

Research Article

Cite this article: Utami, N., Setiawan, A., Hamidah, I., & Koehler, T. (2025). Investigating Problem-Based Worksheets (PBWs) to Improve Understanding in Logic Gates Topic: Stacking and Racking Analyses in Rasch Model. *Educational Process: International Journal*, 16, e2025243. <https://doi.org/10.22521/edupij.2025.16.243>

Received April 8, 2025

Accepted May 6, 2025

Published Online June 10, 2025

Keywords: Conceptual understanding, problem-based learning, worksheet, Rasch analysis, vocational education

Author for correspondence:

Agus Setiawan

✉ agus_setiawan@upi.edu

✉ Universitas Pendidikan Indonesia, Indonesia



OPEN ACCESS

© The Author(s), 2025. This is an Open Access article, distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction, provided the original article is properly cited.

Investigating Problem-Based Worksheets (PBWs) to Improve Understanding in Logic Gates Topic: Stacking and Racking Analyses in Rasch Model

Nisaudzakiah Utami[✉], Agus Setiawan[✉], Ida Hamidah[✉], Thomas Koehler[✉]

Abstract

Background/purpose. This study investigated the effectiveness of Problem-Based Worksheets (PBWs) in improving conceptual understanding of logic gates. PBWs were designed with key characteristics that integrate authentic problem-solving, active student engagement, and conceptual scaffolding to support critical thinking processes. A total of 32 vocational high school (VHS) students in Indonesia were involved in the PBWs class.

Materials/methods. This study employed a quantitative quasi-experimental one-group pretest-posttest design to examine the effects of PBWs on students' conceptual understanding. The intervention was conducted for 90 minutes, and 90 minutes for pre- and post-tests. Data were analyzed using stacking and racking within the Rasch model to measure changes in students' understanding and item characteristics.

Results. The results significantly improved students' conceptual understanding of logic gates after the PBWs intervention. Due to non-normal data distribution and small sample size ($n = 32$), Mann-Whitney U and Wilcoxon tests were used, both confirming statistically significant score increases ($p < 0.05$). Rasch analysis further supported this improvement, with a logit difference of -7.14 between pre- and post-tests. Racking analysis revealed that item difficulty shifted, particularly for items 3, 5, 6, and 8. Despite overall gains, anomalies such as guessing, cheating, and carelessness influenced some students' results, as detected through response patterns and scalogram analysis.

Conclusion. This study concluded that PBWs effectively improved students' conceptual understanding of logic gates in pneumatic. PBWs offered a more engaging and holistic learning experience. These findings highlight the need for consideration of student behaviour during testing to ensure accurate interpretation of educational outcomes.

1. Introduction

Problem-based learning (PBL) is a pedagogical method that fosters complex skills, such as critical thinking, teamwork, and creative problem-solving skills (Das et al., 2023; Maidan et al., 2020; Aritia & Suyanto, 2019). In vocational education, conceptual understanding is fundamental to mastering complex skills. Studies have shown that students' comprehension significantly influences the ability to make informed decisions and apply knowledge effectively in practice (Cano et al., 2014; Pascual & Goikoetxea, 2014). Therefore, implementing innovative instructional models that improve students' understanding is crucial for fostering learning outcomes. Studies have also emphasized the importance of conceptual mastery in vocational education. According to (Shepherd et al., 2022; Corlatescu et al., 2023), improving understanding through effective learning models helps students make informed decisions.

Pneumatics has been revealed to be an essential engineering education concept and is widely used in different industries (Saravanakumar et al., 2017). A big problem with pneumatic learning is understanding logic gates, and when students do not grasp the concept well, they may find it hard to design circuit diagrams. The factors that make learning difficult include (1) the absence of hands-on skills in utilizing theoretical knowledge for real-world applications (Raja Subramanian et al., 2024), (2) traditional learning methods that are not interactive, hence students find it difficult to connect logical concepts with pneumatic systems (Yeni et al., 2024), (3) limited interactive learning media, which could facilitate the exploration of logic and pneumatics concepts in practical scenarios, and (4) limited simulation tools and workstations that replicate industrial conditions more realistically. The main challenge in pneumatic learning is the variation in students' cognitive abilities when grasping the fundamental concepts of pneumatic systems. Learning logic gates in pneumatics is very important, especially for those who are studying automation engineering and system control. However, one big issue is that many students do not understand how logic gates work or how they are used in pneumatic systems (Dai et al., 2021). Even though evaluating students' understanding of these concepts is important, there are limited tests to track their progress. Therefore, teachers cannot properly evaluate progress or find out where students need help.

This study aimed to investigate the effectiveness of PBWs in improving students' conceptual understanding of logic gates using stacking and racking analyses in the Rasch model. To achieve this objective, the following questions were formulated:

1. Is there a significant change in students' conceptual understanding after learning with PBWs?
2. How does conceptual understanding change through PBWs intervention compared to other methods?
3. In addition to interventions, what other factors contribute to changes in students' conceptual understanding?

This study is important because it helps create better ways for students to learn about logic gates using PBWs. Moreover, it examined how PBWs enhanced students' engagement and conceptual understanding in learning logic gates. This study also provided teachers with strategies for incorporating PBWs into lessons to create a practical learning experience. A review of previous studies, methodology, findings, and discussions was conducted, followed by conclusions and recommendations.

2. Literature review

PBL is widely known as a good teaching method in vocational education. Previous studies have shown that it helps students think critically, solve problems, and work together, which are essential in the modern workforce. According to (Hung & Amida, 2020), PBL is all about learning using

structured problem scenarios that stimulate inquiry and knowledge application. Vocational education helps students gain practical skills and understand concepts better by working on real-world problems (Jaganathan, S., Bhuminathan, S., Ramesh, 2024). However, issues like different cognitive abilities and difficulties in conceptual understanding present the need for further investigation into optimized learning strategies (Syahmani et al., 2020; Samsudin et al., 2019).

Understanding concepts is very important in technical education because it helps students use theory in real tasks. As stated by Jaganathan, Bhuminathan and Ramesh (2024), a good understanding helps students to use knowledge in different situations. In engineering and vocational education, inadequate understanding of concepts makes fixing problems and creating good systems difficult. Also, knowing key concepts in vocational education helps solve difficult engineering problems (Hermundstad et al., 2016).

In pneumatic systems, understanding logic gates is important for making circuit diagrams and following industry rules. However, many students find these concepts hard because they are too theoretical and lack hands-on practice (Fries et al., 2021; Hermundstad et al., 2016). Better teaching methods should mix theory with practical work to help students learn. Logic gates work based on truth tables that show how inputs affect outputs. Furthermore, they are key parts of digital systems, helping in computing and automation (Bouchée et al., 2022). Learning helps students understand digital technology in electronics and engineering (Çakıroğlu et al., 2022; Yıldırım & Canpolat, 2019). Logic gates can also be used in pneumatic systems to control mechanical functions based on logic (Rahman et al., 2022). Even though they are important in circuit design, students often have trouble understanding them. This study focused on AND, OR, and NOT gates, which use binary inputs to create outputs based on logic rules (Küçük-Avcı et al., 2024). As stated in (Marhaeni & Andriyani, 2023; Lugosi & Uribe, 2022), students struggle with logic gates because they are too abstract. Hence, they need more active learning.

Old teaching methods do not always help, and new ones like PBWs are needed. Table 1 shows traditional methods compared to PBWs.

Table 1. Comparison of Conventional Worksheet and PBWs

Aspect	Conventional Worksheet	Problem-Based Worksheets
Learning Approach	Direct instruction is often lecture-based (Küçük-Avcı et al., 2024)	Student-centered, focusing on problem-solving and critical thinking (Marhaeni & Andriyani, 2023)
Students' Engagement	Lower engagement, passive learning (Lugosi & Uribe, 2022)	Higher engagement and active learning through real-world problems (Utami et al., 2022)
Collaboration	Limited opportunities for collaboration (Utami et al., 2022)	Encourages teamwork and collaboration (Yusop & Firdaus, 2022)
Skill Development	Focus on memorization and basic understanding (Lugosi & Uribe, 2022)	Fosters problem-solving, critical thinking, and creative thinking skills (Marhaeni & Andriyani, 2023)
Instructional Design	Easier to design and implement (Utami et al., 2022)	Requires more time and effort to design and implement
Effectiveness	Effective for basic knowledge acquisition (Shana & El Shareef, 2022)	More effective for deeper understanding and application of knowledge (Sipayung, 2019)

Logic gates combined with PBWs offer a dynamic method of teaching and learning about pneumatic systems. PBWs present students with real-world scenarios that require the application of logic gates to solve engineering problems. Students learn to design and troubleshoot pneumatic systems by engaging in practical exercises while incorporating logic gates concepts. Studies have shown that hands-on problem-solving experiences in engineering education increase knowledge retention and understanding of theoretical concepts. PBWs incorporating logic gates enable the application of logical reasoning in a structured educational environment, preparing the students for practical challenges in industrial and technical fields.

PBW have emerged as an innovative instructional tool that integrates PBL principles into structured learning materials, facilitating active learning by guiding students through problem identification (J. Lu et al., 2014), data analysis (Hung & Amida, 2020), and solution development (Devika et al., 2024). Unlike conventional worksheets, PBWs foster engagement and deep learning through systematic, inquiry-based exploration (Yeni et al., 2024). Although several studies have shown that PBWs significantly improve conceptual understanding, particularly in STEM education, only few focus on PBWs in the context of pneumatic learning, emphasizing the need for further empirical validation.

Understanding concepts demands an integration of theoretical knowledge with practical application. When working with logic gates, specifically in the framework of PBWs, students develop an understanding of how logical principles control real-world systems. Therefore, applying logic gate concepts in pneumatic systems connects abstract logical operations to tangible actions and outcomes. Conceptual understanding is specifically crucial in engineering and technology, where theoretical principles are often applied to practical situations. For example, mastering the concept of logic gates involves understanding their function and their application in controlling systems such as pneumatic machines. Strong conceptual understanding also promotes critical thinking, particularly when solving complex problems. Although existing literature emphasizes the effectiveness of PBL, PBWs, and Rasch analysis in improving conceptual understanding, investigation on their combined application in pneumatic learning remains limited. Most previous studies have focused on general STEM subjects, with little emphasis on vocational education contexts (Hmelo-Silver & Barrows, 2008). Furthermore, the assessment of conceptual understanding in pneumatic systems has not been extensively explored, creating a gap this study aims to address. By integrating PBWs with Rasch model analysis, the aim was to provide empirical evidence of their effectiveness in improving understanding of logic gates in pneumatics.

Rasch analysis serves as an effective method to objectively and comprehensively measure conceptual understanding (S. Lu & Bi, 2016). It is a statistical method for analyzing assessment data and accurately evaluating learning outcomes (Abu Bakar et al., 2023). This model provides a detailed measurement of students performance by considering both item difficulty and individual ability levels. In vocational education, Rasch analysis has been applied to assess students' conceptual understanding and identify areas requiring targeted intervention (Boone et al., 2014). Studies have shown that it offers a more nuanced evaluation of students' progress compared to traditional assessment methods (Sanusi et al., 2025). Furthermore, the application in pneumatic learning is particularly beneficial for diagnosing cognitive strengths and weaknesses, enabling teachers to tailor instructional strategies accordingly (Linacre, 1994). Rasch analysis also provides accurate assessments of learning and improves instructional decision-making. Based on the description, this study aimed to contribute to the existing body of knowledge by examining the impact of PBWs on conceptual understanding in pneumatic systems, addressing a critical study gap in vocational education.

3. Methodology

This study used a quantitative design with a quasi-experimental one-group pretest-posttest design to examine the effects of PBWs. The experimental process was conducted over three sessions, each lasting 90 minutes. Subsequently, pre-test and post-test assessments were administered, with each taking 45 minutes. This study involved pre-tests, PBWs implementation, post-tests, and data collection on intervention outcomes. Changes in students' understanding and question items were subjected to stacking and racking analyses in the Rasch model using Ministep 5.6.4.0 software. As standard techniques, racking and stacking were introduced by Benjamin Wright to measure changes in students' conceptual understanding and items before and after the intervention (Anselmi et al., 2015). Students' understanding was measured to identify those with significant conceptual understanding changes in response to the learning intervention. In addition, item analysis was conducted to determine items with unique characteristics that were specifically understood differently during the intervention.

3.1. Participants

Participants in this study were selected using a purposive sampling technique and involved 32 11th-grade VHS students, with a composition of 76.5% male and 23.5% female. Participants were selected based on their developmental stage in pneumatic learning and previous experience with the PBWs learning environment.

The inclusion criteria in this study were 11th-grade students who had participated in basic pneumatic learning and had learning experience using the PBWs environment. Meanwhile, the exclusion criteria included students who did not complete all learning sessions or did not provide consent to participate. Of the total 45 students who met the inclusion criteria, 32 were willing to participate and complete the entire series of research activities.

This study was conducted at SMK Cimahi, Indonesia, which was selected because it has advanced infrastructure, a supportive educational environment, a collaborative culture, and good accessibility. These factors make the school an ideal location to investigate the role of PBWs in logic gate learning. Consent to participate was obtained from the students before the study.

3.2. Instrument

Conceptual understanding was measured using several indicators based on Bloom's taxonomy, specifically focusing on cognitive levels such as understanding, applying, and analyzing. The assessment instrument comprised 8 test items. Three professionals with backgrounds in project-based learning and engineering education have validated the content of the instrument used in this study. Regarding the questions' alignment with the learning objectives and the context of the logic gate material, the experts offered their opinions. Additionally, ten students from courses not included in the research sample participated in an early trial (pilot test) to assess the instrument's consistency, clarity, and understandability. The Cronbach's Alpha reliability coefficient showed a high degree of dependability, which was used to examine the pilot test findings. Its value was 0.82. This process seeks to guarantee that the tool utilized to measure efficacy is legitimate and trustworthy. This procedure aims to ensure that the instruments used are valid and reliable in measuring the effectiveness of PBWs.

3.3. Experiment

The study process was conducted in three stages, namely Pre-test, Implementation, and Post-test. During pre-test stage, one class of second-year students from School was selected for pre-testing. A specialized "Test to Measure Students' Conceptual Understanding in Pneumatics" was designed and administered. A total of 8 test items were completed, and responses were collected for

analysis. In particular, teaching strategies and content were standardized across the groups. The main teaching process included: (i) Introducing the problem (understanding the problem, characterizing the problem, and exploring possible solutions), (ii) Group investigation (applying strategies), (v) Analysis and Evaluation of Problem-solving process (evaluation and validation). During the implementation stage, instructional materials were provided, including modules, laptops equipped with Fluidsim software, and PBWs. Each experimental group worked through worksheets, integrating problem-solving tasks. Subsequently, students designed simulations using Fluidsim and presented results to teachers. The final phase of implementation involved a practicum on a pneumatic board, which students could conduct only after receiving approval from teachers based on presentation results. During the post-test stage, students in the PBWs Class completed the "Test to Measure Students' Conceptual Understanding in Pneumatics" to assess the impact of PBWs on learning. Data were also collected on students' perceptions of conceptual understanding improvements. A total of 34 students participated in the post-experiment survey. The pre-test and post-test contained identical items, and 90 minutes were given to complete the assessment.

3.4. Data Analysis

Rasch model is a psychometric tool used in educational assessment to measure latent traits such as students' ability and comprehension. Unlike traditional assessment methods that only show correct or incorrect answers, the Rasch model evaluates both item difficulty and test-takers ability on a general scale, providing a more nuanced perception of students understanding (Engelhard, 2013). This analysis determines the extent of comprehension by relating responses to item difficulty levels (Soeharto & Csapó, 2021). Rasch model, based on probabilistic logic, allows for precise measurement by relating students' ability with the complexity of items on the same scale (Stolt et al., 2022). When applied to PBWs involving logic gates, this model could effectively assess the ability to understand and apply concepts by calibrating responses across varying levels of problem difficulty.

Rasch model provided several advantages in measuring conceptual understanding. First, it produced standardized results, enabling reliable comparisons between different groups or populations. Second, the model considered the relationship between students' abilities and item characteristics, offering a more comprehensive evaluation. Third, it detected invalid items, improving the accuracy of measurement instruments.

Regarding operationalization, initially in ordinal form, pre-test and post-test data were converted into interval data using Rasch Partial Credit Model in WINSTEPS 4.5.5 software (Abu Bakar et al., 2023). This transformation ensured that both students ability levels and item difficulty were measured on a consistent logit scale. Two Rasch analysis methods were used to access changes before and after PBWs intervention. Stacking analysis arranged pre-test and post-test data vertically, where each item appeared once across experimental and control groups. This allowed for a direct comparison of students performance before and after the intervention. Therefore, each student received two ability measurements (pre-test and post-test), with each measured once. Meanwhile, racking analysis arranged pre-test and post-test data horizontally, meaning each item appeared twice in the dataset, with each ability recorded only once. The method helped analyze how the intervention affected item difficulty and students' performance, emphasizing shifts in conceptual understanding. With the application of "stacking" (measuring changes in students' ability) and "racking" (measuring changes in item difficulty), the Rasch model provided a detailed evaluation of students' conceptual understanding and problem-solving development (Boone et al., 2014).

4. Results

4.1. The Difference in Students Understanding in Pre-test and Post-test

Based on the findings of the data normalcy test, which indicated that the data distribution did not satisfy the assumption of normalcy, the Mann–Whitney U and Wilcoxon tests were used in this investigation. The data were not normally distributed, as shown by the significance value below 0.05 obtained from the Shapiro-Wilk test applied to the pretest and post-test data. As a result, we decided to employ a non-parametric test that has greater advantages for data with a non-normal distribution. Additionally, the comparatively small sample size ($n = 32$) supports the justification for employing this method in data analysis.

Table 2 shows a pre-test comprised 32 data points with a Mean Rank of 0.56, while the post-test had 32 data points with a Mean Rank of 0.70. Post-test scores were significantly higher than pre-test scores, confirming an improvement in students' performance after the intervention. Mann-Whitney test results ($U = 46,500$, $p < 0.05$) showed a statistically significant difference between both scores. Since the p -value was less than 0.05, the null hypothesis (H_0) was rejected, confirming the effectiveness of the intervention. Wilcoxon test results ($Z = -6.296$, $p < 0.05$) further supported a significant difference between the two scores. This confirmed that the intervention significantly improved students' conceptual understanding of logic gates. Therefore, the learning process using PBWs was more effective than conventional worksheets in improving conceptual understanding.

Table 2. Mann-Whitney and Wilcoxon tests

Group	Wilcoxon W	Z	U	P
Pre-test and Post-test	574.500	-6.296	46.500	.000

4.2. Changes in students' conceptual understanding of logic gates topic

Table 3 shows the analysis of pre-test and post-test differences, further confirming the improvement. Mean logit for pre-test was 3.57 (S.E. logit .25), while post-test score was -3.57 (S.E. logit .25). The difference between both tests was -7.14, confirming a significant increase in conceptual understanding of logic gates after the intervention.

Table 3. Logit of the mean of pre- and post-test items of experimental groups

Student	Item	Mean/SE (logit)		
		Pre-test	Post-test	Pre-test and Post-test Difference
32	8	3.57/.25	-3.57/.25	-7.14
Description SE = Standard Error				

Figure 1 shows a plotting pre-test and post-test results in pairs, the difference in mean logit score between the two (logit -7.14) appeared as an intercept on the horizontal axis, with a plotted slope = 1. The data points below the curve showed there was no significant change between the first measurement (Pf1.txt) and the second (Pf2.txt). Student 15 was positioned below this line, confirming little to no change between the two assessments. Students positioned between the two curves but close to the identity line experienced some changes, but the differences were not highly significant. The majority, except student 15, confirmed the measurement results remained relatively stable between Pf1 and Pf2 with minor changes.

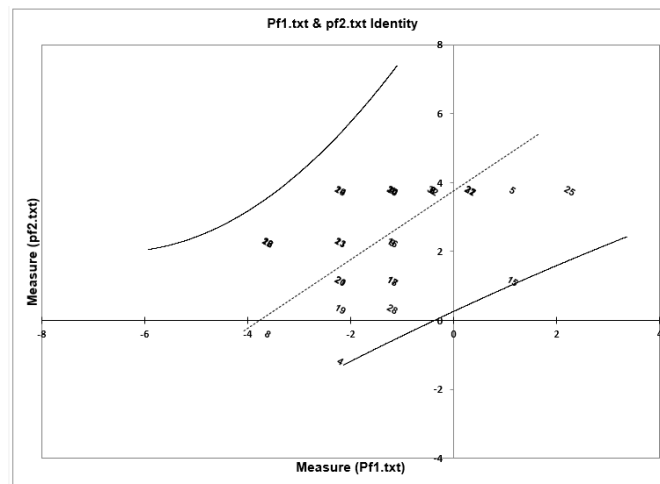


Figure 1. Stacking Plot of Person Measures for Time 1 against Time 2 for the Experimental Group

4.3. The Changes in Item Difficulty Level

Table 4 presents racking analysis results, examining changes in item difficulty between pre-test and post-test in the experimental group. Mean item difficulty in pre-test was logit 0.56, while post-test mean was logit 0.70, resulting in mean difference of logit 0.14. The results showed that items 3, 5, 6, and 8 experienced a significant increase in difficulty, more than the mean difference (logit 0.14). Meanwhile, items 1, 2, 4, and 7 showed less change in difficulty than the mean difference. The most difficult in the experimental group was item 8 (logit 4.49), while the easiest was item 1 (logit -2.91). These showed the variations in item difficulty before and after students were taught using PBWs.

Table 4. Racking Analysis

Item	Pre-test	Post-test	Difference between Pre-test and Post-test
Item 1	1.01	-1.90	-2.91
Item 2	-.24	-1.03	-0.79
Item 3	-1.78	.00	1.78
Item 4	.45	.38	-0.07
Item 5	-1.97	-.45	1.52
Item 6	.20	.72	0.52
Item 7	2.32	2.28	-0.04
Item 8	4.49	5.60	1.11
Mean	.56	.70	0.14

4.4. Conceptual Understanding changes in students' ability and item difficulty levels

In addition to the effects of learning interventions, three other factors could influence changes in students' abilities and the difficulty level of question items: 1) Guessing (i.e., a lucky guess that happens to be correct), 2) Cheating, and 3) Carelessness. These factors could be identified through students' response patterns using a scalogram analysis. For instance, student 18 (Table 5) was unable to correctly answer items 1 (logit -1.90) and 3 (logit 0.00) but correctly explained a more difficult item 7 (logit 2.28). This confirmed a lucky guess, as post-test ability level was higher than expected based on the difficulty of item 7. Cheating was confirmed in the response patterns of students 24, 30, 10, 3, and 7. This was initially detected from the identical pre-test and post-test means (logit -1.27 and -

0.40, as well as logit 3.74 and 3.74). The final response pattern showed carelessness. For example, students 26, 18, and 20 were considered careless, failing to correctly answer an easier question (item 3, logit 0.00) but successfully answered the more difficult item 7 (logit 2.28).

Table 5. Salogram result of several students

Person	Item	Mean Pre-test	Mean Post-test	Pre-post different	Item Response Pattern
	+ 1 2 5 3 4 6 7 8				
26	+1 1 1 0 1 1 1 0	-3.87	2.27	6.14	Careless
18	+0 1 1 0 1 1 1 0	-1.27	1.16	2.43	Careless, Guessing
20	+1 1 1 0 0 1 1 0	-2.40	1.16	3.56	Careless
24	+1 1 1 1 1 1 1 0	-1.27	3.74	5.01	
30	+1 1 1 1 1 1 1 0	-1.27	3.74	5.01	Cheating
10	+1 1 1 1 1 1 1 0	-1.27	3.74	5.01	
3	+1 1 1 1 1 1 1 1	-.40	3.74	4.14	Cheating
7	+1 1 1 1 1 1 1 1	-.40	3.74	4.14	

4.5. Negative Change

Negative changes in conceptual comprehension were detected when post-test logits were lower than pre-test logits. Student 15's ability declined after the intervention (Table 6). The pre-test and post-test had mean logits of 1.26 and 1.16, respectively, with a difference of -0.10.

Table 6. Scalogram result of student 15

ID person	Test	Item Response Pattern	Mean			
			Item Logit	S.E* Logit	Pre-test and Post- test difference	Combined S.E
15	Pre-test	+11101100	1.26	.97	-.10	1.93
	Post-test	+11111000	1.16	.96		

5. Discussion

The results showed a change in students' conceptual understanding of logic gates after the implementation of PBWs. Based on the difference between the two test means, the post-test group experienced a positive change (Ferreira et al., 2013; Sipayung, 2019; Sritun et al., 2020; Dotimineli & Mawardi, 2021; Yerizon et al., 2022; Haruehansawasin & Kiattikomol, 2018). In addition to the effect of the intervention, other factors contributed to changes in conceptual understanding. These included students' ability, the difficulty level of test items, and negative influences such as carelessness and misconceptions (Kyriakopoulou & Skopeliti, 2024; Lin et al., 2016; Yazbec et al., 2019). However, not all learning objectives related to logic gates were in accordance with PBWs. Negative changes occurred when students were not fully epistemological in the learning process, particularly at the observation and analysis stages, where the working principles of logic gates in pneumatic systems were examined. These activities, conducted in groups, were essential for concept formation based on real-world problems. The use of PBWs was expected to improve students' involvement, facilitating the explanation and connection of concepts related to pneumatic logic gates based on individual reasoning. According to the validity analysis using Rasch model, the instrument

showed strong validity. As stated in (da Silva et al., 2019; Jansen et al., 2024), a valid test item ensures that the collected data accurately represents learning outcomes.

Changes in conceptual understanding were observed in student 15 (Table 8), whose pre-test logit was 1.26, but declined to 1.16 in the post-test, confirming a decrease in performance. Similar changes were noted in question 8, which had a pre-test logit score of 4.49 and a post-test logit score of 5.6, confirming an increase in difficulty. Despite the higher post-test difficulty, students' scores improved. Studies have shown that the distribution of correct answers influences perceptions of test difficulty. For example, when easier items are answered correctly more often, logit measure may increase, reflecting a higher level of ability. Conversely, when correct answers are more evenly distributed across different difficulty levels, logit measure may reflect a different outcome (Tavakol et al., 2024; Wallmark et al., 2023). The increase in correct answers for item 8, corresponding to cognitive level C4 (applying), did not reduce difficulty evenly. PBWs significantly improved performance on questions 1- 7, but item 8 remained particularly challenging. Item 8 required students to design circuit diagrams based on real-world problems, integrating three logic gates, namely AND, OR, and NOT. Most students only implemented two of the three gates, or vice versa. Only 9 out of 32 students were able to complete the task successfully after the intervention, compared to none before the intervention. However, for the other items, PBWs led to significant improvements.

Based on analysis, students number 26, 18, 20, 24, 30, 10, 3, and 7 experienced negative changes (Table 7), with answer patterns not following the expected hierarchy of question difficulty. Besides the impact of the learning intervention, other external factors influenced students' test results, including correct guesses, cheating and carelessness. For instance, student 18 could not answer easy questions but managed to answer more difficult ones, confirming the influence of a "lucky guess" factor. In addition, indications of cheating were identified based on similar patterns among students 24, 30, 10, and 3, whose pre-test and post-test means remained identical. Carelessness was also evident in student 26, who failed to answer easy questions but correctly answered more difficult ones. According to Wise (2015), effortless test-taking behaviours, such as carelessness, can introduce systematic negative bias in test scores. Moreover, external conditions, such as the testing environment and socioeconomic background, can also affect test performance. A notable example is Spain's exclusion from PISA 2018, which was attributed to abnormal student response behaviours, emphasizing how environmental factors can lead to unreliable test scores (van Grieken et al., 2024). The classroom atmosphere and overall learning environment are crucial in improving students' performance. A positive and conducive learning environment can improve performance, regardless of the initial academic foundation (Zhang, 2011). These results further emphasized the superiority of Rasch model, particularly the ability to relate pre-test and post-test results, item difficulty, and measured content (Planinic et al., 2019; Hope et al., 2024). The results originated from stacking and racking analyses, where staking identified students who experienced changes and racking determined which items showed higher or lower difficulty changes (Laliyo et al., 2022; Öz et al., 2024). Although the instrument measurements in this study were not data-rich, the analytical power of the Rasch model provided a detailed representation of changes in both ability and item difficulty.

6. Conclusion

This study successfully developed PBWs to facilitate the teaching of logic gates in the context of pneumatic and evaluated their effectiveness in enhancing students' conceptual understanding. The results of the first question demonstrated a notable improvement in students' understanding, as evidenced by increased total test scores and supported by stacking and racking analyses. These findings indicate that the PBWs intervention effectively promoted meaningful conceptual change. The PBWs approach provided a more holistic and engaging learning experience. By integrating real-world problem solving and collaborative activities, PBWs surpassed traditional instructional methods, fostering deeper understanding and active student participation. However, findings from the Rasch

model-based item difficulty analysis revealed complexities in the assessment of conceptual understanding. Despite overall performance gains, some test items exhibited increased difficulty post-intervention due to the uneven distribution of correct responses, which elevated their logit values. This suggests that improvements in total scores may not uniformly reflect gains across all item types. Several contributing factors were identified. These included: (1) inconsistent responses from high-ability students, especially on low-difficulty items; (2) deviations from the expected item difficulty hierarchy in the Rasch model; (3) 22% of low-ability students correctly answering high-difficulty items, potentially due to guessing or uncontrolled external factors; and (4) observed instances of carelessness, cheating, or random guessing during the assessment process.

7. Suggestion

Future research should explore strategies to address the discrepancies observed in students' performance after the intervention. Specifically, investigating methods to minimize inconsistent responses among high-ability students and analyzing factors influencing deviations from the expected difficulty hierarchy could provide deeper insights. Another important limitation to acknowledge is the absence of a delayed post-test to measure students' retention of conceptual understanding after the intervention. Without such data, it is difficult to determine the long-term effectiveness of the PBWs approach. Therefore, future research is encouraged to incorporate longitudinal assessments to evaluate the sustainability of learning gains over time.

Declarations

Author Contributions. N.U.: Introduction, methodology, data analysis. A.S.: review-editing and writing, literature review, conceptualization. I.H.: review-editing and writing, original manuscript preparation. T.K.: review-editing and writing. All authors have read and approved the published on final version of the article

Conflicts of Interest. The authors declare no conflict of interest.

Funding. The Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia supported this study through Grant PKPI PMDSU (2127/E4/DT.04.02/2024).

Data Availability Statement. The data that supports the findings of this study are confidential as bound by the human ethics application

Acknowledgements. The authors are grateful to all parties who contributed to data processing and research support

References

- Abu Bakar, S. N., Mahmood, N., Ismail, A., Hashim, A., Hamsan, N. Y., Nor Hamid, N. I., & Che Yakzam, R. (2023). Validity and reliability of competency analysis instrument for cooperative board members using Rasch measurement model approach. *International Journal of Academic Research in Business and Social Sciences*, 13(6), 2408–2419. <https://doi.org/10.6007/ijarbss/v13-i6/17543>
- Anselmi, P., Vidotto, G., Bettinardi, O., & Bertolotti, G. (2015). Measurement of change in health status with Rasch models. *Health and Quality of Life Outcomes*, 13(1), 1–7. <https://doi.org/10.1186/s12955-014-0197-x>
- Aritia, E., & Suyanto, S. (2019). The effect of problem-based Learning model and concept map strategy for problem solving and understanding of the ecosystem concept of high school students. *Journal of Physics: Conference Series*, 1233(1). <https://doi.org/10.1088/1742-6596/1233/1/012005>
- Boone, W. J., Yale, M. S., & Staver, J. R. (2014). Rasch analysis in the human sciences. *In Rasch Analysis in the Human Sciences*. <https://doi.org/10.1007/978-94-007-6857-4>

- Bouchée, T., de Putter - Smits, L., Thurlings, M., & Pepin, B. (2022). Towards a better understanding of conceptual difficulties in introductory quantum physics courses. *Studies in Science Education*, 58(2), 183–202. <https://doi.org/10.1080/03057267.2021.1963579>
- Çakıroğlu, Ü., Atabaş, S., Aydın, M., & Özyılmaz, I. (2022). Creating concept maps with augmented reality: A case of eclipse of the lunar and solar topic. *Research and Practice in Technology Enhanced Learning*, 17(1), 1–22. <https://doi.org/10.1186/s41039-022-00191-1>
- Cano, F., García, Á., Justicia, F., & García, A. B. B. (2014). Learning approaches and reading comprehension: The role of student questioning and prior knowledge. *Revista de Psicodidactica*, 19(2), 247–266. <https://doi.org/10.1387/RevPsicodidact.10186>
- Corlatescu, D., Watanabe, M., Ruseti, S., Dascalu, M., & McNamara, D. S. (2023). The automated model of comprehension version 3.0: Paying attention to context. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13916 LNAI, 229–241. https://doi.org/10.1007/978-3-031-36272-9_19
- da Silva, J. G., Alvarenga, A. V., & Costa-Felix, R. P. . (2019). Evaluating the reliability and validity of a questionnaire used to collect data about the periodic evaluation of ultrasound therapy equipment. *IFMBE Proceedings*, 70(1), 825–828. <https://doi.org/10.1007/978-981-13-2119-1>
- Dai, J. Y., Yeh, K. L., Kao, M. T., Yuan, Y. H., & Chang, M. W. (2021). Applying petri-net to construct knowledge graphs for adaptive learning diagnostics and learning recommendations. *Journal of Research in Education Sciences*, 66(3), 61–105. [https://doi.org/10.6209/JORIES.202109_66\(3\).0003](https://doi.org/10.6209/JORIES.202109_66(3).0003)
- Das, S., Nandgaonkar, V., Eklarker, R., Balkhande, B. W., & Pande, S. D. (2023). Exploring the dynamics of pbl-based learning: A study on collaboration, reflection, engagement, critical thinking, and student success. *Design and Implementation of Higher Education Learners' Learning Outcomes (HELLO)*, 146–158. <https://doi.org/10.4018/978-1-6684-9472-1.ch009>
- Devika, S. V., Siddapuram, A., Kaur, R., & Kondala Rao, P. (2024). Enhancing program outcomes and academic performance through problem-based learning in an antennas and wave propagation course. *Journal of Engineering Education Transformations*, 38(Special Issue 1), 140–145. <https://doi.org/10.16920/jeet/2024/v38is1/24223>
- Dotimineli, A., & Mawardi, M. (2021). Development of STEM integrated PBL-based student worksheets in energetic materials of first-year students. *Journal of Physics: Conference Series*, 1788(1), 12045. <https://doi.org/10.1088/1742-6596/1788/1/012045>
- Engelhard, G. (2013). Invariant measurement: Using rasch models in the social, behavioral, and health sciences. *Invariant Measurement: Using Rasch Models in the Social, Behavioral, and Health Sciences*, 1–288. <https://doi.org/10.4324/9780203073636>
- Ferreira, M. L. P., Almeida, R. M. V. R., & Luiz, R. R. (2013). A new indicator for the measurement of change with ordinal scores. *Quality of Life Research*, 22(8), 1999–2003. <https://doi.org/10.1007/s11136-012-0288-2>
- Fries, L., Son, J. Y., Givvin, K. B., & Stigler, J. W. (2021). Practicing connections: A framework to guide instructional design for developing understanding in complex domains. *Educational Psychology Review*, 33(2), 739–762. <https://doi.org/10.1007/s10648-020-09561-x>
- Haruehansawasin, S., & Kiattikomol, P. (2018). Scaffolding in problem-based learning for low-achieving learners. *Journal of Educational Research*, 111(3), 363–370. <https://doi.org/10.1080/00220671.2017.1287045>

- Hermundstad, A. L., Diller, T. E., Williams, C. B., & Matusovich, H. M. (2016). Exploring conceptual understanding in heat transfer: A qualitative analysis. *ASEE Annual Conference and Exposition, Conference Proceedings*. <https://doi.org/10.18260/p.26845>
- Hmelo-Silver, C. E., & Barrows, H. S. (2008). Facilitating collaborative knowledge building. *Cognition and Instruction*, 26(1), 48–94. <https://doi.org/10.1080/07370000701798495>
- Hope, D., Kluth, D., Homer, M., Dewar, A., Goddard-Fuller, R., Jaap, A., & Cameron, H. (2024). Exploring the use of Rasch modelling in “common content” items for multi-site and multi-year assessment. *Advances in Health Sciences Education*, 1–12. <https://doi.org/10.1007/s10459-024-10354-y>
- Hung, W., & Amida, A. (2020). Problem-based learning in college science. *Active Learning in College Science: The Case for Evidence-Based Practice*, 325–339. <https://doi.org/10.1007/978-3-030-33600-4>
- Jaganathan, S., Bhuminathan, S., Ramesh, M. (2024). Problem based learning - An overview. *Journal of Pharmacy and Bioallied Sciences*, 16(2). <https://doi.org/10.4103/jpbs.JPBS>
- Jansen, B. J., Aldous, K. K., Salminen, J., Almerexhi, H., & Jung, S. gyo. (2024). A discussion of the validity of data analytics. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 139–145. https://doi.org/10.1007/978-3-031-41933-1_12
- Küçük - Avcı, Ş., Akgün, Ö. E., & Balkan - Kırıyıcı, F. (2024). Comparison of the effects of PBL in 3D virtual environment and F2F on learning and spatial skills. *Education and Information Technologies*, 29(17), 12686. <https://doi.org/10.1007/s10639-024-12686-z>
- Kyriakopoulou, N., & Skopeliti, I. (2024). Rethinking Engagement: Insights from conceptual change research. *Psychology*, 29(2). <https://doi.org/10.12681/psy>
- Laliyo, L. A. R., Sumintono, B., & Panigoro, C. (2022). Measuring changes in hydrolysis concept of students taught by inquiry model: Stacking and racking analysis techniques in Rasch model. *Heliyon*, 8(3), 9126. <https://doi.org/10.1016/j.heliyon.2022.e09126>
- Lin, J. W., Yen, M. H., Liang, J. C., Chiu, M. H., & Guo, C. J. (2016). Examining the factors that influence students’ science learning processes and their learning outcomes: 30 years of conceptual change research. *Eurasia Journal of Mathematics, Science and Technology Education*, 12(9), 2617–2646. <https://doi.org/10.12973/eurasia.2016.000600a>
- Linacre, J. M. (1994). Sample Size and Item Calibration or Person Measure Stability. *Rasch Measurement Transactions*, 7(4), 328. <http://www.rasch.org/rmt/rmt74m.htm>
- Lu, J., Bridges, S., & Hmelo-Silver, C. E. (2014). Problem-based learning. *In The Cambridge Handbook of the Learning Sciences, Second Edition*. <https://doi.org/10.1017/CBO9781139519526.019>
- Lu, S., & Bi, H. (2016). Development of a measurement instrument to assess students’ electrolyte conceptual understanding. *Chemistry Education Research and Practice*, 17(4), 1030–1040. <https://doi.org/10.1039/c6rp00137h>
- Lugosi, E., & Uribe, G. (2022). Active learning strategies with positive effects on students’ achievements in undergraduate mathematics education. *International Journal of Mathematical Education in Science and Technology*, 53(2), 403–424. <https://doi.org/10.1080/0020739X.2020.1773555>
- Maidan, Halim, A., Safitri, R., & Nurfadilla, E. (2020). Impact of Problem-based Learning (PBL) model through Science Technology Society (STS) approach on students’ interest. *Journal of Physics: Conference Series*, 1460(1), 12145. <https://doi.org/10.1088/1742-6596/1460/1/012145>

- Marhaeni, N. H., & Andriyani. (2023). Students' worksheet through problem based learning in determinant and inverse matrix material. *AIP Conference Proceedings*, 2733(1), 141780. <https://doi.org/10.1063/5.0141780>
- Öz, E., Bulut, O., Cellat, Z. F., & Yürekli, H. (2024). Stacking: An ensemble learning approach to predict student performance in PISA 2022. *Education and Information Technologies*, 13110. <https://doi.org/10.1007/s10639-024-13110-2>
- Pascual, G., & Goikoetxea, E. (2014). Comprehension of university texts: Effects of domain-knowledge and summary. *Reading Psychology*, 35(2), 101–126. <https://doi.org/10.1080/02702711.2012.664612>
- Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, 15(2), 20111. <https://doi.org/10.1103/PhysRevPhysEducRes.15.020111>
- Rahman, M. M., Adeniran, O., Shourabi, N. B., & Owolabi, O. (2022). Initial impact of evidence-based and experiment-focused teaching approach in a computer architecture course in computer science. *Proceedings - 2022 International Conference on Computational Science and Computational Intelligence*, 1996–2002. <https://doi.org/10.1109/CSCI58124.2022.00359>
- Raja Subramanian, R., Venkatesh, K., Manikumar, T., Raja Sudharsan, R., & Aravindrajan, V. (2024). SHAPE: Design and evaluation of a transformative model for engineering education. *Journal of Engineering Education Transformations*, 38(Special Issue 1), 153–158. <https://doi.org/10.16920/jeet/2024/v38is1/24225>
- Samsudin, A., Suhandi, A., Rusdiana, D., Kaniawati, I., Fratiwi, N. J., Zulfikar, A., Muhaemin, M. H., Hermita, N., Mansur, Wibowo, F. C., Supriyatman, Malik, A., & Costu, B. (2019). Optimizing students' conceptual understanding on electricity and magnetism through cognitive conflict-based multimode teaching (CC-BMT). *Journal of Physics: Conference Series*, 1204(1), 12027. <https://doi.org/10.1088/1742-6596/1204/1/012027>
- Sanusi, N., Zulkifli, H., & Hamzah, M. I. (2025). Challenges, opportunities, and effects of alternative assessment approaches in teaching practices: a systematic literature review. *International Journal of Evaluation and Research in Education*, 14(2), 1105–1113. <https://doi.org/10.11591/ijere.v14i2.32283>
- Saravanakumar, D., Mohan, B., & Muthuramalingam, T. (2017). A review on recent research trends in servo pneumatic positioning systems. *Precision Engineering*, 49, 481–492. <https://doi.org/10.1016/j.precisioneng.2017.01.014>
- Shana, Z. A., & El Shareef, M. A. (2022). Science teachers' use of analogies: Findings from classroom practices. *European Journal of Educational Research*, 11(2), 1023–1036. <https://doi.org/10.12973/EU-JER.11.2.1023>
- Shepherd, T., Trinder, M., & Theophilus, M. (2022). Does virtual reality in the perioperative setting for patient education improve understanding? A scoping review. *Surgery in Practice and Science*, 10. <https://doi.org/10.1016/j.sipas.2022.100101>
- Sipayung, T. N. (2019). The development of mathematics' worksheets based on problem solving at the numbers topics. *Journal of Physics: Conference Series*, 1211(1), 6–15. <https://doi.org/10.1088/1742-6596/1211/1/012071>
- Soeharto, S., & Csapó, B. (2021). Evaluating item difficulty patterns for assessing student misconceptions in science across physics, chemistry, and biology concepts. *Heliyon*, 7(11), 8352. <https://doi.org/10.1016/j.heliyon.2021.e08352>

- Sriatun, S., Ekohariadi, Sondang, M., Winanti, T., Ismayanti, E., & Nurlaela, L. (2020). Effectiveness in mathematics learning with the problem solving approach in vocational high school. *Journal of Physics: Conference Series*, 1511(1), 12116. <https://doi.org/10.1088/1742-6596/1511/1/012116>
- Stolt, M., Kottorp, A., & Suhonen, R. (2022). The use and quality of reporting of Rasch analysis in nursing research: A methodological scoping review. *International Journal of Nursing Studies*, 132. <https://doi.org/10.1016/j.ijnurstu.2022.104244>
- Syahmani, Suyono, & Imam Supardi, Z. A. (2020). Effectiveness of i-smart learning model using chemistry problems solving in senior high school to improve metacognitive skills and students' conceptual understanding. *Pedagogika*, 138(2), 37–60. <https://doi.org/10.15823/p.2020.138.3>
- Tavakol, M., O'Brien, D. G., Sharpe, C. C., & Stewart, C. (2024). Twelve tips to aid interpretation of post-assessment psychometric reports. *Medical Teacher*, 46(2), 188–195. <https://doi.org/10.1080/0142159X.2023.2241624>
- Utami, W. P., Louise, I. S. Y., & Rampean, B. A. O. (2022). Analysis of teacher's need through student's worksheet development based on problem based learning on acid-base material. *AIP Conference Proceedings*, 2645, 113898. <https://doi.org/10.1063/5.0113898>
- Van Grieken, R., Tena, J. D., Pires, L., Sanz, I., & AvendañoMiranda, L. L. (2024). The importance of test order in external and standardized test results: The case of PISA 2018. *PLoS ONE*, 19(9), 309980. <https://doi.org/10.1371/journal.pone.0309980>
- Wallmark, J., Ramsay, J. O., Li, J., & Wiberg, M. (2023). Analyzing Polytomous Test Data: A Comparison Between an Information-Based IRT Model and the Generalized Partial Credit Model. *Journal of Educational and Behavioral Statistics*, 49(5), 753–779. <https://doi.org/10.3102/10769986231207879>
- Wise, S. L. (2015). Effort Analysis: Individual Score Validation of Achievement Test Data. *Applied Measurement in Education*, 28(3), 237–252. <https://doi.org/10.1080/08957347.2015.1042155>
- Yazbec, A., Borovsky, A., & Kaschak, M. P. (2019). Examining the impact of text style and epistemic beliefs on conceptual change. *PLoS ONE*, 14(9), 220766. <https://doi.org/10.1371/journal.pone.0220766>
- Yeni, S., Grgurina, N., Saeli, M., Hermans, F., Tolboom, J., & Barendsen, E. (2024). Interdisciplinary integration of computational thinking in k-12 education: A systematic review. *Informatics in Education*, 23(1), 223–278. <https://doi.org/10.15388/infedu.2024.08>
- Yerizon, Arnawa, I. M., Fitriani, N., & Tajudin, N. M. (2022). Constructing calculus concepts through worksheet based problem-based learning assisted by geogebra software. *HighTech and Innovation Journal*, 3(3), 282–296. <https://doi.org/10.28991/HIJ-2022-03-03-04>
- Yıldırım, T., & Canpolat, N. (2019). An investigation of the effectiveness of the peer instruction method on teaching about solutions at the high-school level. *Eğitim ve Bilim*, 44(199), 127–147. <https://doi.org/10.15390/EB.2019.7966>
- Yusop, F. D., & Firdaus, A. (2022). Alternative assessments in malaysian higher education: Voices from the field., 1–309. <https://doi.org/10.1007/978-981-16-7228-6>
- Zhang, Q. (2011). The application of multiple linear regress equation on analysing the effect factors of student scores. *2011 International Conference on Multimedia Technology, ICMT 2011*, 6087–6089. <https://doi.org/10.1109/ICMT.2011.06002588>

About the Contributor(s)

Nisaudzakiah Utami is a dedicated PhD candidate in Technical and Vocational Education at Universitas Pendidikan Indonesia. Her current areas of research interest are educational technology, vocational education, and higher-order thinking skills.

Email: udzakiahnisa@upi.edu

ORCID: <https://orcid.org/0009-0008-9186-4766>

Agus Setiawan is an Associate Professor and the Director of the Faculty of Engineering and Industrial Education at Universitas Pendidikan Indonesia, Indonesia. His doctorate was awarded by Tohoku University, Japan. His undergraduate degree is from IKIP Bandung, and his master's degree is from Institut Teknologi Bandung. His current areas of interest are Technical and Vocational Education and Training.

Email: agus_setiawan@upi.edu

ORCID: <https://orcid.org/0000-0001-8848-0720>

Ida Hamidah is a Professor of Research Methodology and scientific publication at Universitas Pendidikan Indonesia. Her doctorate was awarded by Institut Teknologi Bandung. Her undergraduate and master's degrees are from IKIP Bandung. Her current areas of interest are Material Physics, Engineering, and Science Education.

Email: idahamidah@upi.edu

ORCID: <https://orcid.org/0000-0001-7720-8326>

Thomas Köhler has been Professor of Educational Technology at the Institute for Vocational Education of Dresden University of Technology since 2005, where he serves as deputy director. He studied at Friedrich-Schiller-University Jena (Germany) and Swarthmore College (USA), holds degrees in psychology and sociology, and received his Ph.D. in communications psychology from FSU Jena in 1999. Main topics of my research deal with distance education, virtual organizations, tele-cooperation and computer-mediated communication, online research methods, and computer user's personality.

Email: Thomas.Koehler@tu-dresden.de

ORCID: <https://orcid.org/0000-0002-9438-5786>

Publisher's Note: *The opinions, statements, and data presented in all publications are solely those of the individual author(s) and contributors and do not reflect the views of Universitepark, EDUPIJ, and/or the editor(s). Universitepark, the Journal, and/or the editor(s) accept no responsibility for any harm or damage to persons or property arising from the use of ideas, methods, instructions, or products mentioned in the content.*
