LOW-STAKES ASSESSMENT WITHOUT EFFECTIVE FEEDBACK IS WORTHLESS: A DESCRIPTIVE-ANALYTICAL STUDY OF PROGRAMMATIC ASSESSMENT IN UNDERGRADUATE DENTAL EDUCATION

(PENILAIAN RISIKO RENDAH TANPA UMPAN BALIK YANG EFEKTIF TIDAK BERHARGA: STUDI DESKRIPTIF-ANALITIK TENTANG PROGRAMMATIC ASSESSMENT PADA PENDIDIKAN SARJANA KEDOKTERAN GIGI)

JHDS.unjani.ac.id/jite
Doi:10.xxxx/jite

Article History
Received:11/06/2025
Accepted: 15/06/2025

Kertamaya Sundawan[1*]
[1]Dental Education Unit, Faculty of Dentistry, Universitas Jenderal achmad Yani, Cimahi, Indonesia

*Corresponding author

kertamaya.s@fkg.unjani.ac.id

ABSTRACT

Assessment plays a vital role in the education process. Assessment plays a significant role in the learning process in medical education, serving not only to evaluate learning outcomes but also to declare someone competent. Formative assessment is typically low-stakes, designed to stimulate learning. This low-risk assessment is effective when it is integrated into the ongoing teaching and learning program to facilitate timely, specific, and actionable feedback for students. The aim is to evaluate the effectiveness and sustainability factors of programmatic assessment based on the results of data analysis. The research method used is a descriptive analytical study. This study employed a descriptive-analytical design using a survey of undergraduate student grade data in the Periodontology study program, Faculty of Dentistry, Jenderal Ahmad Yani University. Inclusion criteria were students with longitudinal low-stakes grades, while those not participating in the low-stakes learning process were excluded. Data were analyzed using statistical methods to calculate mean and standard deviation, with institutional approval ensuring confidentiality. The results show that the average low-risk score is 83.94, while the high-risk score has an average of 68.47, with a passing percentage of 100% on low-risk scores, resulting in a very good predicate, and only 7% on high-risk assessments. It suggests that there is no holistic approach to

programmatic assessment, which is influenced by several factors, including students not receiving timely feedback, poorly designed final exam questions, and internal factors such as nervousness during the final exam. The conclusion is that programmatic assessment should combine low-stakes and high-stakes assessments. Effective low-stakes assessments can enhance students' preparation for high-stakes exams. Low-stakes assessments can also increase students' motivation to study more deeply and regularly, but this depends on the quality of the feedback provided.

Keywords: assessment; feedback; programmatic assessment

ABSTRAK

Penilaian berperan penting dalam proses pendidikan. Penilaian berperan sangat besar dalam proses pembelajaran dalam pendidikan kedokteran, selain digunakan untuk mengevaluasi capaian pembelajaran, penilaian sering digunakan untuk menyatakan seseorang kompeten. Penilaian formatif biasanya berisiko rendah yang dimaksudkan untuk merangsang pembelajaran. Penilaian berisiko rendah ini efektif jika terintegrasi dalam program belajar mengajar yang sedang berlangsung untuk memfasilitasi umpan balik yang tepat waktu, spesifik, dan dapat ditindaklanjuti bagi mahasiswa. Tujuannya untuk mengevaluasi faktor efektivitas dan keberlanjutan penilaian program dari hasil analisis data. Metode penelitian yang digunakan adalah studi deskriptif analitik dengan survei data nilai mahasiswa program studi Periodonsia, Fakultas Kedokteran Gigi, Universitas Jenderal Ahmad Yani. Kriteria inklusi adalah mahasiswa yang memiliki nilai longitudinal low-stakes, sedangkan yang tidak mengikuti proses pembelajaran low-stakes dikeluarkan. Data dianalisis menggunakan metode statistik dengan perhitungan rata-rata dan standar deviasi, serta mendapat persetujuan institusi dengan menjaga kerahasiaan data.. Hasil penelitian menunjukkan bahwa skor penilaian formatif rata-rata adalah 83.94 sedangkan skor penilaian sumatif memiliki rata-rata 68.47 dengan persentase kelulusan pada penilaian formatif sebesar 100% mendapatkan predikat sangat baik dan hanya 7% pada penilaian sumatif. Hal ini menunjukkan bahwa belum ada penilaian program yang holistik, hal ini terkait dengan beberapa faktor, antara lain mahasiswa tidak diberikan umpan balik, Evalusai soal ujian akhir yang belum dilaksanakan dengan baik, dan Faktor internal seperti rasa gugup mahasiswa saat ujian akhir. Kesimpulan ini

adalah penilaian programatik harus menggabungkan penilaian berisiko rendah dan berisiko tinggi. Penilaian berisiko rendah yang efektif dapat meningkatkan persiapan mahasiswa menghadapi ujian berisiko tinggi. Penilaian berisiko rendah juga dapat meningkatkan motivasi mahasiswa untuk belajar lebih mendalam dan teratur, tetapi hal ini bergantung pada kualitas umpan balik yang diberikan.
Kata Kunci: p*enilaian; programmatic assessment; umpan balik*

INTRODUCTION

The field of medical education has long recognized the importance of assessment in the learning process. Assessment serves as a critical tool for evaluating student performance, guiding curriculum development, and ensuring the competence of future healthcare professionals. Programmatic assessment represents a holistic approach to evaluating student learning and competence in medical education. This method involves collecting longitudinal data at various points in time, providing a comprehensive view of student performance over an extended period.[1,2] This method aims to optimize the learning function, decision-making function, and curriculum quality assurance function of assessment.[3]

Programmatic assessment in dental education combines low-stakes and high-stakes assessments to provide a comprehensive evaluation of student progress and competence. While low-stakes assessments offer ongoing feedback and opportunities for improvement, high-stakes assessments measure overall achievement at the end of a learning period.[4] The use of multiple assessment modalities, as seen in the programmatic assessment model for medical education, can be adapted to dental education to ensure comprehensive evaluation across multiple competency domains. However, this will be a major challenge in dental education.[5]

Programmatic assessment comprises several core components that are integrated to achieve comprehensive and continuous evaluation within the education program. One of the components is the formative and summative assessment system, which is currently undergoing a paradigm shift from the formative and summative approaches to the continuum of stakes approach.[6] Low-stakes assessment is formative, while high-stakes assessment is a summative assessment. Low-stakes assessments can build student confidence by providing opportunities to practice and develop skills, while high-stakes assessments are those that have major or significant consequences for the student's academic future. The results of this assessment usually have a direct impact on important decisions such as graduation, so it has the greatest weight in determining the final grade [7,8]

Generally, countries in Asia, including Indonesia, have a high Power Distance Index (PDI).

A high PDI culture indicates a strong hierarchy, reflected in significant power differences between individuals, such as between lecturers and students. The socio-cultural aspect of a strong hierarchy, where health professionals are considered to hold a higher position in the social structure, can affect various aspects of the interaction gap between patients and health workers (9,10). This hierarchy has a significant impact on the learning process.

In Indonesia, the socio-cultural context plays a significant role in the education system, including medical education. Although the use of flow-stakes assessment in dental education has significant benefits, it can be faced challenges originating from the Indonesian culture and social system, which have certain characteristics. Therefore, the research aims to evaluate the effectiveness and sustainability factors of programmatic assessment based on the results of data analysis.

METHOD

The research design for this study employed a descriptive-analytical method. This approach allows for systematic examination and interpretation of the collected data. This approach enables a comprehensive examination of the data studied, allowing it not only to describe the current state but also to analyze the underlying factors and potential implications.[11]

The research method used in this study was a survey of student grade data. The survey method selected for this study aligns with the analytical descriptive approach, facilitating the collection of quantitative data. The data used is the value data of undergraduate students registered in the Periodontology study program at the Faculty of Dentistry, Jenderal Ahmad Yani University. The inclusion criteria in this writing are students who have longitudinal low-stakes grades, while the exclusion criteria are students who do not follow the low-stakes learning process. The collected data were analyzed using statistical analysis, involving the calculation of the average and standard deviation for each item in the data. The Institution approved this study, and the data values in this writing will be kept confidential.

RESULT

This study utilizes grade data on the Periodontal subject to analyze differences in student performance on two types of assessments: low-stakes and high-stakes assessments. In this study, students' daily grades were used as a representation of the low-stakes assessment, which is low-stakes and does not directly influence the final results, but is vital for assessing the development of skills and understanding during the learning process. Meanwhile, the value of the multiple-choice question (MCQ), which has the highest weight, is used as a high-stakes assessment, where the results of this exam have a significant impact on students' final achievement. Student grades can be seen in Table 1.

Table 1. Low and high stakes results in periodontal course

Information	Low Stakes	High Stakes
AVERAGE	83.94	68.47

MAX	85.00	84.84
MIN	80.00	36.89
MEDIAN	84.50	70.09
SD	1.28	9.36
SUM	12.339	10.065

Based on the data presented, a comparison of the results between the low-stakes assessment and the high-stakes assessment in the periodontal subject is provided. The total score for low-stakes is 12.339, higher than the total score for high-stakes, which is 10.065. It indicates that, overall, the score for the low-stakes assessment is higher than the score for the high-stakes assessment. The average score for low-stakes is 83.94, while the average score for high-stakes is lower, at 68.47. It suggests that, in general, student performance in low-stakes assessments is better than in high-stakes assessments, as the greater pressure may influence the high-stakes exam.

The maximum score on the low-stakes assessment reached 85.00, slightly higher than the maximum score on the high-stakes evaluation of 84.84. Although the difference is slight, this indicates that the highest scores achieved on both types of assessments are nearly identical. A significant difference is observed in the minimum score, where the minimum score on the low-stakes is 80.00, whereas on the high-stakes, the minimum score is much lower, specifically 36.89. It suggests that some students experience greater difficulty with the high-stakes exam, which may be due to the higher level of stress or the exam's complexity.

The median score for the low-stakes was 84.50, while for the high-stakes it was 70.09. It indicates that half of the students who took the low-stakes assessment scored above 84.50, whereas for the high-stakes assessment, the lower median indicates a more narrow distribution of scores. The standard deviation for the low-stakes was 1.28, indicating minimal variation or spread around the mean, which reflects consistency in student scores. In contrast, the standard deviation for the high-stakes was 9.36, indicating much greater variation. It suggests that there was greater volatility in the high-stakes results, with some students scoring exceptionally high and others scoring exceptionally low.

DISCUSSION

From these results, students tend to get better and more consistent grades in low-stakes assessments, while high-stakes assessments show more varied results and lower average grades. Basically, low-stakes assessments are specifically designed to help the learning process and support student development. Low-stakes assessments serve to provide direct feedback to students and lecturers, to improve and adjust teaching throughout the learning process, not just at the end. [12]

Learning in dentistry involves programmatic assessment, which is a thorough and ongoing evaluation method. This approach aims to gather diverse data regarding student performance and competence through multiple methods. There are two primary categories of assessment in this approach: low-stakes and high-stakes assessments. Low stakes assessments conducted at the Faculty of Dentistry UNJANI include group assignments,

problem-based learning, and quizzes using the edpuzzle platform, while high stakes assessments conducted include multiple choice question (MCQ) and Oral Assessment (OA) exams.

Figure 1. Problem-based Learning in Small Class.

Figure 2. Multiple Choice Question (MCQ) Examination.

Group work and group discussions are methods used as low-stakes assessment evaluations that allow for peer interaction and collaborative learning. This approach is consistent with findings that emphasize the importance of ongoing assessment in the classroom. Group activities enable lecturers to observe and assess students' understanding, communication skills, and ability to work collaboratively.[13] The inability of lecturers to provide feedback on group learning and PBL can have a significant impact on students' learning process. In addition, the lack of constructive feedback can hinder the development of students' critical thinking skills and analytical abilities. It can ultimately harm the overall quality of education and reduce the effectiveness of the learning process in group learning and PBL systems. Constructive feedback from lecturers is essential in tutorial learning. Research shows that the majority of students (55.9%) consider the feedback given by lecturers to be still at a moderate level, while the rest receive more detailed feedback in the form of correct answers and explanations.[5] Differences in perceptions and practices of feedback between lecturers can create inconsistencies in students' learning experiences. It can lead to confusion and frustration among students, especially if they receive different or even conflicting feedback from different lecturers. On the other hand, other studies have found that low-stakes feedback and motivational approaches are essential elements in the learning context. [14]

The use of EdPuzzle for quizzes is an innovative integration of technology into low-stakes assessment. As highlighted in Section 1.4. EdPuzzle quizzes can be retaken multiple times. EdPuzzle's flexibility extends beyond its core functionality, allowing educators to create interactive video lessons that cater to a variety of learning styles. The platform's ability to support repeated quiz attempts

encourages a growth mindset, prompting students to engage more deeply with the content. The lack of feedback on some EdPuzzle quizzes can limit the platform's full potential to enhance student learning. While video content can be low-stakes, the lack of immediate feedback on quiz performance can hinder students' ability to gauge their understanding and identify areas for improvement.

Low-stakes assessments can increase students' motivation to study more deeply and regularly, but this depends on the quality of the feedback provided. Studies have found that an excessive frequency of low-stakes assessments without adequate feedback can actually hinder students' independent learning and negatively impact their performance on high-stakes exams.[15] Without constructive feedback, low-stakes assessments become less effective in motivating students. Research indicates that learning-oriented students are more likely to value low-stakes assessments and utilize them for self-study if they receive constructive feedback.[16] Therefore, the quality of feedback is more important than the quantity of low-stakes assessments in motivating students. Feedback should be specific, actionable, and help students identify their weaknesses and areas for development.[17] Thus, constructive feedback is key to maximizing the motivational benefits of low-stakes assessment in dental education.

Students' preparation for high-stakes exams can be enhanced through the use of effective low-stakes assessments. Studies show a positive correlation between performance on low-stakes assessments and final exam results (high-stakes).[18] The development of a comprehensive and integrated low-stakes assessment system can help students identify areas for improvement before high-stakes exams. Additionally, providing constructive and specific feedback on low-stakes assessments can motivate students to make improvements and deepen their understanding. The implementation of active and collaborative learning strategies in high-stakes exam preparation can also help students consolidate their knowledge and improve their exam readiness. It supports the assumption that students' performance on low-stakes assessments can predict their performance on high-stakes exams. Therefore, good low-stakes assessments can help students prepare more optimally for high-stakes exams.[15,18] However, internal factors such as anxiety and academic self-efficacy also play an important role. A significant negative correlation was found between academic self-efficacy and students' anxiety in facing final semester exams.[19] Students with higher academic self-efficacy tend to have lower levels of anxiety. In addition, negative emotions such as anger, fear, and sadness experienced after failing an exam can hinder motivation to study and prepare for the next exam.[20] This can be minimized by providing feedback. Providing constructive feedback can help students understand their mistakes and improve their academic self-efficacy. Emotional management strategies and relaxation techniques can also be taught to students to manage anxiety and negative

emotions that arise during the exam period. Additionally, fostering a supportive learning environment and promoting open communication between lecturers and students can help alleviate stress and enhance exam readiness. Academic self-efficacy can also affect students' learning strategies and persistence in facing academic challenges. Increasing academic self-efficacy through psychological interventions or self-development programs can be an effective approach to reducing exam anxiety. Additionally, developing emotional regulation skills can help students manage negative emotions following exam failure and maintain their motivation to study.

Another factor that causes failure in high-stakes exams, namely, flawed or poorly designed MCQs, can harm student achievement, especially for high and medium achievers.[21] The quality of distractors in MCQs plays a significant role in differentiating between low and high achievers. A study found that over 45% of distractors in MCQ tests were relatively to very weak, with some even having a negative impact by drawing out more high achievers than low achievers.[22] This emphasizes the need for continuous evaluation and improvement of MCQ questions to ensure their effectiveness in measuring student understanding. It is essential for educators to carefully consider the selection and formulation of distractors to accurately distinguish students' levels of understanding accurately.

Interestingly, the flawed MCQ test showed better reliability compared to the standard test, a difference that was statistically significant ($p < 0.05$).[21] This suggests that the quality of MCQ design has a complex impact on student assessment outcomes. Ineffective distractions can obscure actual differences in ability between students. Additionally, involving students in the feedback process after the MCQ test can help identify areas for improvement in item design and enhance overall learning quality. Feedback is designed to be two-way, which is why it is one-way due to the socio-cultural impact of hierarchy in the learning environment. To address the socio-cultural impact of hierarchy in learning, educators can create a more inclusive environment and encourage active participation of all students in the feedback process. Strategies such as small group discussions, anonymous Q&A sessions, or the use of interactive technology can help minimize communication barriers and encourage greater student engagement. In addition, educators can practice facilitation skills to ensure that all voices are heard and valued in the process of evaluating and improving MCQ questions [12,22]

Constructivism has emerged as a dominant paradigm in Indonesian education, focusing on students constructing their own knowledge rather than simply receiving knowledge transfer from lecturers. This approach emphasizes students' active involvement in authentic and meaningful learning, both individually and in a learning community. However, there are still challenges in implementing a more interactive approach. Research indicates that Indonesian students tend to exhibit more passive and idealistic learning patterns compared to students

from other countries, such as those from Sri Lanka and the Netherlands.[23] This can be attributed to the high hierarchical culture in Indonesian education. To overcome this challenge, a systematic effort is needed to change the mindset and learning practices in Indonesia. This can include training lecturers in interactive teaching methods, adjusting the curriculum to encourage active student participation, and creating a more collaborative learning environment.

Continuous training also encourages lecturers to continuously improve themselves to meet the "knowledge demands" of students.[24] Interestingly, some studies have found a gap between teaching practices and assessments conducted by lecturers. For example, in Iran, the majority of instructors still adhere to traditional methods and show no inclination to adopt innovative practices.[25] This emphasizes the importance of ongoing feedback training to encourage innovation in teaching. In addition, ongoing feedback training can help lecturers identify areas for improvement in their teaching methods. This process also encourages self-reflection and critical evaluation of teaching practices, which, in turn, can enhance the effectiveness of classroom learning. Furthermore, ongoing training enables lecturers to stay current with the latest developments in their field of education and discipline, thereby providing more relevant and meaningful learning experiences for students. It highlights the absence of a holistic assessment in programmatic evaluation, as the current approach tends to be limited to assessing only certain components, thereby failing to provide a complete and accurate picture of student competency. Through a holistic assessment, essential dimensions such as cognitive abilities, practical skills, emotional aspects, and professional attitudes can be addressed, leading to a better understanding of student development and learning needs. It emphasizes the need for improvements in the design and implementation of programmatic assessment by starting with feedback training to change the view of pedagogical learning to an andragogic.

CONCLUSION

Programmatic assessment should combine low-stakes and high-stakes assessments. Effective low-stakes assessments can enhance students' preparation for high-stakes exams. Low-stakes assessments can also increase students' motivation to study more deeply and regularly, but this depends on the quality of the feedback provided. Without constructive feedback, low-stakes assessments become less effective in motivating students. Lecturers need education on constructive feedback and how to deliver it effectively. Based on the three conclusions above, the recommendation for medical education institutions is to provide curriculum development training for lecturers, with a focus on regular feedback.

CONFLICT OF INTEREST

The authors reported no potential conflict of interest.

ACKNOWLEDGEMENT

Thank you to the professional who contributed to this manuscript.

REFERENCES

1. Shrivastava SRBL, Shrivastava PS. Capacity building of teachers to conduct fair online assessments. Indian J Anaesth. 2024 Jan 1;68(1):121–2.
2. Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, Govaerts MJB, Heeneman S. Twelve tips for programmatic assessment. Med Teach. 2015;37(7):641–6.
3. Maas MJM, Sluijsmans DMA, Van Der Wees PJ, Heerkens YF, Nijhuis-Van Der Sanden MWG, Van Der Vleuten CPM. Why peer assessment helps to improve clinical performance in undergraduate physical therapy education: A mixed methods design. BMC Med Educ. 2014 Jun 13;14(1).
4. Torre D, Rice NE, Ryan A, Bok H, Dawson LJ, Bierer B, et al. Ottawa 2020 consensus statements for programmatic assessment–2. Implementation and practice. Med Teach [Internet]. 2021;43(10):1149–60. Available from: https://doi.org/10.1080/0142159X.2021.1956681
5. Bissell V, Dawson LJ. Assessment and feedback in dental education: a journey. Br Dent J. 2022 Sep 23;233(6):499–502.
6. Greviana N, Ainin DQ, Sari SM, Claramita M. I Am Hpe Recommendation Programmatic Assessment: Bagaimana Penerapannya? Jurnal Pendidikan Kedokteran Indonesia-The Indonesian Journal of Medical Education. 2022;11(4):457–66.
7. Harrison CJ, Könings KD, Dannefer EF, Schuwirth LWT, Wass V, van der Vleuten CPM. Factors influencing students' receptivity to formative feedback emerging from different assessment cultures. Perspect Med Educ. 2016 Oct 1;5(5):276–84.
8. Ganjarani GH, Herlina S, Firmansyah M. High Stakes Assessment Dan Dampak Terhadap Pembelajaran Di Fakultas Kedokteran Dan Kesehatan: Systematic Literature Review. 2022.
9. Susilo AP, Dalen J Van, Scherpbier A, Tanto S, Yuhanti P, Ekawati N. Nurses ' roles in informed consent in a hierarchical and communal context. 2013;20(4):413–25.
10. Claramita M. Doctor-patient communication in a culturally hierarchical context of Southeast Asia : a partnership approach Doctor-patient communication in a culturally hierarchical context. 2012;
11. Wolniak. The concept of descriptive analytics. Scientific Papers of Silesian University of Technology Organization and Management Series. 2023;2023(172).
12. Ismail SM, Rahul DR, Patra I, Rezvani E. Formative vs. summative assessment: impacts on academic motivation, attitude toward

learning, test anxiety, and self-regulation skill. Language Testing in Asia. 2022 Dec 1;12(1).

13. Kumar Upadhyay A, S SN, Kumar D. Development Of Grade Level Assessment Tool In Social Science (Glat-Ss) For Grade Vii & Viii Based On Karnataka State Education Board For Typically Developing Children And Children With Hearing Impairment. 2024.
14. Kusuma JW, Hamidah H. Online Learning: Student's Perception off Lecturer's Competence and its Influence on Achievement Motivation. International Journal for Educational and Vocational Studies. 2021 Mar 30;3(1):37.
15. Hossain S, Begum N, Talukder MHK. Teachers' evaluation of formative assessment on summative assessment in undergraduate medical education. Bangladesh Journal of Medical Education. 2014 Apr 12;3(1):18–21.
16. van Wijk E V., van Blankenstein FM, Donkers J, Janse RJ, Bustraan J, Adelmeijer LGM, et al. Does 'summative' count? The influence of the awarding of study credits on feedback use and test-taking motivation in medical progress testing. Advances in Health Sciences Education. 2024;
17. Elbra-ramsay C. Roles , relationships and emotions : Student teachers ' understanding of feedback as interpersonal. 2021;
18. Chand SP, Pillay KK, Pillay K. Understanding the fundamental differences between formative and summative assessment Global Scientific and Academic Research Journal of Education and Literature. Understanding the fundamental differences between formative and summative assessment. 2024; Available from: https://gsarpublishers.com/gsarjel-home-page/
19. Schunk DH, DiBenedetto MK. Self-efficacy and human motivation. In: Advances in Motivation Science. Elsevier Ltd; 2021. p. 153–79.
20. Omotoy JF. Examining College Students' Self-Efficacy In The Online Learning Environment System During The Covid-19 Pandemic Implications For Higher Education Institutions. Revista de Gestao Social e Ambiental. 2023 Jul 6;17(5).
21. Jabeen H, Fatima Z, Awan RUN. Impact of Academic Stress on Motivation and Academic Achievement of Undergraduates. UMT Education Review [Internet]. 2024 Jun 28;7(1):78–90. Available from: https://journals.umt.edu.pk/index.php/uer/article/view/4017
22. Al Ameer AY. Assessment of the Quality of Multiple-Choice Questions in the Surgery Course for an Integrated Curriculum, University of Bisha College of Medicine, Saudi Arabia. Cureus. 2023 Dec 13;
23. Suhendi A, Purwarno P, Chairani S. Constructivism-Based Teaching and Learning in Indonesian Education. KnE Social Sciences. 2021 Mar 11;76–89.

24. Sherly, Gultom S, Daryanto E, Nasrun. Management of Continuous Professional Development through Competency-Based Training Model for Junior High School Teachers. Emerging Science Journal. 2023 Feb 1;7(1):190–206.
25. Taghizadeh M, Mazdayasna G. Language assessment course at Iranian state universities: An evaluation of the incorporation of assessment principles into the course content. Heliyon. 2023 Jan 1;9(1).