Analysis of practical and test-based evaluations

Alin ZAMFIROIU[1,2], Iris REYCHAV[3], Roger MCHANEY[4], Daniela TIMISICĂ[1,5]

[1] National Institute for Research & Development in Informatics - ICI Bucharest, 8-10, Maresal Averescu Avenue, sector 1, Bucharest, Romania

[2] The Bucharest University of Economic Studies, 6 Piata Romana, 1st district, Bucharest, 010374 Romania

[3] Department of Industrial Engineering and Management, Ariel University, Ariel, Israel

[4] Management Information Systems, Kansas State University, Manhattan, Kansas, USA

[5] Politehnica University of Bucharest, 313 Independence Av., Bucharest, 060042, Romania

alin.zamfiroiu@ici.ro, irisre@ariel.ac.il, mchaney@ksu.edu, daniela.timisica@ici.ro

Abstract: *This paper explores the correlation between different teaching evaluation methods, specifically 3 practical assessment formats and quizzes, and related impact on students' academic performance. The study draws from educational theory, using four significant stages: teaching, learning, evaluation, and feedback. The feedback phase is a crucial aspect of education that is dependent on the evaluation process used. The paper investigates various evaluation methods, including written examinations, practical projects and labs, test/quizzes, collaborative assessments, and portfolio-based assessments. Statistical analysis is conducted on a dataset of students' evaluations to determine which assessment methods yielded optimal results. The study suggests that quizzes tend to produce narrower score intervals and lower variances compared to practical assessment approaches. Correlations between evaluation methods are analyzed, and a significant correlation is observed between practical assignments and quiz performance. The findings contribute to a deeper understanding of evaluation methods in education and provide insights into selecting appropriate assessment approaches to improve students' learning outcomes.*

Keywords: educational environment, learning, evaluation, quizzes, assessments, evaluation methods.

1. Introduction

Educational systems can be viewed as four significant stages (Tuychieva and Xudoyorov, 2022), (Valieva, 2022) which include (see Figure 1):

- **teaching** – by teachers;
- **learning**- by students;
- **evaluation** – by teachers for students;
- **feedback** – by teachers for students.

https://doi.org/10.58503/icvl-v18y202315

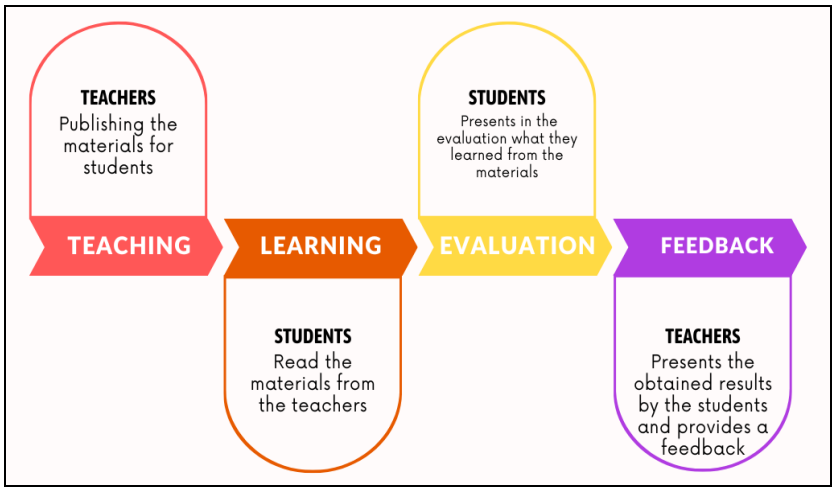

Figure 1. Linear Process of Learning

Each of these four phases is important to ensure a student moves through essential material, absorbs key elements, and demonstrates mastery. Teacher feedback reinforces learning and ensures any missing elements are communicated back to the learner. Within this process, evaluation of learning and feedback become key elements for an effective teacher. Depending on the evaluation method, various feedback approaches can be used for each student. For example, feedback may be provided for each question or in summary for related groups of questions. It may be oriented to individuals or to multiple students who work in a team. To facilitate learning, it becomes important to choose the correct method of evaluation for various learning objectives. This will impact the manner in which feedback will be provided by the teacher.

In this paper, we evaluate the third phase of this four-stage process: evaluation. Evaluation can utilize multiple approaches and methods. We examine 4 of these and then statistically analyze effectiveness to help teachers determine the appropriateness of each.

2. Background of evaluation methods

Student evaluation in a technological field can take place according to a variety of method. Several current approaches include the following (Baigi et al., 2022; Mohan, 2023). See Figure 2.

- **Written examinations** or **practical assessments**. Traditional assessment methods involve evaluating students' practical knowledge in areas such as programming, algorithms, data structures, or fundamental computer science concepts. For such assessments, the teacher must create requirements in a way that the student presents his acquired knowledge in a practical way (Paiva, et al., 2022). For this method of assessment the

teacher will assign deadlines and wait for the students to submit their work. The teachers reciew the submissions and provide feedback. The feedback will be provided to evaluate exactly what a student submits, and will be personalized.

- **Practical projects and labs**. Assessment through practical projects and labs allows students to apply knowledge and skills in real projects or practical scenarios. This may involve developing software, creating applications, configuring and administering systems, or solving practical IT problems (Bunse et al., 2022). Evaluation of practical projects assumes that the students are present physically or virtually. Feedback can be provided verbally or in recorded messages.
- **Quizzes**. Quizzes or tests can be used to assess students' theoretical knowledge and skills as well as practical knowledge by providing concrete examples and motivating the student to choose an optimal decision or approach. On platforms such as Moodle, quizzes are easy to create and can be used to evaluate students (Freitas et al., 2016; Lopez-Tocon, 2021). This method of the evaluation requires the teacher to prepare feedback before the evaluation phase. Generally, the teacher prepares feedback for each question and students see feedback depending on the answer they select. The advantage is that the teacher will prepare everything before the evaluation, and after the evaluation phase no additional work is required. The disadvantage is that feedback will be general and similar for all students.
- **Collaborative assessment**. Teamwork and collaboration are important aspects of technical fields. Therefore, collaborative assessment is relevent and often used as an approach for assessing learners' ability to work in a team, communicate, and collaborate with other team members (McConnell, 2002). The feedback for this type of evaluation is provided to the entire team. Students often are responsible for extracting relevant portions of feedback to better understand their strengths and weaknesses. Teachers see the high level result of deliverables so they collectivedly evaluate all work from team members.
- **Portfolio-based assessment**. Digital portfolios can be used to collect and assess projects, source code, documentation, and other work developed by students during courses or projects (Sanjaya et al., 2022). This method is important because the evaluation represents a collection of work developed by the students over a longer time period. Feedback is usually provided periodically and then culiminates at the end of the semester. Sometimes, this means that students will not have a final opportunity to change their deliverables. In other situations, a project can be requested by the teachers in more phases, and feedback will be for provided following each phase of the project before incorporation into a portfolio.

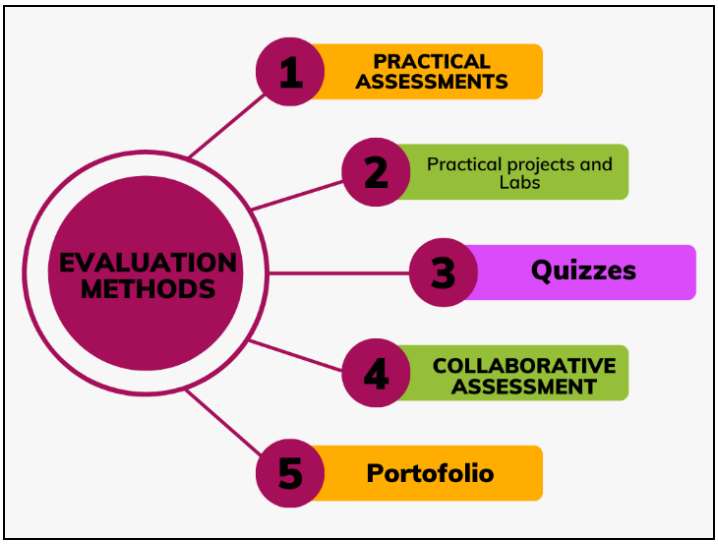

Figure 2. Evaluation Methods

3. Data sets and methodologies

Various evaluation methods are more effective under different circumstances. We analyze a sample of students to examine several methods of assessment. This is intended to allow us to observe which assessment method seems to encourage students to perform better.

3.1 Data and subjects

Subjects in this study were technology students enrolled in a programming class. An overall course grade was assessed according to the breakdown shown in Table 1.

Table 1. Weights for evaluation

Laboratory activity	10%
Practical assignment	20%
Project	10%
Exam - Quiz	60%

The first three elements are practical assessments which take place during the semester, and the final exam is assessed as a quiz. The purpose of this study is to observe the degree to which students score better on practical assessments or quizzes. This evaluation process is represented graphically in Figure 3.

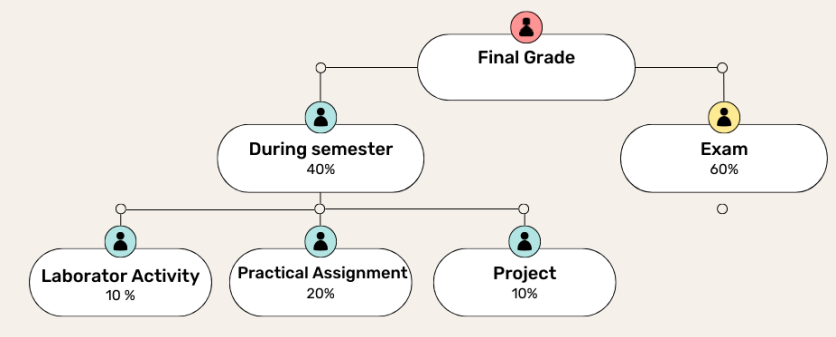

Figure 3. Evaluation process with weights

The study included 73 students divided into 5 groups. All 5 groups have the same content presented and the same evaluation approach. All students were in their 4th year and had a similar background in programming. The current course covered mobile device programming using the Java programming language.

The 4 types of evaluation used for this study can be considered as part of Self-Regulated Learning (Steinherr & Vay, 2023). This method applies because students develop their projects and learn by applying theoretical concepts to their project.

Evaluation through the quizzes was primarily a theoretical evaluation but some practical questions were included to ensure students were able to apply concepts in a practical way.

3.2 Applied analysis approach

The study examined 4 evaluation items. These included laboratory activities, practical assignments, projects and exam/quizzes. The portfolio and group approaches were not included. An average was calculated for each approach used and these were compared to determine the impact of each form of assessment. To observe the representativeness of the scores obtained by the students, standard deviation and variance were calculated for each respective evaluation approach.

Finally, we presented the correlation between the 4 types of student assessment. This allowed us to see the degree of student engagement correlated with the grades obtained in the other elements.

4. Results

Student grades are presented in Table 2. The first three assessment elements indicate the minimum student grade was 0 and the maximum was 10. This shows some students were very involved in the laboratory and practical work during the semester but other students did not get involved in this activity at all. This is normal for this course.

Table 2. Minimum and maximum grades for each method of evaluation

	Laboratory activity	Practical assignment	Project	Exam - Quiz
Minimum	0.00	0.00	0.00	4.50
Maximum	10.00	10.00	10.00	9.42

For the exam category, the minimum and maximum were 4.5 and 9.42. This shows that the evaluation through Quizzes reduced the interval for the obtained degree. To observe this conclusion in a statistical way we calculated the mean, standard deviation, and variance for all values in each evaluation method. The results are presented in Figure 3.

The mean was highest for the project evaluation at 8.51. The smallest means were obtained for the lab activity and practical assignment at 6.08 and 6.11. This indicates that when students have more time to resolve tasks, their outcomes are more likely to have been prepared more in detail.

We examined the standard deviation and variance. The interval of values was smaller for the exam category. That suggests that the quiz evaluation reduced the variance and standard deviation. It also helped increase the obtained mean. The largest variance was for the practical assignment at 6.12. This was determined to be due to lack of student prepartion. Values are summarized in Figure 4.

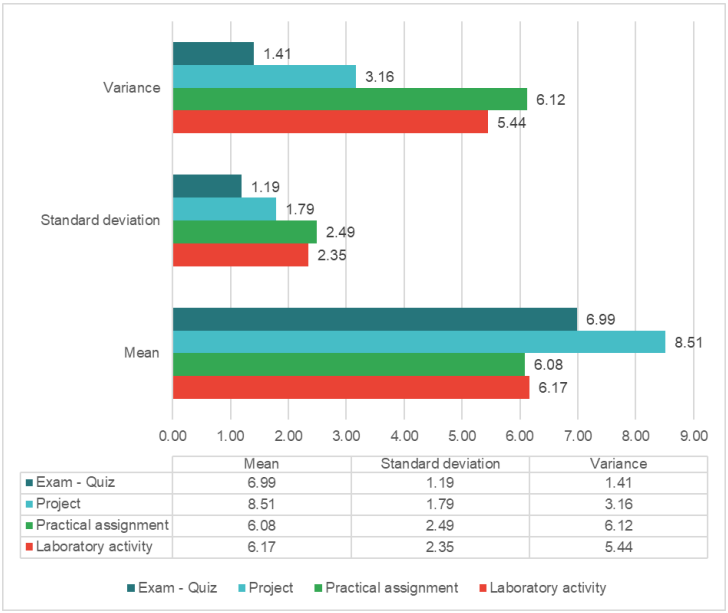

Figure 4. Mean, standard deviation, and variance

In the next phase of our research, we calculated correlations between grades obtained for each method of evaluation with other methods. For that, we created a matrix. See Table 3.

Table 3. Correlation for all students' degrees

Correlation	Laboratory activity	Practical assignment	Project	Exam - Quiz
Laboratory activity	1	0.37	0.05	0.21
Practical assignment		1	0.15	0.34
Project			1	0.11
Exam - Quiz				1

These correlations are presented graphically in Figure 5.

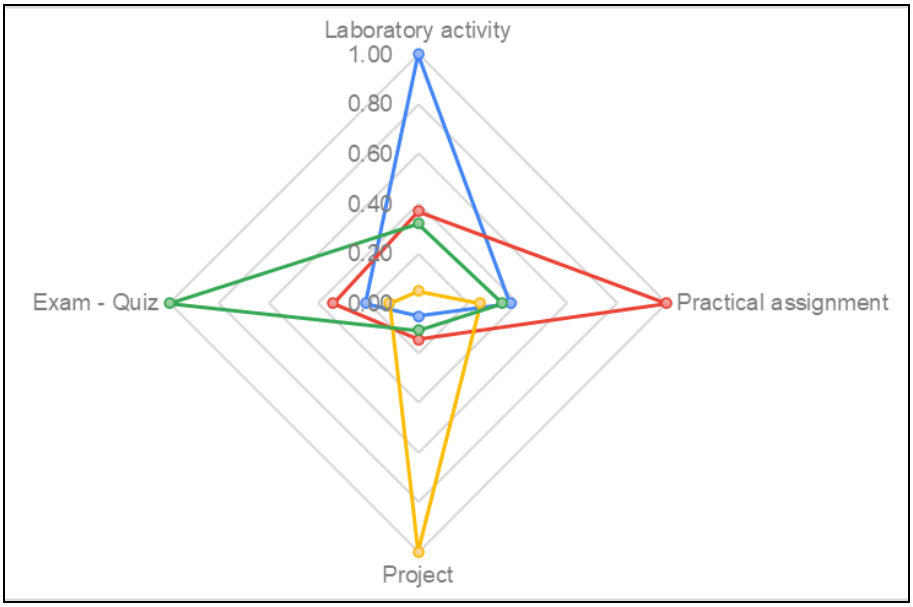

Figure 5. Correlations between evaluation methods

In Figure 5 we see the smallest correlation was between the project method evaluation and the other three methods because all three points are in the center of the graph. This shows that if the student has more time to resolve an assigned task he can improve the project. This means that results are different for those obtained in a practical assessment or in a quiz.

To further invesigate the correlation between the practical assignment and the quiz form of the evaluation, we applied a t-Test and Z-Test. These outcomes of these analyses are presented in Table 4.

Table 4. t-Test and z-Test applied for Practical assignment and Quiz

	Practical assignment	Exam - Quiz
Mean	6.08	6.99
Variance	6.20	1.43
Observations	73.00	73.00
Pearson Correlation	0.34	
Hypothesized Mean Difference	0.00	
df	72.00	
t Stat	-3.27	
P(T<=t) one-tail	0.00	
t Critical one-tail	1.67	
P(T<=t) two-tail	0.00	
t Critical two-tail	1.99	
Z	-2.80	
P(Z<=z) one-tail	0.00	
z Critical one-tail	1.64	
P(Z<=z) two-tail	0.01	
z Critical two-tail	1.96	

We can observe that p-value is 0.00. That means the obtained results on the practical assignment and the quiz (final exam) are statistically significant. So, even if the variance is different for these two sets, the values are correlated and significant.

5. Discussion and conclusions

A variety of evaluation methods can be used in teaching and those should determine the method of feedback provided to students. Several interesting outcomes were highlighted by our study. First, the analysis between practical assessments and quizzes, in the context of educational evaluation methods, suggested that quizzes tend to yield narrower score intervals and lower variances compared to practical assessments. Quizzes are typically more standardized and provide a consistent measure of students' knowledge and understanding. Reduced variability in quiz scores may be attributed to the controlled nature of quiz formats, where students are required to select predefined answers or make specific decisions based on given scenarios. In contrast, practical assessments involve subjective evaluations of students' performance, which can introduce greater variability in grading criteria and outcomes as well as greater variability in student approaches to solving a predefined problem. In quizzes, students generally have a fixed amount of time for completion while practical assignments are more open-ended and can vary based on student time commitments.

The analysis also revealed a significant correlation between students' performance on practical assignments and their quiz scores. This finding suggests that students who excel in practical assignments are likely to perform well in quizzes as well. One possible explanation is that practical assignments provide students with hands-on experience, allowing them to apply theoretical concepts in real-world scenarios. This practical application of knowledge may enhance their understanding and retention of the material, leading to improved quiz performance. Additionally, the correlation implies that teachers who emphasize practical assignments as part of the evaluation process may effectively assess and reinforce students' comprehension of the subject matter.

It is important to consider the strengths and limitations of each evaluation method. Practical assessments, such as projects and labs, provide opportunities for students to demonstrate their skills and problem-solving abilities in a realistic context. These assessments align with the demands of real-world situations and encourage critical thinking and creativity. On the other hand, quizzes, particularly when well-designed, can effectively assess students' theoretical knowledge and conceptual understanding. Quizzes also offer advantages in terms of ease of administration, automated grading processes, and timely feedback provision.

Teachers should consider the balance between practical assessments and quizzes based on learning objectives, subject matter, and desired learning outcomes. A combination of both evaluation methods can provide a comprehensive assessment of students' competencies, covering theoretical knowledge, practical application, and problem-solving abilities. By incorporating a variety of assessment approaches, teachers can gain a holistic understanding of students' strengths and areas for improvement, enabling personalized feedback and targeted instructional interventions.

Furthermore, the findings of this study contribute to the ongoing discussions on evaluation practices in education. Teachers must carefully select and design evaluation methods that align with the specific learning goals and the type of content being covered. Additionally, the study highlights the importance of providing personalized feedback to students. Feedback serves as a valuable tool for promoting student engagement, self-reflection, and continuous improvement.

The study had several limitations. For example, the analysis focused on a specific sample of students in a single university’s Java programming course for mobile devices. The results may not be directly generalizable to other academic disciplines or locations. The study examined 4 evaluation methods, and other approaches, such as peer assessment, portfolio approaches, group assessments and oral examinations, were not included. Future research should explore additional evaluation methods and investigate the impact on students' learning outcomes in diverse educational contexts.

Acknowledgments

This work was supported by a grant of the Ministry of Research, Innovation and Digitalization, CCCDI - UEFISCDI, project number PN-III-P2-2.1-PED-2021-2142, within PNCDI III and the Core Program within the National Research Development and Innovation Plan 2022-2027, carried out with the support of MCID, project no. 23 38 01 01, “Contributions to the consolidation of emerging technologies specific to the Internet of Things and complex systems".

REFERENCES

Baigi, S. F. M., Aval, R. N., Sarbaz, M. & Kimiafar, K. (2022) Evaluation tools for digital educational games: A systematic review. *Acta Medica Iranica*. 60(8). 508-512.

Bunse, C., Kennes, L. & Kuhr, J. C. (2022, May) On using distance labs for engineering education. In *2022 IEEE/ACM 4th International Workshop on Software Engineering Education for the Next Generation (SEENG)*. IEEE. pp. 5-11.

Freitas, S. A. A., Silva, R. D. C., Lucena, T. F. R., Ribeiro, E. D. N., De Lima, V. C. & Da Silva, R. M. (2016, January) Smart quizzes in the engineering education. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*. IEEE. pp. 66-73.

McConnell, D. (2002) The experience of collaborative assessment in e-learning. *Studies in continuing education*. 24(1), 73-92.

Mohan, R. (2023) *Measurement, evaluation and assessment in education*. PHI Learning Pvt. Ltd.

Nodira, T. & Rashid, X. (2022) Problems Of Innovation Management In The Higher Education System. *Web of Scientist: International Scientific Research Journal*. 3(11), 155-164.

Paiva, J. C., Leal, J. P. & Figueira, Á. (2022) Automated assessment in computer science education: A state-of-the-art review. *ACM Transactions on Computing Education (TOCE)*. 22(3), 1-40.

Sanjaya, D. B., Suartama, I. K. & Suastika, I. N. (2022) The Effect of the Conflict Resolution Learning Model and Portfolio Assessment on the Students' Learning Outcomes of Civic Education. *International Journal of Instruction*. 15(1), 473-488.

Steinherr, V. M. & Vay, C. (2023) Leadership Education in a Technology-Enhanced Learning Environment: The Relation Between Self-Regulated Learning and Self-Leadership. *ECIS 2023 Research Papers*. 370. https://aisel.aisnet.org/ecis2023_rp/370.

Susanna, V. (2022) Information and Communication Technologies in Education. *Eurasian Journal of Learning and Academic Teaching*. 6, 89-93.