Fairness in Learning Evaluation of Database Management Course

Shinta Yulia Puspitasari, Prati Hutari Gani, and Villy Satria Praditha
School of Computing, Telkom University Bandung, Indonesia
Corresponding Author's Email: shintayulia@telkomuniversity.ac.id

ABSTRACT

BACKGROUND & CONTEXT

Evaluation is essential for measuring the effectiveness of the learning process and student competency achievement. At the Faculty of Informatics, Telkom University, assessments are categorized into Participatory Activities (PA), Cognitive/Knowledge (CK), and Projects (PJ). This study focuses on the Database Management course, in which significant differences in grades across 13 parallel classes were detected during the 2022/2023 and 2023/2024 academic years. These discrepancies were linked to variations in assessment methods and the parameters used by different lecturers, leading to inconsistencies in evaluating student performance.

PURPOSE OR GOAL

The main objective of this research is to develop a comprehensive learning evaluation system that addresses all aspects of assessment categories while incorporating fairness principles. Additionally, the study aims to evaluate the proposed system and its implementation in the Database Management course during the 2024/2025 academic year.

APPROACH OR METHODOLOGY

This study employed the Analysis, Design, Development, Implementation, and Evaluation (ADDIE) model to develop a fair learning evaluation system. Its effectiveness was examined through statistical analysis of student learning outcomes, lecturer Focus Group Discussions (FGDs), and student feedback gathered through a student satisfaction questionnaire.

OUTCOMES

Quartile values Q1 (60.83), Q2 (67.32), and Q3 (74.11) indicated that the final scores follow a normal distribution. The average scores for project (85.76) and participatory activities (77.96) were higher than cognitive scores (58.56), while midterm exams revealed the widest spread and lowest scores, suggesting mid-semester challenges that eased by the final exam. Student satisfaction was 96.46% for assessment objectivity and 95.88% for alignment of assessments with learning outcomes.

CONCLUSIONS

This study successfully developed and implemented a fair learning evaluation system encompassing Participatory Activity, Cognitive Knowledge, and Project categories, aligned with the nationwide Key Performance Indicator (IKU-7). Student grades followed a normal distribution with clearly interpretable assessment indicators, and lecturers applied the process

KEYWORDS

Learning evaluation, fairness, assessment category, IKU-7, normal distribution.

https://doi.org/10.52202/086293-0002

Introduction

In higher education, learning evaluation is essential for measuring the achievement of student learning outcomes (Aiman Faiz et al., 2022). However, the objectivity of this process can be compromised if test instruments, measurement procedure, and assessments method do not account for fairness. Fairness is a condition where each student is given material with the same learning content, structure, test instruments, and uses measurement procedures and assessment methods with the same standards. Fairness is equal treatment and unbiased process (Tierney, 2016).

The Informatics Undergraduate Program at Telkom University has around 13 parallel classes per year. When multiple classes study the same subject under different lecturers without standardized evaluation strategies, grade disparities and bias can arise. This issue was evident in the Database Management course, where during the 2022/2023 and 2023/2024 academic years, significant grade gaps were observed across the 13 parallel classes. As illustrated in Figure 1 and Figure 2 the inconsistency in grades final scores is primarily attributed to the lack of standardized assessment methods. Each lecturer independently designed exam questions and conducted assessments, resulting in varying evaluation practices across classes.

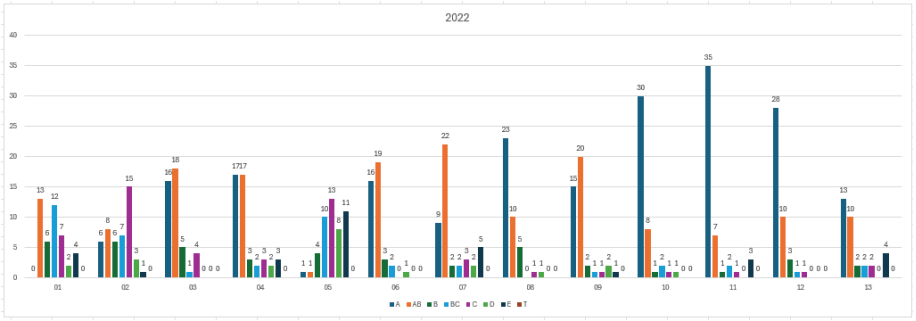


Figure 1: Comparison of final grades in 2022/2023

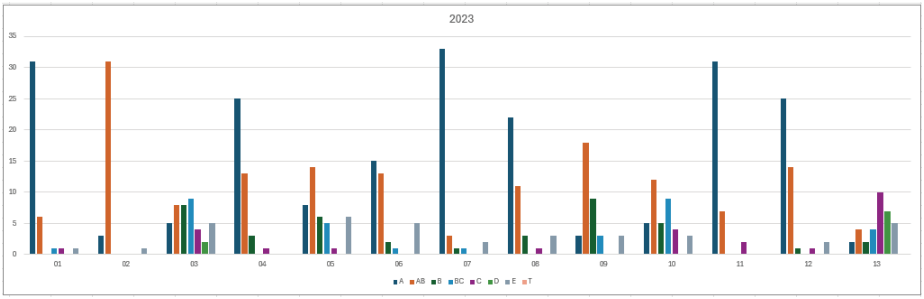


Figure 2: Comparison of final grades in 2023/2024

In conducting assessments, Telkom University follows the Ministry of Education regulations, specifically the Higher Education Key Performance Indicators (IKU), including IKU-7, which requires student evaluations to cover Participatory Activity, Cognitive/Knowledge, and Project categories. Despite having the same learning material structure and the same assessment indicator references, each lecturer of 13 parallel classes conducted assessments individually using different assessment instruments and exam questions. With a grade range of A, AB, B, BC, D, E, and T (i.e. missing component), one class obtained 97.14% of A and AB (success

rates), while another class only 17.65%. Six of the 13 classes had total A and AB scores >90%, three classes <50%, and the rest ranged from 55% to 82%. For grades D and E (failure rates), one class accumulated 35%, but in six classes the failure rate was below 5%, with two of them was as low as 0.00%. Although the overall average across the parallel classes was satisfactory (>70%), the score distribution varied widely between classes.

Problem Statement

The evaluation methods in the 13 parallel classes of Database Systems course, each taught by different lecturers, did not adequately ensure fairness. The lack of a standardized and holistic assessment framework led to grading disparities and inconsistent evaluation practices. To resolve this, research is needed to develop a comprehensive learning evaluation system that accommodates parallel class conditions, complies with IKU-7 assessment guidelines, and ensures accurate, consistent, and fair measurement of learning outcomes.

Theoretical Framework

Learning Evaluation

Learning evaluation is crucial in educational process, not only to measure achievement but also to foster students' motivation in their learning. For this reason, fairness in the evaluation system is especially important. Tierney (2016) states fairness in educational evaluation involves clear criteria, opportunities for students to correct mistakes, and consistency among evaluators. It ensures that each student receives the same learning content, test instruments, and is assessed using identical procedures and standards. This emphasizes that fair evaluation is not merely a technical matter of assigning grades but also relates to students' perceptions of the lecturer's objectivity. However, in practice, many evaluations still focus primarily on final outcomes, without considering the extent to which the assessment process is truly perceived as fair by students.

According to (Faiz et al., 2022), field of education recognizes four interrelated components: testing, measurement, assessment, and learning evaluation. Testing refers to tools comprising tasks or questions designed to measure specific student behaviors. Measurement involves assigning numerical values to the learning process or outcomes based on established standards. Assessment interprets these measurements to support decision-making, while evaluation is the final step, involving judgments about the effectiveness of learning processes. These components are hierarchical and interconnected, forming a continuous cycle in educational practice. When mapped, fairness as described by Tierney (2016) needs to be present at every stage: clear testing instruments (testing), objective scoring (measurement), consistent interpretation of results (assessment), and transparent final judgment (evaluation). Thus, fairness and the evaluation cycle together form a comprehensive and equitable evaluation system.

Key Performance Indicators

The concept of evaluation fairness (Tierney, 2016) requires clear standards, which in Indonesian higher education are embodied in Key Performance Indicators (KPIs). The KPIs published by the Directorate General of Higher Education (2020) provide a performance measurement framework that is more adaptive and emphasizes concrete results. To ensure accurate learning evaluation aligned with national education objectives, Telkom University integrates these standards, particularly KPI #7 (IKU-7), which focuses on collaborative and participatory learning. This approach divides learning evaluation into three categories: Participatory Activities (PA), Cognitive/Knowledge (CO), and Project (PJ)

Analyze, Design, Develop, Implement, Evaluate (ADDIE) Model

To ensure that the learning evaluation system is both meaningful and systematically implemented, this study uses the ADDIE Model approach as a methodological framework to

https://doi.org/10.52202/086293-0002

develop a comprehensive learning evaluation method that addresses fairness. This model was chosen for its systematic and flexible approach to instructional design, which enables the learning evaluation process to be well-planned, structured, and effective (Almomen et al., 2016; ELM Learning, n.d.; Adeoye et al., 2024). Figure 3 shows the Addie model framework.

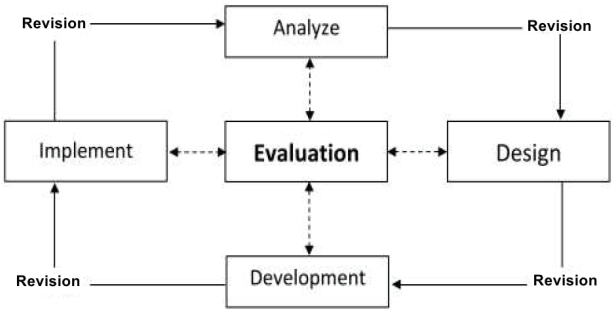


Figure 3: ADDIE Model (Fitria & Muhamad, 2021)

Boxplot

To support the evaluation phase in the ADDIE model, boxplots were used to identify patterns and fairness in student performance. Boxplots summarize data using five key statistics: minimum, first quartile (Q1), median (Q2), third quartile (Q3), and maximum. They indicate central tendency, variability, and potential outliers without assuming normality (Xingrong Sun, 2021). Figure 4 shows a rectangular box that stretches from Q1 to Q3, representing the middle 50% of the data. Outliers beyond this range appear as individual points, signaling anomalies that may indicate scoring inconsistencies. The size of the box in the boxplots reflects the variation in student achievement and pedagogically can guide evaluation of consistency of understanding as well as inform the need for differentiated instructional strategies.

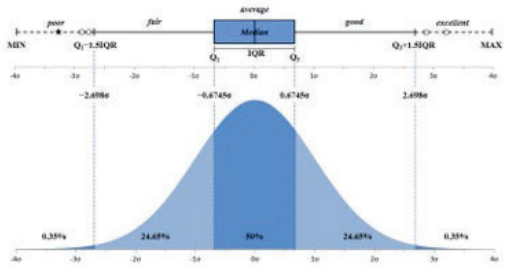


Figure 4: Boxplot

In contrast, Figure 5 illustrates left- and right-skewed distributions using both box plots and probability density curves. Skewness indicates asymmetry in data distribution. In this study, boxplot visualizes how student scores are distributed across different evaluation methods. It allows researchers to detect imbalances in grade distribution, such as clusters of particularly low or high scores, and to identify extreme values that may require further investigation. Pedagogically, understanding skewness is essential for interpreting whether an assessment fosters equitable learning outcomes. A left skew suggests generally strong performance, while a right skew indicates widespread difficulty or misalignment between instruction and assessment design. Identifying such trends helps educators refine teaching strategies, address learning gaps, and ensure fairness in grading practices.

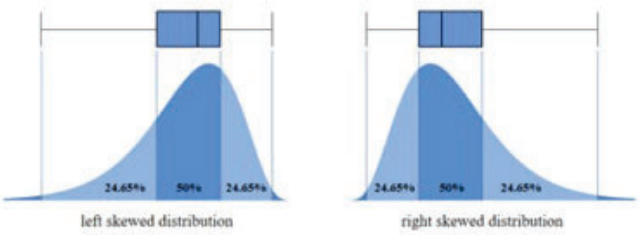


Figure 5: Comparison of left skewed and right skewed distribution

Methodology: ADDIE Model

Analyse

The first stage of ADDIE is Analysis (Adeoye et al., 2024) which identifies requirements for learning evaluation process that ensures fairness and aligns with the IKU-7 assessment standards. The activities in this stage to determine evaluation components are:

1) Organizing materials by expected competencies. With curriculum change in 2024/2025, the structure of materials and the weighting of Course Learning Outcomes (CLOs) must be adjusted. Previously in 2022/2023 and 2023/2024, the Database Management course had three CLOs; in 2024/2025, there are four CLOs.
2) Identifying factors that could cause unfairness and analysing historical grade data. Exploratory data analysis was conducted on the grades supplemented by interviews with course coordinators. The analysis revealed that: (a) lecturers independently designed exam questions and conducted assessments, resulting in inconsistent evaluation practices across classes; (b) grading practices among lecturers varied, with some lecturers being more lenient in awarding high grades and others stricter, contributing to disparities among parallel classes.
3) Defining assessment types and media for each CLO and evaluation aspects i.e. Participatory Activities (PA), Cognitive/Knowledge (CO), and Project Outcomes (PJ). The University employs blended learning, so it is necessary to determine which assessments are conducted conventionally and which use LMS media (CELOE).
4) Specifying fairness aspects for each test instrument. Assessment methods must align with the type of test, media used, and the aspects being assessed to ensure fairness throughout the learning evaluation process.

Design

After the analysis phase, the design phase focuses on developing evaluation methods aligned with the learning objectives and contextual needs defined in the analysis stage. In this phase, assessment weights were assigned to each learning outcome based on the depth and complexity of the material. These weights were then mapped to the test instrument for each IKU-7 assessment category (PA, CO, PJ) as shown in Table 1. From the test instrument, the frequency, number of attempt, duration, and assessment media were determined, as presented in Table 2. The complete learning evaluation plan was compiled into RPS (Student Learning Plan) document, which details the material structure, learning strategies, and evaluation strategies.

Table 1: Mapping Indicator, Test Instrument, and Scoring Weight

Category	Description	Test (Weight %)
(PA) Participatory Activities	Class engages in active, student-driven discussions, with lecturers as facilitators: guiding conversations, posing questions, and observing. Students are expected to analyze	Forum (12%) Individual Assignment (16%)

https://doi.org/10.52202/086293-0002

	cases to develop recommended solutions, often supported by group discussions to test and refine their ideas.	
(CO) Cognitive/ Knowledge	In this approach, students act as the 'protagonists' of the learning process, taking primary responsibility for analyzing and resolving the case.	Quiz/ Post-Test (12%) Mid-Term Exam (20%) Final Exam (25%)
(PJ) Project	Classes are divided into groups to collaboratively work on a long-term assignment. Each group is presented with real problem or a complex question and is tasked to develop a work plan and a collaborative model. At the end of the project, groups present a product where audiences can provide constructive feedback	Team-Based Project (15%)

Table 2: Assessment Strategy Development

Category	Test Instrument	Frequency	Attempt & Duration	Media
PA	Forum	Every week	Unlimited attempt while forum is still open	LMS
PA	Individual Assignment	One for each Sub-CLO	Once with duration of work is 1 week	LMS
CO	Quiz/ Post-Test	One for each Sub-CLO	Twice (take the highest score), with duration of work is 1 week	LMS
CO	Mid-Term Exam	Once in mid semester	Once with duration of work is 2 hours	Manual/ Paper based
CO	Final Exam	Once at the end of semester	Once with duration of work is 2 hours	LMS
PB	Team-Based Project	Twice in one semester	Once with duration of work is 2-3 weeks	LMS

Develop

Following the design phase, the development stage implements the Course Learning Plan (RPS) into functional products for learning evaluation. Activities include creating assessment rubrics, compiling question banks for assignments, exams, and mini-projects, and preparing student feedback tools. Since a blended learning system is used, some assessments are conducted manually (e.g., midterm exams), while others are managed through the e-learning platform (CELOE).

To maintain fairness in the learning evaluation system, several strategies are applied:

- **Question compilation**: The coordinating lecturer collaborates with all class lecturers to gather suggestions. Questions are selected based on alignment with course learning outcomes, difficulty level, and weighting in accordance with the RPS.
- **Rubric development**: The coordinator prepares answer rubrics, which are validated collectively by the lecturer team. This ensures unified assessment standards and consistent interpretation across all parallel classes.
- **Assessment scoring (measurement)**: Each class lecturer assesses their own students' activities through observation for PA and PB indicators. For CO indicators (essay-based Midterms), cross-review is conducted, with each lecturer grading all students' answers to ensure consistent measurement standards. Additionally, for other CO indicators (non-essay, quizzes, Final Exam), an automated grading through LMS features can be used.

The summary of the fairness strategy is presented in Table 3.

Table 3: Fairness strategy

Test Instrument	Rubric	Measurement by
Forum (PA)	Ratings (Rubric)	Class Lecturer
Individual Assignment (PA)	Advance Grading (Rubric)	Class Lecturer
Quiz/ Post-Test (CO)	-	Automation Grading
Mid-Term Exam (CO)	Manual Rubric	Cross Review: all parallel lecturers
Final Exam (CO)	-	Automation Grading
Team-Based Project (PB)	Advance Grading (Rubric)	Class Lecturer

Implement

The implementation stage applies the developed evaluation system in the classroom over one semester (Adeoye et al., 2024). During this phase, the system is integrated into practical teaching, while data are collected on its application. Student responses are observed to assess how well the system provides fair opportunities for all learners.

Evaluate

This stage aims to identify areas for revision and improvement in the learning evaluation system (Adeoye et al., 2024). Both quantitative and qualitative analyses are conducted. **Qualitative analysis:** Student perceptions of fairness, clarity, and transparency are gathered using the student satisfaction questionnaire questionnaire, administered each semester as part of routine evaluations. Lecturers also conduct peer and self-reflection to assess the effectiveness of learning and assessment processes. Feedback is collected on teaching facilities, clarity of rubrics, student engagement, and achievement of learning outcomes, often through Focus Group Discussions (FGDs) held mid-semester and at semester's end. **Quantitative analysis**: Data from student performance and assessment outcomes are analyzed to measure the effectiveness and consistency of the implemented system.

Results and Discussion

Figure 6 shows the distribution of final grades across three academic years. In 2022/2023 shows the widest interquartile range (IQR = 19,20) and the widest overall range (min=37,12, max=93,09), with median (75,52) positioned closer to Q3 (82,60). This pattern suggests a left skewed distribution, indicating higher variability in student performance, with some students scoring significantly lower than the others. In 2023/2024, variability decreased, although its IQR (13.44) remains larger than that of 2024/2025. The median (78.29) is still closer to Q3 (83.82), showing a similar skewness pattern to the previous year, albeit with more controlled dispersion. Outliers remain frequent. However, in 2024/2025 the distribution was the most balanced. The smallest IQR (13.31) and the median (67.34) are of equal distance from Q1 and Q3, indicating a more symmetrical distribution. The high variability observed in 2022/2023 and 2023/2024 was likely due to the lack of a standardized exam design, which resulted in differences in question difficulty across classes. In contrast, the implementation of a standardized evaluation system in 2024/2025 successfully improved consistency, albeit with lower average scores. This suggests a more objective assessment approach.

Contrastingly, Figure 7 illustrates grade distribution across three assessment categories in 2024/2025. Both the Participatory Activities and Project Outcomes average scores were above 70. This trend suggests higher motivation in collaborative and participatory activities than in assessments requiring independent cognitive effort. However, since these categories were assessed by class lecturer Table 3 potential subjectivity exists. In contrast, the Cognitive category had a significantly lower average score of 58.57, indicating greater challenge in

https://doi.org/10.52202/086293-0002

independent cognitive tasks. The lower performance can be attributed to the fact that this category was evaluated collectively by all instructors through a cross-review process, supplemented by automated grading Table 3, which minimized subjectivity and ensured a more objective and precise assessment. Additionally, the Cognitive component carried the highest weighting among all assessment categories, which amplified its influence on overall student performance.

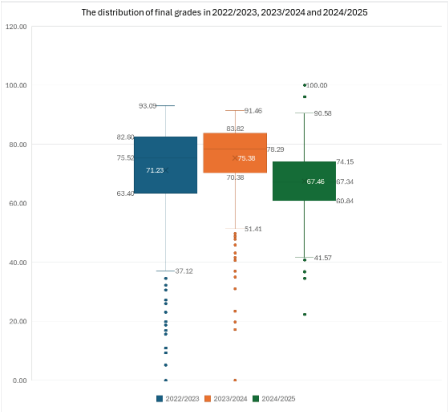


Figure 6: Comparison of grades distribution across 2022/2023, 2023/2024 and 2024/2025

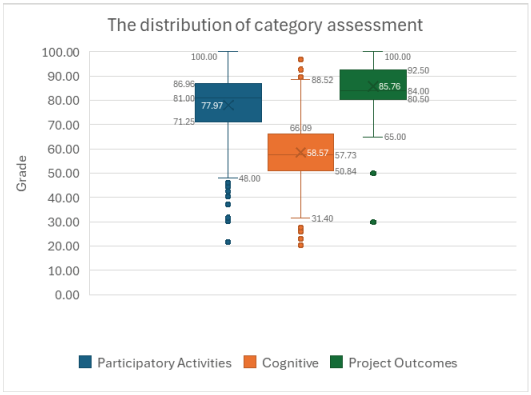


Figure 7: Comparison of assessment categories

This study also revealed a notable pattern: test instruments that allowed multiple attempts and extended completion time Table 2 tended to yield higher average scores than those conducted once within a limited time frame of a few hours. As shown in Figure 8, assessments such as Forum, Individual Assignment, Quiz/Post-Test, and Team-Based Project demonstrated higher average scores compared to the Midterm and Final Exams. This finding suggests that providing flexibility in assessment duration and opportunities for repeated attempts can enhance student performance by reducing time pressure and promoting deeper engagement with the material.

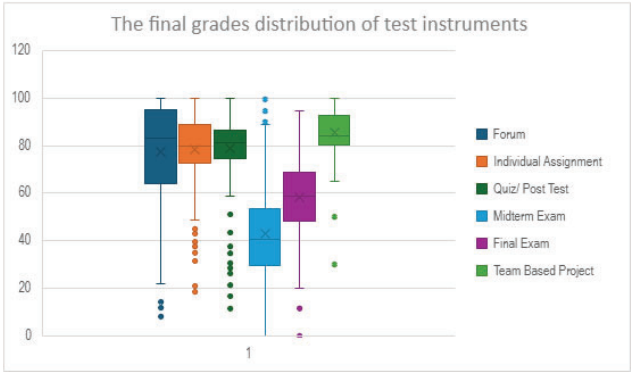


Figure 8: Comparison of test instrument

presents the distribution of student responses regarding their satisfaction with the objective assessment and feedback, with the majority expressing positive views. Of the 510 respondents, 59.02% (301 students) reported being Satisfied and 36.86% (188 students) were Very Satisfied. In contrast, only a small fraction expressed dissatisfaction—3.73% (19 students) were Dissatisfied and just 0.39% (2 students) were Very Dissatisfied. Overall, 95.88% of students indicated satisfaction, highlighting a strong approval of the fairness and

clarity of the assessment and feedback processes, and suggesting a high level of alignment between the assessment approach and student expectations.

Table 4: Objective assessment and feedback

Level of Satisfaction	Count
Satisfied	301
Very Satisfied	188
Very dissatisfied	2
Dissatisfied	19
Grand Total	510
Satisfied and Very Satisfied Percentage	95.88%

Table 5 then summarizes student perceptions of the alignment between assessment methods and intended learning outcomes, revealing overwhelmingly positive feedback. Among 509 respondents, 59.14% (301 students) reported being Satisfied and 37.33% (190 students) were Very Satisfied. Only a small minority expressed dissatisfaction, with 2.75% (14 students) were Dissatisfied and 0.79% (4 students) were Very Dissatisfied. Altogether, 96.46% of students expressed satisfaction, indicating a strong approval that the assessments were appropriate and effectively aligned with the intended learning outcomes. This suggests that the evaluation methods successfully measured what students were expected to learn.

Table 5: Appropriate assessment to learning outcomes

Level of Satisfaction	Count
Satisfied	301
Very Satisfied	190
Very dissatisfied	4
Dissatisfied	14
Grand Total	509
Satisfied and Very Satisfied Percentage	96.46%

On December 20, 2024, all lecturers of Database Management course met to reflect on the teaching and learning process. They shared feedback and suggestions, especially about assessment methods and the use of rubrics. Key points included that forums were less effective during face-to-face classes and would work better during online weeks. It was also suggested that remedial opportunities be offered not only at the end of each CLO but also after specific assessments like midterms or finals. Positive feedback highlighted the benefits of cross-lecturer grading and the consistent use of rubrics for fair evaluation. While the grading system was considered appropriate, it was recommended to spread out assignments more evenly throughout the semester. Lecturers also suggested adding more practice exercises, preferably within presentation slides, to help improve student understanding.

Conclusions and Recommendations

This study successfully developed and implemented a learning evaluation system that emphasizes fairness. Results showed that student learning outcomes grades were normally distributed, and the assessment format for each indicator was interpretable. Each lecturer conducted the learning evaluation with consistent standards and structured rubrics. This ensures alignment between assessment categories and expected learning outcomes, minimizes inequality, and incorporates objective grading mechanisms. Its implementation in the Database Management course during the 2024/2025 academic year demonstrated the feasibility of applying fairness principles by standardizing assessment design, providing balanced weight distribution, and integrating automated grading tools alongside instructor cross-review. The evaluation results indicate that the proposed system not only improved

https://doi.org/10.52202/086293-0002

consistency across classes but also offered students equitable opportunities to demonstrate their competencies, thereby fulfilling the research objective of creating a fair and comprehensive learning evaluation framework.

Based and feasibility and anticipated impact, the recommendations are proposed: (1) maintain and enhance the standardized evaluation system as it promotes consistency and fairness across parallel classes, despite a slight decrease in average scores; (2) redesign of cognitive assessment using interactive tools such as digital quizzes, adaptive assessment, scenario-based formative task, and evaluation with immediate feedback to facilitate remedial; (3) Participatory Activities and Project Outcomes category should remain central while integrating reflective components that connect practical application with theoretical concepts. This enhances conceptual understanding without reducing the practical learning; (4) conduct periodic analysis of student grade distribution and feedback to promote fairness and ensure sustainability in evaluation.

Future research should focus on repeated observations to examine the ongoing impact of a standardized evaluation system on equity and academic achievement. Experimental studies could also explore the integration of adaptive learning technologies and personalized feedback mechanisms to address persistent cognitive challenges.

References

Adeoye, M. A., Wirawan, K. A. S. I., Pradnyani, M. S. S., & Septiarini, N. I. (2024). Revolutionizing Education: Unleashing the Power of the ADDIE Model for Effective Teaching and Learning. *JPI (Jurnal Pendidikan Indonesia)*, *13*(1), 202–209. https://doi.org/10.23887/jpiundiksha.v13i1.68624.

Almomen, R. K., Kaufman, D., Alotaibi, H., Al-Rowais, N. A., Albeik, M., & Albattal, S. M. (2016). Applying the ADDIE—Analysis, Design, Development, Implementation and Evaluation—Instructional Design Model to Continuing Professional Development for Primary Care Physicians in Saudi Arabia. *International Journal of Clinical Medicine*, *07*(08), 538–546. https://doi.org/10.4236/ijcm.2016.78059.

Elm, L. (n.d.). *What is ADDIE? Your Complete Guide to the ADDIE Model*. Https://Elmlearning.Com/Hub/Instructional-Design/Addie-Model/.

Faiz, A., Putra, N. P., & Nugraha, F. (2022). Memahami Makna Tes, Pengukuran (Measurement), Penilaian (Assessment), Dan Evaluasi (Evaluation) Dalam Pendidikan. *Jurnal Education and Development*. https://doi.org/10.31004/obsesi.v5i2.972.

Fitria, H., & Muhamad, N. (2021). *Model ADDIE (Analysis, Design, Development, Implementation And Evaluation) Dalam Pembelajaran Pendidikan Agama Islam*.

Tierney, R. D. (2016). Fairness in Educational Assessment. In *Encyclopedia of Educational Philosophy and Theory* (pp. 1–6). Springer Singapore. https://doi.org/10.1007/978-981-287-532-7_400-1.

Acknowledgement

The authors would like to thank Telkom University especially School of Computing for providing the facilities and all technical support to complete this research. This work was supported by Academic Development Division of Telkom University, under the Classroom Action Research Grant in fiscal year 2024-2025. as well as by additional financial support from the Directorate of Research and Community Service. We also extend our gratitude to Database Management team teaching of 2024/2025, and our five students who served as research assistant over the past year. Finally, we sincerely thank our reviewers for their constructive comments and suggestions that significantly improved the quality of this manuscript.

Copyright Statement

Copyright © 2025 Shinta Yulia Puspitasari, Prati Hutari Gani, and Villy Satria Praditha: The authors assign to the Research in Engineering Education Network (REEN) and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full, and this copyright statement is reproduced. The authors also grant a non-exclusive license to REEN to publish this document in full on the World Wide Web (prime sites and mirrors), on portable media and in printed form within the REES and RCEE & RHEd 2025 proceedings. Any other usage is prohibited without the express permission of the authors.