b - a \in P$  and  $b \in P$ , it follows that  $(b - a)b \in P$   $b^2 - ab \in P$ 

It means  $b^2 > ab$ . Since  $a^2 \le ab$  and  $b^2 > ab$ , then it is proved that  $a^2 \le ab < b^2$ .

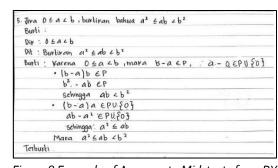


Figure 8 Example of Answers to Mid-tests from PY-Chat Class

Translating into English:

If  $0 \le a < b$ , prove that  $a^2 \le ab < b^2$ Proof: Given  $0 \le a < b$ . Asked: Prove that  $a^2 \le ab < b^2$ . Since  $0 \le a < b$ , then it follows that  $b - a \in P$ and  $a - 0 \in P \cup \{0\}$ ,  $(b - a)b \in P$   $b^2 - ab \in P$ It means  $ab < b^2$   $a(b - a) \in P \cup \{0\}$   $ab - a^2 \in P \cup \{0\}$ It means  $a^2 \le ab$ . Then  $a^2 \le ab < b^2$ . Proved.

In Figure 7 and Figure 8, one of the proofs from students in PBL and PYChat classes managed to prove it correctly. Proofs from the PBL class started from and then continued to prove. Both steps were well done and correct. On the other hand, the PYChat class solution started from the opposite order namely, and then continued to prove. It was also the right solution. In conclusion of the proving, both solutions of PBL and PYChat class were precisely accurate.

wa	ca buetkan bahwa ne(KS): K Inf S.
Jawal	
Mich	I inf S=y
	diburtiran bahwa ky abalah batas bawah dan KS
	Sebarang KS & MKS dimang K &R down S &S
	y adalah batus batusah dari s, mara untuk setiap seS balaku
	<u>y                                    </u>
	#H3 ≤ KS
ini be	arti ky addah botus bolwah dani KS
Schagut	nya, arcan diburtikan ky batas bawah terberar danis
	an x adalah batas bawah lain dari KS, makan untuk semua sES berh
	x cks
	<u>x</u> <u>c</u> s
arting	x x adulah batus bawah dan S
Karena	y = inf S, maka
	¥ ≤ y language and man farman
	X £ Eq

Figure 9 Example of Answers to Final tests from the PBL Class

Translating Figure 9 into English:

Let  $S \subseteq \mathbb{R}$  is nonempty set and bounded, k > 0 and  $kS = \{ks : s \in S\}$ , then prove that  $\inf(kS) = k\inf S$ .

Solution:

Let  $\inf S = y$ .

Will be proven that ky lower bound in kS.

Let any  $ks \in kS$  where  $k \in \mathbb{R}$  and  $s \in S$ .

Because y is lower bound of S, then for any  $s \in S$  follows that

$$y \le s$$
$$ky \le ks$$

It means that ky is lower bound of kS.

Next will be proven ky is the largest lower bound of kS, then for any  $s \in S$  follows that

$$x \le ks$$
 $x/k \le s$ 

It means that  $^{\chi}/_{k}$  is lower bound of S.

Since  $y = \inf S$ , then

$$x/k \le x$$

$$x \le ky$$

It means that ky is the largest bound of S. Then its proved that  $\inf(kS) = ky = k\inf S$ .

```
3. Misaltern S & R tak kerong from terbators, K > 0, down ks = {ks:se}} hasker bautilean bahwa tap (ks) - k tap S.

Jauriban.

Misaltern ta S = U

Afan fibilitern bahwa ku adaletah basha barah
Ambil Sembaran Ks & KS dingana Ke R down scS

karena u adalet tambar stas forman Ke R down scS

karena u adalet tambar stas forman Ke R down scS

karena u adalet stas forman star s. maka unterk setiap s & S berlak

S & ku

ks & ku

joni ber arti ku adaletah batas samas tam ks

Awan fibiliterna Ku bahar samas tam ks

Awan fibiliterna Ku bahar samas tam ks

Awan fibiliterna ku adaletah batas samas tam down ks, mandan
marka unterk semua s & S berlaku

ks & w

**E stase W

Fritzinga W adaletah batas samas dari s

kaman u = samp s

Marka v & W

Ini ber arti Ku adaletah batas samas tam bawah ter begar dari ks

Terbaki ing KS = Ku = K sup S
```

Figure 10 Example of Answers to Final tests from the PYChat class

Translating into English:

Let  $S \subseteq \mathbb{R}$  is nonempty set and bounded, k > 0 and  $kS = \{ks : s \in S\}$ , then prove that  $\inf(kS) = k \inf S$ 

Solution:

Let  $\inf S = u$ .

Will be proven that ky lower bound in kS.

Let any  $ks \in kS$  where  $k \in \mathbb{R}$  and  $s \in S$ .

Since u is lower bound of S, then for any  $s \in S$  follows that

$$u \le s$$
$$ku \le ks$$

It means that ku is lower bound of kS.

Then will be proven ku is the largest lower bound of kS.

Let w is another lower bound of S, then for any  $s \in S$  follows that

$$ks \le w$$
$$s \le {}^{W}/_{k}$$

It means that  ${}^{W}/_{k}$  is lower bound of S.

Since  $u = \inf S$ , then

$$w/k \le u$$

 $W \leq \kappa u$ 

It means that ku is the largest bound of S. Then its proved that  $\inf(kS) = ku = k \inf S$ .

In Figures 9 and 10, one of the proofs from PBL and PYChat class also managed to prove it correctly. PYChat and PBL class had similar solution. Both started from

 $\inf S = u$  and then taking any element of kS to prove that ku was the lower bound of kS. They continued to prove ku was  $\inf kS$  so they got conclusion that  $\inf kS = ku = k\inf S$ .

Based on the description above, the PYChat and PBL classes had similar solutions in the mid-test and the final test. Both classes proved deductively, accurately, and communicatively. On the other hand, when doing some tasks, the PYChat class had a slight advantage because they got assistance from ChatGPT regardless of whether the solution from ChatGPT was right or wrong. On the other hand, the students in the PBL class did not get the assistance.

The practicality of the PYChat model was evaluated through the field test involving assessing the appropriateness between the learning activities and syntax of the model and the success of all groups in solving the problems in the worksheets at each meeting. The research results indicated that the implementation of the PYChat model has well supported the syntax. While in the classroom, the students presented some solutions to the problems and discussed them. Then, the learning continued by giving some tasks to be solved outside of class sessions. The students worked in groups with support from ChatGPT and YouTube videos. In addition, each group routinely submitted the solutions before each meeting at LMS (Learning Management System). Figure 11 is evidence that all groups had successfully submitted the solutions.

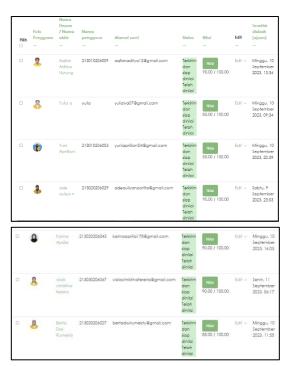


Figure 11 Evidence of All Groups had Successfully Submitted the Task Solutions

#### Discussion

The success of implementing the model was analyzed based on three main criteria namely validity, effectiveness, and practicality (Nieveen & Folmer, 2013). Those observed criteria could be at the stages of developmental research, specifically in the expert validation and the field test. Based on the presentation of the validity results, the PYChat model already met the required standards for quality of (1) suitability to learning model Socratic Method, (2) suitability to Real Analysis content, and (3) PYChat module format, content, and language were valid before being used in learning. Also, the effectiveness of the PYChat model was elevated by the results gained in the field test. Moreover, comparing the solutions between PY-Chat and PBL indicated that the model tends to contribute positively to the learning process. The result aligned with previous research that stated ChatGPT became a highly-practical tool in providing assistance or guidance when solving problems (Halaweh, 2023; Tlili et al., 2023). Other research also suggested that the PBL model and YouTube videos influenced students' problem-solving abilities (Mairing, 2020a, 2021). The effectiveness of PYChat as a learning model was in line with Fidan and Tuncel (2019), who stated that PBL would be more effective on students' learning outcomes or problem-solving ability when assisted by media. In addition to being a valid and effective model, the PYChat model has also proven practical in its implementation. The evidence was from the appropriateness of PYChat syntax with the classroom activities during the field test.

The success of the PYChat model was substantiated by the transformation observed in students' comprehension of proof-related matters. Previously, they were grappling with abstract concepts, but they now exhibit competence in effectively addressing abstract proof problems (Saraswati et al., 2021). Students struggled to initiate the proof process, lacking critical, logical, creative, and systematic thinking skills, along with difficulty expressing ideas nonverbally (Widiati & Stephani, 2018). This struggle was attributed to their limited exposure to resolving abstract problems (Mairing, 2020a). These students were used to solving concrete mathematical problems rather than abstract problems before entering college. By addressing the issue, the PYChat model was proven effective in solving their difficulties based on presentation. PYChat also helped them learn how to prove deductively and logically by giving them clues even if the clues were not perfectly correct. Minggi et al. (2021) suggested providing examples of proofs or step-by-step proof assistance helps inexperienced students in solving mathematical proof problems. However, it indirectly encouraged students to engage in critical thinking through the analysis and evaluation of responses provided by

ChatGPT. Students needed to demonstrate more critical thinking in evaluating information, fostering creativity and innovation, as well as developing and presenting new ideas through what was generated by ChatGPT (Guo & Lee, 2023; Marbun, 2023; van den Berg & du Plessis, 2023).

The PYChat learning model is also convenient and practical. This evidence became the findings of this research concerning the practicality aspect of the model. Moreover, the model is aligned with contemporary trends, as students are closely associated with technology and AI, particularly ChatGPT. A model closely related to technology will inherently motivate students, thus aiding in their cognitive development. In addition to being practical in enhancing students' proving abilities, this model is also practical, engaging, and motivating (Dewi et al., 2020; Hasanah et al., 2023).

The PYChat model was also in line with the principles of the Socratic method of learning theory, namely (1) questioning, (2) dialectical inquiry, (3) reflection, (4) active engagement, (5) critical thinking, and (6) open-ended (Delić & Bećirović, 2016; Rahmawati & Suwarjo, 2021). By the first principle, questioning, instead of directly providing students with material, they learned by solving given problems. Furthermore, this model also facilitated dialogue among group members in problem-solving, demanding active involvement from students in learning the material, which aligned with the principles of dialectical inquiry and active engagement. Subsequently, with the assistance of YouTube videos and ChatGPT responses, students analyzed and evaluated the information they obtained from these sources, consistent with the principles of reflection and critical thinking. Proof problems inherently offer various proofs,

and each group or individual will inevitably have different proofs stemming from discussions and information gleaned from YouTube videos and ChatGPT responses, reflecting the principle of open-ended.

#### Implication of Research

Educators should embrace learning models integrated with AI, such as ChatGPT, and incorporate them into their learning. By implementing AI technology in education, teachers can provide students with a more dynamic and personalized learning experience. It can help improve the efficiency of learning and create new opportunities to introduce innovative learning models. Therefore, the collaboration between teachers and AI could advance the education sector to a higher level, preparing students for a future increasingly connected to technology.

#### Limitation

The limitation of this research is the effectiveness of the PYChat model for students with good cognitive, affective, and conative abilities, but not for students who are lacking. This limitation happened because the model requires students to be critical in learning and finding their own answers. It also requires good analysis to process information from ChatGPT's responses. This constraint was overcome by grouping all students heterogeneously, meaning each group member includes students who have excellent and poor abilities. However, students with low motivation and abilities still struggle to follow the flow of this learning model. On the other hand, researchers cannot monitor each student individually to ensure they can join the learning process effectively.

#### CONCLUSION

The PYChat learning model was proven valid based on the average score of validity model by two experts was 84. It means that the PYChat model is included in the very valid category. The research findings indicated that the PYChat model had an initial ability score of 58.44 and a proving ability score of 69.48, whereas PBL scored 69.13 and 68.89, respectively. Using AN-COVA and the multiple comparison test, there were significant differences between PYChat and PBL classes in the initial and the proving abilities. Furthermore, the proving ability of students in the PYChat class was significantly better than in the PBL class. Therefore, the PY-Chat model effectively enhanced the students' proving ability. The effectiveness of the PYChat model could also be observed by comparing solutions when the students were also working on some problems. The PYChat model made a positive contribution to the learning process. Some of the less accurate solutions from the PBL class were due to a lack of guidance when solving the problems, and it can be assisted by ChatGPT and YouTube videos to get the correct solutions. In addition, the PYChat model has proven practical in its implementation. The evidence was from the PYChat syntax in line with the classroom activities during the field test.



### 2% Overall Similarity

Top sources found in the following databases:

- 1% Internet database
- Crossref database
- 0% Submitted Works database

- · 2% Publications database
- Crossref Posted Content database

#### **TOP SOURCES**

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Advanced Mathematics 3, 1985.  Crossref	<1%
2	repository.uir.ac.id Internet	<1%
3	researchgate.net Internet	<1%
4	Helena Roquet, Alexandra Vraciu, Florentina Nicolás-Conesa, Carmen  Crossref	<1%
5	Zhou Yan. "QoS-based space-frequency prefiltering for TDD MC-CDMA Crossref	<1%
6	coursehero.com Internet	<1%
7	iosrjournals.org Internet	<1%
8	Wangwei Xuan. "GDASC: Practical teaching mode structure for IPE cou  Crossref	<1%



9	imo.universite-paris-saclay.fr Internet	<1%
10	"Advanced Intelligent Computing Technology and Applications", Spring  Crossref	<1%
11	Arini Aha Pekuwali. "Prediction of student learning outcomes using the  Crossref	·<1%
12	Susilahudin Putrawangsa, Muhammad Syawahid. "Model Kegiatan Mic Crossref	<1%
13	mdpi.com Internet	<1%
14	Eric D'Hoker, Michael B. Green. "Identities between modular graph for  Crossref	<1%
15	repository.unp.ac.id Internet	<1%