

Development and Calibration of a Moodle-Based Computerized Adaptive Test Using Rasch Model for Electrical Competencies in Vocational Schools

Fitria Wahyu Ningsih*, Samsul Hadi

Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

*Corresponding Email: fitriawn00@gmail.com

ARTICLE INFO

Article history

Received: Sep 04, 2025

Revised: Jun 24, 2026

Accepted: Jun 26, 2026

Keywords

Computerized Adaptive Test

Vocational Electrical Competency

Rasch Model

Moodle-Based Assessment

4D Development Model

ABSTRACT

This study developed a Moodle-based Computerized Adaptive Test (CAT) to assess vocational students' competencies in electrical installation and troubleshooting. Conventional assessments often lack efficiency, security, and personalized feedback, motivating the development of an adaptive evaluation system. Using the 4D research and development model, 150 test items were created and calibrated with the Rasch Model. After removing misfit and overfit items, 106 items were confirmed valid and reliable (reliability = 0.96). The instrument provides a balanced difficulty distribution, enabling accurate measurement across different ability levels. Expert validation rated the content as "Highly Feasible" (94.75%) and the media as "Feasible" (89.75%), while student responses showed positive acceptance (76%–80%). Implementation demonstrated that the CAT effectively differentiates student abilities and provides immediate diagnostic feedback for formative assessment. The study offers an adaptive, Moodle-integrated assessment tool and recommends expanding the item bank to better measure higher-order thinking skills. This is an open access article under the [CC-BY-SA](#) license.



Introduction

The rapid development in the Society 5.0 era has brought significant transformation in the world of education, especially through digitalization that allows the learning process to be more efficient, personalized, and relevant to the needs of the times Wijaya et al (2024). In Vocational High Schools

(SMK), the use of technology is the main key to ensuring the achievement of student competencies that are in accordance with industry standards and the demands of the ever-evolving world of work. One crucial aspect in education is learning evaluation, which plays a central role in the entire process, especially in vocational competencies such as installation and troubleshooting of electrical systems. This competency is very important considering the dependence of various sectors on a stable and secure supply of electrical energy, as well as the increasing need for skilled electrical technicians in Indonesia Kementerian Energi dan Sumber Daya Mineral Republik Indonesia (2019). SMK graduates, especially from the Electrical Engineering expertise program, are expected to be able to meet the demands of this job market by focusing on installation capabilities according to standards (for example, PUIL) and expertise in identifying and repairing electrical system disturbances.

However, learning to achieve these competencies in vocational high schools often faces significant challenges, such as limited computer facilities and students' difficulty mastering basic cognitive concepts, which can impact their readiness for the industrial world. Data Badan Pusat Statistik (2024) shows that the open unemployment rate for vocational high school graduates is 9.01%, the highest compared to other levels. This highlights a gap between competencies and industry needs, necessitating more adaptive and accurate evaluation through the development of technology-based instruments.

Conventional methods used at SMKN 1 Kramatwatu Serang, such as Paper-Based Tests (PBT) and Computer-Based Tests (CBT), have several limitations. These types of tests have significant weaknesses. Question security is often not guaranteed, potentially leaked or accessed by unauthorized parties. The creation process is also time-consuming and paper-intensive (Hadi, 2013). In addition, the assessment process is time-consuming and resource-intensive, especially for Paper-Based Tests (PBT) which require manual correction and duplication of questions. Paper-Based Tests (PBT) also have a relatively high risk of cheating and are susceptible to human error Tananda et al (2023). This often means students cannot receive quick and personalized feedback on test results. This delay in feedback negatively impacts student motivation and learning engagement Corral et al (2021) and increases the administrative burden for teachers Fahrurrizal et al (2024).

Furthermore, these conventional test instruments are based on Classic Test Theory (CTT), that has long been the standard for instrument validation. It is inherently limited by its dependency on the specific sample being tested, which hinders the objective measurement of student ability across different difficulty levels. The weakness of CTT is that it does not consider the difficulty level of

individual items (Bond et al., 2020). Existing test instruments are not fully able to measure students' abilities accurately and vary according to individual competency levels. The questions presented should be more dynamic so that they can adjust the level of difficulty dynamically. This is related to Item Respond Theory (IRT), an approach used to model the relationship between individual abilities and their responses to an item (question) in a test Ayala (2022).

In developing an instrument based on Item Respond Theory (IRT), it can be analyzed using the Rasch Model by simultaneously considering the level of difficulty of the questions and individual abilities. The Rasch Model provides a robust framework for analyzing IRT-based instruments, as it ensures that the probability of a correct response is strictly a function of the person's ability and the item's difficulty, thereby enhancing the instrument's fundamental validity Bond et al (2020).

Computer Adaptive Test (CAT) is a sophisticated form of computer-based testing that tailors the exam to each examinee's ability level by selecting items from a calibrated bank based on previous responses Thompson and Weiss (2021). The implementation of the Rasch model within adaptive testing ensures that item difficulty and person ability are measured on the same scale, providing an objective basis for adaptive item selection Bond et al (2020). The CAT system selects questions adaptively based on student responses, shortening the exam duration, and increasing the efficiency and precision of the estimation of participant ability (Falani, 2023b; Wulandari et al., 2020).

Recent studies have extensively explored the transition from conventional testing to digital platforms. For instance, Falani (2023) demonstrated that CAT significantly improves assessment precision and validity compared to standard CBT. In its implementation, Moodle platform is often chosen due to its flexibility in adjusting to learning needs. As a popular Learning Management System (LMS), Moodle can be integrated with plugins or external systems to support adaptive tests, where questions will adjust the level of difficulty based on the participant's abilities in real-time Sáiz-Manzanares et al (2021).

While previous research has successfully implemented CAT in general subjects like mathematics or language Wulandari et al (2020) its application in highly specialized vocational competencies, specifically electrical installation and troubleshooting remains limited. Most vocational schools still utilize Moodle as a static platform for linear testing, failing to leverage its potential for adaptive evaluation Sáiz-Manzanares et al (2021). The gap addressed in this study lies in the integration of an adaptive engine within the Moodle environment specifically designed for technical-vocational tasks. This research fills this void by developing a Moodle-based CAT system where the item bank is meticulously calibrated using the Rasch Model. This study provides a

practical framework for implementing high-precision, personalized assessment in a vocational context, ensuring that student competency is measured with higher accuracy and efficiency than traditional methods. By considering the relevance of vocational education and technological developments, this study focuses on the development and implementation of Moodle-based CAT for electrical system installation and troubleshooting competencies at SMKN 1 Kramatwatu Serang.

Method

This study uses the Research and Development (R&D) method with the 4D development model (Define, Design, Develop, Disseminate). The main objective is to develop a Moodle-based Computer Adaptive Test (CAT) as an evaluation tool for electrical installation and troubleshooting. The research was conducted over a six-month period, from January to June 2025, at SMKN 1 Kramatwatu Serang. The systematic process commenced with the Define phase in January 2025, which involved material and needs analysis. This was followed by the Design and Develop phases from February to Mei 2025, where the item bank was calibrated and the Moodle-based system was integrated. The final Disseminate phase, including effectiveness testing with thirty Grade XI students from the Electrical Power Installation Engineering program, was carried out in June 2025. Critical validation steps, including expert judgment and psychometric analysis using the Rasch Model, were integrated throughout this timeline to produce a reliable and functionally effective adaptive testing system. The research procedure is illustrated in Figure 1.

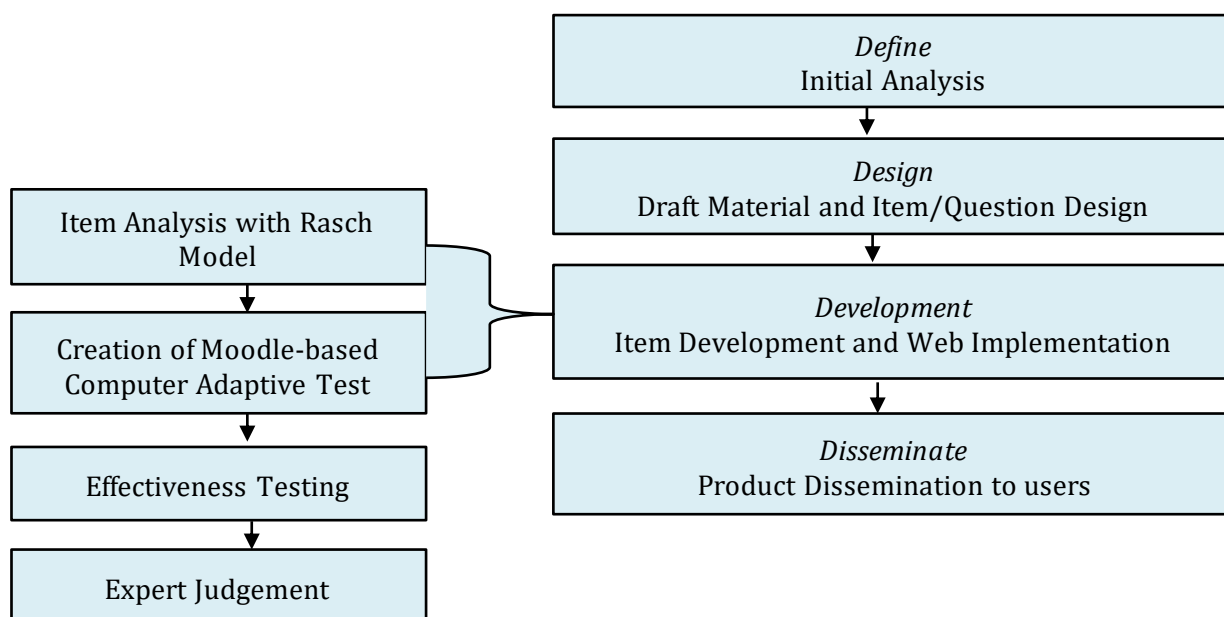


Fig 1: Procedure of 4D Models

This study employed a quantitative research approach, specifically focusing on the psychometric quality of test items. The participants were 11th-grade students enrolled in the Electrical Power Installation Engineering (TITL) program. Data were gathered through multiple-choice tests and validation questionnaires. The quality of test items—including validity, reliability, difficulty level, and model fit—was analyzed quantitatively using the Rasch Model. Furthermore, the product's feasibility was evaluated by subject matter experts using a Likert scale, while user responses were analyzed using descriptive statistics to assess the effectiveness and acceptance of the developed CAT product.

To further evaluate the quality of these 106 valid items, we mapped each item according to the revised Bloom's Taxonomy (C1–C6), as presented in Table 1. This distribution ensures that the assessment covers a comprehensive range of cognitive processes, from basic knowledge retention to higher-order analytical and evaluative competencies necessary for professional electrical troubleshooting. Difficulty Item Categories explain in Table 1.

Table 1. Difficulty Item Categories

No.	Bloom's Level	Frequency	Percentage
1	C1 (Remembering)	15	14.1%
2	C2 (Understanding)	20	18.9%
3	C3 (Applying)	25	23.6%
4	C4 (Analyzing)	20	18.9%
5	C5 (Evaluating)	16	15.1%
6	C6 (Creating)	10	9.4%
	Total	106	100%

The item analysis highlights that while the instrument is highly reliable (0.96), the Wright Map indicates a concentration of items within the low-to-moderate difficulty range. While this is effective for identifying foundational competency gaps, future iterations of the item bank will prioritize the development of more complex troubleshooting scenarios to better differentiate high-ability students.

Result and Discussion

Result Calibration using Rasch Model in RStudio

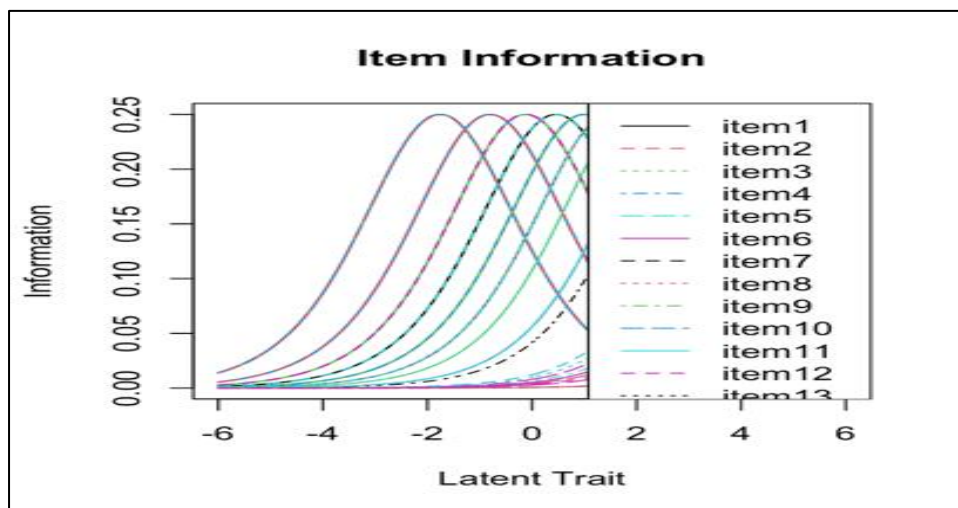
The psychometric properties of the instrument were evaluated through a systematic calibration process using the Rasch Model. This analysis was performed in RStudio using the eRm package to ensure the validity, reliability, and suitability of the item bank. The calibration procedure followed a rigorous filtering process to ensure that only items meeting established quality criteria were

included in the final Computerized Adaptive Test (CAT) instrument. The question bank is calibrated using the Rasch Model to ensure item validity and reliability using RStudio with eRm. Based on the final analysis results, a total of 44 items were eliminated due to not meeting quality standards. The eliminated items consisted of 13 items with negative discrimination, 13 items that did not fit the model (misfit), and 18 items that were redundant (overfit). Thus, the final recommended instrument set consists of 106 test items that have been proven valid. The final classification of item difficulty is presented as follows in Table 2.

Table 2. Difficulty Item Categories

No.	Difficulty Level	Number of Items
1	Very difficult	14
2	Difficult	8
3	Moderate	50
4	Easy	34
Total Valid		106

From 150 initial questions, 106 items were declared valid, with a distribution of varying difficulty levels, namely 14 very difficult questions, 8 difficult questions, 50 medium questions, and 34 easy questions as seen in the following table. Meanwhile, the graphic results show that most of the most informative items are in latent abilities between -3 and 0, meaning that the questions are more suitable for participants with low to medium abilities. The graph can be seen below in Figure 2.

**Fig 2:** Item Information Function Graph (IIF)

The reliability of this test instrument is very high, namely 0.96. Furthermore, in the web implementation, the difficulty level of the questions has been dynamically adjusted so that it can

measure according to the test participants' abilities. The test results are displayed comprehensively in Figure 3.

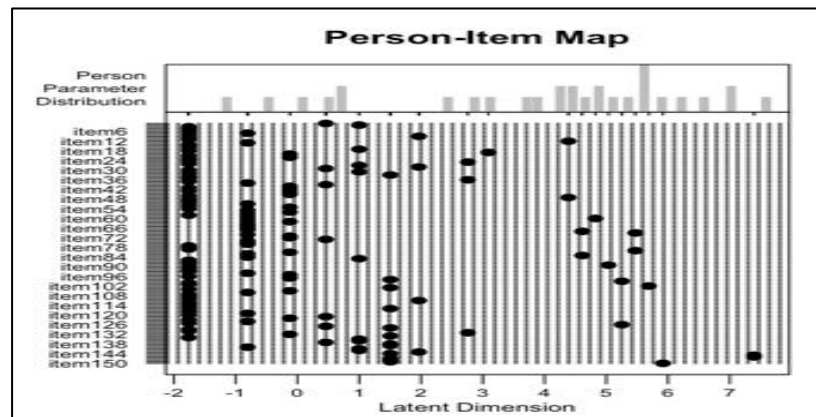


Fig 3: Person-Item Map

This graph is one of the most informative outputs in Rasch analysis, as it visually summarizes the relationship between respondent ability levels (Persons) and item difficulty levels (Items) on the same scale (logit scale). Based on the Wright map analysis, it can be concluded that the instrument used is still too easy for participants with high abilities. The instrument appears to be more suitable for respondents with low to moderate ability levels. This is because most of the items have low difficulty levels, making them less effective in distinguishing respondents with higher abilities.

In its validation, the Computer Adaptive Test (CAT) was assessed through validation by two material experts and two media experts, as well as through responses from end users, namely students. The results of the expert judgment regarding content feasibility are presented in Table 3.

Table 3. Expert Judgment of Content

No.	Indicator	Teacher1	Teacher2	Total	Percentage
1	Truth of the Content of the Material	15	16	31	97%
2	Current Material	13	16	29	91%
3	Breadth & Depth of Material	13	16	29	91%
4	Conformity to Learning Objectives	16	16	32	100%
Total Average				60	

The validation results by two vocational school teacher Content experts showed an average score of 60, which is categorized as "Very Appropriate." In detail, the aspect of the accuracy of the material content obtained a feasibility percentage of 97%, the Content's currentness of 91%, the breadth and depth of the Content also 91%, and the suitability to learning objectives reached 100%. These results reflect that the evaluation Content has met the criteria of factual integrity, curriculum

relevance, and comprehensive content coverage. The Calculation result of Expert Judgment of Media can see in Table 4.

Table 4. Expert Judgment of Media

No.	Indicator	Lecture1	Lecture2	Total	Percentage
1	Test Reliability	16	15	31	97%
2	Ease of Maintenance	16	13	29	91%
3	Easy to Operate	16	12	28	88%
4	Application Compatibility	15	14	29	91%
5	Complete Documentation	16	13	29	91%
6	Clear Visual Communication	16	12	28	88%
7	Creativity in Implementing Ideas	14	13	27	84%
8	Interactive Layout	16	12	28	88%
Total Average				114.5	

The assessment by two media experts resulted in an average score of 114.5, categorized as “Feasible”. Quantitative evaluation of the indicators showed that the test reliability reached 97%, ease of maintenance and application suitability were each 91%, ease of operation and clarity of visual communication were 88%, creativity in implementing ideas was 84%, and interactive layout was 88%. These findings indicate that the product has high technical quality, with an intuitive interface design and can be operated efficiently by users. Regarding the participant involvement, this study utilized a total sampling technique (saturated sampling) due to the specific focus on the eleventh-grade students of the Electrical Power Installation Engineering. The population consisted of 30 students, all of whom were involved in every stage of the research. Specifically, the same students participated in the initial pilot testing for Rasch Model item calibration and the subsequent effectiveness testing for user response analysis. The results from these respondents indicated a positive level of acceptance. The assessment indicators were adopted from Yuniarti et al., (2024), where the user response results for this product ranged from 76% to 80%, falling into the “Good” category. The response percentage graph of the students can be seen in Figure 4.

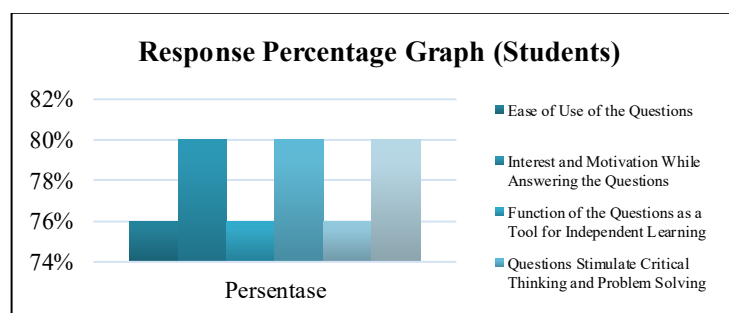


Fig 4: Response Percentage graph (Students)

Figure 4 illustrates the distribution of student responses across several evaluation indicators, including ease of use, system navigation, and clarity of instructions. The data shows that the majority of students provided positive feedback, placing the Computerized Adaptive Test system in the 'Good' category. This high level of acceptance indicates that the Moodle-based adaptive interface is user-friendly and effectively assists vocational students in completing competency assessments without significant technical obstacles.

Implementation Result on Moodle Web

The Computerized Adaptive Test (CAT) system features an integrated interface designed to ensure students focus on problem-solving rather than technical navigation. The main navigation structure has been strategically placed at the top of the page, featuring links to the CAT Guide for user orientation, Start Test as the primary call-to-action, and the Developer Profile. The overall layout is designed to deliver an intuitive and well-guided user experience. This is illustrated in Figure 5 below.



Fig 5: Front Page Interface

The Computerized Adaptive Test (CAT) system features an integrated interface consisting of a material page, a CAT guide, developer profiles, and a "Start Test" portal. The interface for the adaptive test is designed to be intuitive, ensuring that students can focus on problem-solving rather than technical navigation. The system delivers questions using an adaptive algorithm, providing real-time feedback such as an "Attempt Summary" and "Attempt Graph" upon completion. A critical departure from traditional testing is the real-time feedback loop upon completion. That includes their final score, duration, and the specific reason the test stopped (e.g., reaching the target precision).

Email : joves@mpv.uad.ac.id

Website : <http://journal2.uad.ac.id/index.php/joves>

The product interface features a home menu consisting of a material page, a Computerized Adaptive Test (CAT) guide page, a developer profile page, and a Start Test Menu page that can be seen below in Figure 6.

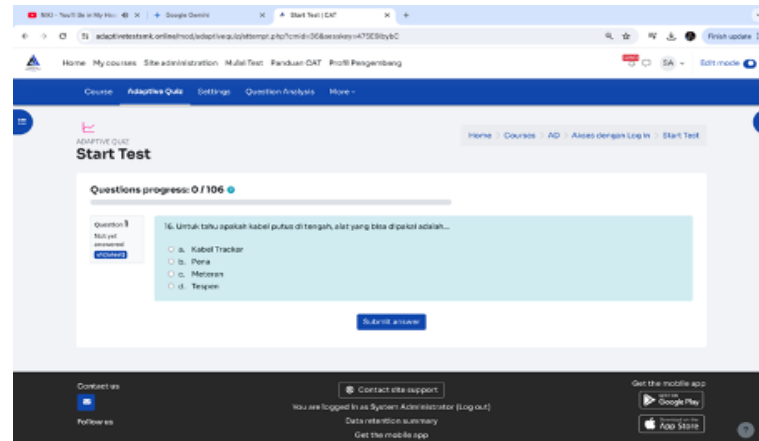


Fig 6: Web Page for the Adaptive Test

The system delivers questions using an adaptive algorithm, providing real-time feedback such as an "Attempt Summary" and "Attempt Graph" upon completion. A critical departure from traditional testing is the real-time feedback loop upon completion. That includes their final score, duration, and the specific reason the test stopped (e.g., reaching the target precision).

Adaptive performance visualization is presented through an "Attempt Graph" chart that shows the "Ability Measure" and movement of question difficulty levels, as well as an "Answer Distribution" chart that displays the number of correct and incorrect answers based on difficulty level. The Graph of score ability of users on the Moodle that can be seen below in Figure 7, which tracks the "Ability Measure" and the fluctuating difficulty levels of the questions served.

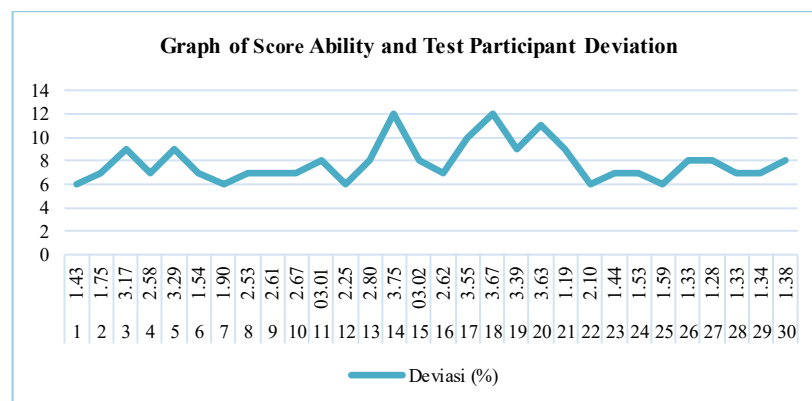


Fig 7: Graph of Score Ability and Test Participant Deviation

Based on the results of user response measurements, the implementation of Computerized Adaptive Test (CAT) on 30 students showed ability scores varying from 1.2 to 3.7. The distribution of abilities includes 9 students (30%) with low abilities (≤ 1.6), 3 students (10%) with medium abilities (1.7–2.0), 2 students (6.7%) with high abilities (2.1–2.4), and 16 students (53.3%) with very high/advanced abilities (≥ 2.5). This distribution aligns with the perspective that latent class models in The Computerized Adaptive Test (CAT) allow for a more personalized mapping of vocational student profiles.

From a Rasch Model perspective, this distribution signifies that the item bank's difficulty parameters are well-aligned with the students' latent traits. The fact that 53.3% of students reached the advanced level (≥ 2.5 logits) suggests that the adaptive algorithm successfully navigated the item bank to find 'maximum information' items for high-performing students. In a fixed-linear test, these students would likely achieve near-perfect scores, masking their true upper limits. However, by serving items of proportional difficulty in real-time, this The Computerized Adaptive Test (CAT) system maintains a consistent Standard Error of Measurement (SEM) across all ability levels, providing a more authentic and granular profile of vocational competence that is essential for industry-standard certifications.

Furthermore, the Graph of Score Ability and Test Participant Deviation demonstrates that as the system adaptively presents items, the deviation percentages remain relatively controlled, primarily fluctuating between 6% and 12% across the various ability levels. This indicates that the The Computerized Adaptive Test (CAT) system has good discrimination power in distinguishing student abilities individually and adaptively. Since the graph tracks real-time progress, it provides the longitudinal accuracy necessary for monitoring competency development (Sáiz-Manzanares et al., 2021), ensuring that the evaluation is both dynamic and responsive to the unique technical mastery of each vocational student.

These findings further indicate that the system effectively addresses the “ceiling effect” and “floor effect” often encountered in traditional vocational assessments. The distribution of ability scores across a wide logit range (1.2 to 3.7) demonstrates high individual-item targeting precision, which proves the system's high discrimination power. Ultimately, the developed instrument precisely differentiates between students who have only mastered foundational concepts and those possessing advanced diagnostic skills in electrical troubleshooting, providing a more objective basis for competency certification in vocational education.

Discussion

Recent studies emphasize that the efficiency of computerized adaptive testing is highly dependent on the quality and structure of the item bank, as cognitively designed item pools significantly improve information efficiency and ability estimation accuracy Luo and Yang (2024). The findings of this study significantly expand the application of Computerized Adaptive Testing (CAT) and the Rasch Model within the vocational education landscape. Compared to previous research by Hadi et al (2022) which was limited by a small bank of 60 items, this study provides a more extensive bank of 150 items (106 validated). This larger scale is essential for ensuring that the adaptive algorithm can accurately cover a wider range of student ability levels and maintain measurement precision.

Previous psychometric studies also emphasize that larger and well-calibrated item banks improve CAT targeting accuracy and information efficiency across heterogeneous ability distributions Mao et al (2022). Advanced item selection strategies in CAT allow the system to maximize test information while minimizing redundant item exposure, resulting in more efficient and targeted ability estimation compared to linear testing models Xi et al (2022).

The high reliability coefficient obtained in this study reinforces the strength of Rasch-based measurement as an objective framework for evaluating technical competencies. Rasch analysis ensures invariant measurement by aligning item difficulty and person ability on the same latent scale, which is essential for maintaining measurement consistency across adaptive assessments (Akhtar and Kovacs, 2024; Bond et al., 2020).

The stability of item parameters is a critical requirement in adaptive testing environments, as unstable items may distort ability estimation. Empirical evidence shows that Rasch-calibrated item banks exhibit high parameter stability across administrations, making them particularly suitable for vocational competency assessment Wyse (2021). A key differentiator of this research is the transition from proprietary software to open-source platforms. While previous psychometric analyses, such as those by Hizqiyah et al (2023) relied heavily on Winsteps, this study utilizes RStudio. A significant methodological advancement in this study is the use of open-source statistical tools for item calibration and adaptive testing implementation. Open-source CAT frameworks integrated with R-based analysis have been shown to provide cost-effective, flexible, and scalable solutions without compromising psychometric quality Oladele (2025). This methodological flexibility is crucial for vocational schools like SMKN 1 Kramatwatu Serang, as it allows for large-

scale data handling without the financial burden of paid licenses, thus ensuring the sustainability of technology-based assessment in resource-constrained environments.

Adaptive testing has proven effective in measuring complex diagnostic skills by dynamically adjusting item difficulty based on student responses. In addition, studies employing latent profile and latent class approaches in technology-based learning contexts reveal that students exhibit heterogeneous performance patterns and engagement profiles, which are often obscured by linear assessment models Pan (2022). This finding supports the present study's results, where the adaptive testing framework was able to differentiate between foundational, intermediate, and advanced troubleshooting competencies more precisely, thereby mitigating the ceiling and floor effects commonly observed in conventional assessments.

Adaptive formative assessment systems are able to capture higher-order reasoning processes more accurately than linear tests, particularly in technical and problem-solving domains Yang et al (2022). While prior implementations of Moodle-based The Computerized Adaptive Test (CAT) by Rahim et al (2024) and Nurdiansyah et al (2020) focused on mathematics and physics, this research successfully extends the adaptive framework into the complex realm of electrical troubleshooting. These skills are often difficult to quantify accurately through traditional linear tests, yet this The Computerized Adaptive Test (CAT) framework successfully captures the dynamic technical reasoning required for identifying and repairing electrical system disturbances. This evolution proves that a The Computerized Adaptive Test (CAT) framework can effectively quantify diagnostic skills, such as identifying and repairing electrical system disturbances that traditional linear tests often fail to capture accurately.

Furthermore, while studies by Yang et al (2022) and Nadia et al (2022) support the efficiency of The Computerized Adaptive Test (CAT), they often overlook specific infrastructure challenges in Indonesian vocational schools. This study demonstrates that despite limited Computerized facilities and student cognitive challenges, a Moodle-based The Computerized Adaptive Test (CAT) can reduce testing time while increasing precision. By precisely differentiating between foundational mastery and advanced troubleshooting skills, this system provides a robust solution to the "ceiling" and "floor" effects. In addition, adaptive termination criteria have been shown to reduce test length without compromising estimation accuracy, thereby enhancing the practical efficiency of computerized adaptive testing systems Wang et al (2022).

This shift signifies a transition from measuring static knowledge to quantifying dynamic technical reasoning, establishing a new benchmark for competency-based assessment that is fairer

and more accurate for vocational students. Theoretically, this shift signifies a transition from measuring static knowledge to quantifying dynamic technical reasoning. Consequently, the integration of a Rasch-calibrated item bank within the Moodle ecosystem not only increases administrative efficiency but also establishes a new benchmark for competency-based assessment that is fairer and more accurate for vocational students.

Implications and Limitations

The results of this study carry significant practical implications for vocational stakeholders. For teachers and schools, this Moodle-based The Computerized Adaptive Test (CAT) provides a time-efficient certification tool that delivers real-time diagnostic data, allowing for immediate remedial intervention. Theoretically, this research establishes a new benchmark for competency-based assessment by shifting from measuring static knowledge to quantifying dynamic technical reasoning.

However, the study is characterized by certain limitations that provide a basis for further development. The primary limitation is the relatively small sample size of 30 students from a single institution, which may affect the generalizability of the findings to broader vocational school contexts in Indonesia. Furthermore, based on the Wright Map and Item Information Function analysis, the current item bank is more suitable for respondents with low to moderate ability levels. Most of the items have low difficulty levels, making the instrument less effective in distinguishing students with higher-order technical abilities. These limitations indicate that although the system is highly reliable, future research should focus on expanding the item bank with more challenging diagnostic tasks and involving a larger and more diverse group of participants.

Conclusion

This study demonstrates that a Moodle-based Computerized Adaptive Test (CAT) integrated with the Rasch Model provides an effective and objective solution for assessing complex vocational competencies, particularly in electrical installation and troubleshooting. By shifting assessment from static knowledge recall to dynamic technical reasoning, the proposed system offers a more precise, efficient, and personalized evaluation aligned with industry standards. Beyond improving measurement accuracy and reducing administrative burden, this framework establishes a robust benchmark for competency-based certification in vocational education. Practically, this system can be directly adopted by vocational schools as an alternative to conventional certification assessments, enabling faster, fairer, and more diagnostic evaluation of students' technical

competencies. Future research should expand the item bank with a greater emphasis on higher-order thinking skills to enhance sensitivity for high-performing students, as well as explore the integration of AI-driven diagnostic feedback and cross-disciplinary implementation to broaden the impact of adaptive assessment in technical education.

References

- Akhtar, H., & Kovacs, K. (2024). Measurement precision and user experience with adaptive versus non-adaptive psychometric tests. *Personality and Individual Differences*, 225, 112675. <https://doi.org/10.1016/j.paid.2024.112675>
- Bond, T. G., Yan, Z., & Heene, M. (2020). *Applying the Rasch model: Fundamental measurement in the human sciences* (4th ed.). Routledge.
- Badan Pusat Statistik. (2024, November 5). *Tingkat pengangguran terbuka (TPT) sebesar 4,91 persen dan rata-rata upah buruh sebesar 3,27 juta rupiah per bulan*. <https://www.bps.go.id/id/pressrelease/2024/11/05/2373/unemployment-rate-was-4-91-percent-.html>
- Corral, G., et al. (2021). Immediate feedback in a flipped classroom improves learning. *Journal of Medical Education and Curricular Development*, 8. <https://doi.org/10.1177/23821205211029280>
- de Ayala, R. J. (2022). *The theory and practice of item response theory* (2nd ed.). Guilford Press.
- Ebenbeck, N., Bastian, M., Mühling, A., & Gebhardt, M. (2024). Duration versus accuracy: What matters for computerised adaptive testing in schools? *Journal of Computer Assisted Learning*, 40(6), 3443–3453. <https://doi.org/10.1111/jcal.13074>
- Fahrurrizal, F., Marwan, M., & Bahri, S. (2024). Efektivitas pengelolaan pemberian tunjangan profesi dalam optimalisasi kinerja profesional guru. *Research and Development Journal of Education*, 10(2), 725–735. <https://doi.org/10.30998/rdje.v10i2.25030>
- Falani, I. (2023). Desain dan validasi aplikasi tes literasi matematika berbasis komputer dengan pendekatan item response theory. *Jurnal Pendidikan Matematika*, 12(2).
- Hadi, S. (2013). *Pengembangan computerized adaptive test berbasis web*. Aswaja Pressindo.
- Hadi, S., Haryanto, H., AM, M. A., Marlina, M., & Rahim, A. (2022). Developing classroom assessment tool using learning management system-based computerized adaptive test in vocational high schools. *Journal of Education Research and Evaluation*, 6(1), 143–155.

<https://doi.org/10.23887/jere.v6i1.35630>

Hizqiyah, I. Y. N., Widodo, A., Sriyati, S., & Ahmad, A. (2023). Development of a digital problem solving skills test instrument: Rasch model analysis. *Jurnal Penelitian Pendidikan IPA*, 9(4), 1658–1663.

<https://doi.org/10.29303/jppipa.v9i4.2671>

Kementerian Energi dan Sumber Daya Mineral Republik Indonesia. (2019). *Tenaga teknik ketenagalistrikan bersertifikat beri rasa aman menggunakan listrik*.

<https://www.esdm.go.id/id/media-center/arsip-berita/tenaga-teknik-ketenagalistrikan-bersertifikat-beri-rasa-aman-menggunakan-listrik>

Luo, H., & Yang, X. (2024). Efficiency of computerized adaptive testing with a cognitively designed item bank. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1353419>

Mao, X., Zhang, J., & Xin, T. (2022). The optimal design of bifactor multidimensional computerized adaptive testing with mixed-format items. *Applied Psychological Measurement*, 46(7), 605–621.

<https://doi.org/10.1177/01466216221108382>

Nadia, Kamila, V. Z., Islamiyah, Wardhana, R., & Setyadi, H. J. (2022). Penerapan computerized adaptive test (CAT) pada ujian semester menggunakan algoritma teori respon butir model 2PL berbasis web (Studi kasus: Kelas XI SMAN 1 Muara Jawa Kota Samarinda). *Adopsi Teknologi dan Sistem Informasi (ATASI)*, 1(1), 36–51. <https://doi.org/10.30872/atasi.v1i1.43>

Nurdiansyah, A. I., Sriyono, & Eko, S. K. (2020). Pengembangan program computerized adaptive test hasil belajar fisika SMK. *Jurnal Inovasi Pendidikan Sains*, 1(1).

Oladele, J. I. (2025). Computerised adaptive method for assessing university undergraduates' mental well-being within an African context: An open-source set-up with Concerto. *MethodsX*, 14, 103140. <https://doi.org/10.1016/j.mex.2024.103140>

Pan, X. (2022). Exploring the multidimensional relationships between educational situation perception, teacher support, online learning engagement, and academic self-efficacy in technology-based language learning. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.1000069>

Rahim, A., Hadi, S., Marlina, M., Susilowati, D., Muti'ah, M., & Irhas, I. (2024). Transfigurations of the computerized mathematics final exam: How was the results of computerized adaptive test-based assessment? *Jurnal Pendidikan Matematika dan IPA*, 15(2), 257–274.

<https://doi.org/10.26418/jpmipa.v15i2.69485>

Sáiz-Manzanares, M. C., Marticorena-Sánchez, R., Muñoz-Rujas, N., Rodríguez-Arribas, S., Escolar-

- Llamazares, M. C., Alonso-Santander, N., Martínez-Martín, M. Á., & Mercado-Val, E. I. (2021). Teaching and learning styles on Moodle: An analysis of the effectiveness of using STEM and non-STEM qualifications from a gender perspective. *Sustainability*, 13(3). <https://doi.org/10.3390/su13031166>
- Tananda, O., Nasir, M., Milana, M., & Muslim, M. (2023). Perbandingan antara sistem computer based test dan paper based test pada hasil pembelajaran mata pelajaran PMKR di SMKN 1 Sumatera Barat. *JTPVI: Jurnal Teknologi dan Pendidikan Vokasi Indonesia*, 1(1), 67–74. <https://doi.org/10.24036/jtpvi.v1i1.10>
- Thompson, N. A., & Weiss, D. J. (2021). A history of computerized adaptive testing. In B. P. Veldkamp & C. Sluijter (Eds.), *Theoretical and practical advances in computer-based educational measurement* (pp. 3–33). Springer. https://doi.org/10.1007/978-3-030-43485-4_1
- Veldkamp, B. P., & Sluijter, C. (Eds.). (2021). *Theoretical and practical advances in computer-based educational measurement*. Springer.
- Wang, Z., Wang, C., & Weiss, D. J. (2022). Termination criteria for grid multiclassification adaptive testing with multidimensional polytomous items. *Applied Psychological Measurement*, 46(7), 551–570. <https://doi.org/10.1177/01466216221108383>
- Wijaya, W. M., Syarifah, L. S., & David, T. (2024). Enhancing e-learning in vocational schools: Key characteristics of instructional design. *Jurnal Akuntabilitas Manajemen Pendidikan*, 12(1), 68–78. <https://doi.org/10.21831/jamp.v12i1.75984>
- Wulandari, F., Hadi, S., & Haryanto, H. (2020). Computer-based adaptive test development using fuzzy item response theory to estimate student ability. *Computer Science and Information Technology*, 8(3), 66–73. <https://doi.org/10.13189/csit.2020.080302>
- Wyse, A. E. (2021). How days between tests impacts alternate forms reliability in computerized adaptive tests. *Educational and Psychological Measurement*, 81(4), 644–667. <https://doi.org/10.1177/0013164420979656>
- Xi, C., Tu, D., & Cai, Y. (2022). Dual-objective item selection methods in computerized adaptive test using the higher-order cognitive diagnostic models. *Applied Psychological Measurement*, 46(5), 422–438. <https://doi.org/10.1177/01466216221089342>
- Xiao, J., & Bulut, O. (2022). Item selection with collaborative filtering in on-the-fly multistage adaptive testing. *Applied Psychological Measurement*, 46(8), 690–704. <https://doi.org/10.1177/01466216221124089>

Email : joves@mpv.uad.ac.id

Website : <http://journal2.uad.ac.id/index.php/joves>

Yang, A. C. M., Flanagan, B., & Ogata, H. (2022). Adaptive formative assessment system based on computerized adaptive testing and the learning memory cycle for personalized learning.

Computers and Education: Artificial Intelligence, 3. <https://doi.org/10.1016/j.caeai.2022.100104>

Yuniarti, N., Hariyanto, & Zamtinah. (2024). *Media pembelajaran vokasional*. UNY Press.