ORIGINAL ARTICLE

Gamifying Assessments in Medical Education: Impact of Gamification on Student's Test Scores

Madiha Ata[1*], Shazia Irum[2], Fahad Azam[3], Abida Shaheen[4], Shahzad Riyaz[5] and Muhammad Waqas Rabbani[6]

[1]*Department of Health Professions Education, Indus University of Health Sciences, Indus Hospital & Health Network, Karachi, Pakistan*
[2]*Department of Health Professions Education, Shifa College of Medicine, Shifa Tameer-e-Millat University, Islamabad, Pakistan*
[3]*Department of Pharmacology and Therapeutics, Shifa College of Dentistry, Shifa Tameer-e-Millat University, Islamabad, Pakistan*
[4]*Department of Pharmacology & Therapeutics, Shifa College of Medicine, Shifa Tameer-e-Millat University, Islamabad, Pakistan*
[5]*Department of Gastroenterology, Shifa International Hospital, Islamabad, Pakistan*
[6]*Department of Behavioral Sciences, Shifa College of Medicine, Shifa Tameer-e-Millat University, Islamabad, Pakistan*

Abstract

Background: Gamification in academic settings is a specific type of Technology Enhanced Assessment (TEA) designed to engage learners in an interactive learning environment and increase motivation.

Objective: To explore the role of gamification in improving the academic performance of medical students in Pakistan.

Methods: A mixed-methods approach was employed for the present study, which was conducted from March to May 2022. 251 medical students attempted a formative assessment through online platforms with and without gamification. Students were interviewed for the collection of qualitative data.

Results: Among the study participants, 54.2% were males and 45.8% were females, with a mean age of 19.26±1.14 years and 20.16±0.96 years in the first and second-year MBBS students respectively. The findings revealed a significant gender difference in test scores between students who used gamified and non-gamified assessment tools. However, the comparison of mean formative assessment scores with and without gamification did not yield statistically significant results. Overall, the students had a good perception of the introduction of gamification in assessment and expressed a desire to experience it more frequently.

Conclusion: Gamified assessments can improve academic outcomes by promoting engagement and can mitigate gender disparities in assessment scores. Further research is warranted to explore the specific mechanisms by which gamification influences learning outcomes and to investigate its effectiveness across diverse student populations.

Keywords: *Gamification, formative assessment, engagement, motivation, gender differences, test anxiety.*

INTRODUCTION

Recent assessment trends focus on using existing tools and technology to create more effective assessments, leading to the development of Technology-Enhanced Assessment (TEA). Purposeful gamification is a specialized form of TEA that is being utilized in education to engage learners, especially millennials, to increase their motivation and enhance their learning [1]. Various studies have found that good game design elements can foster feelings of relatedness, competence, and autonomy in a learner [2]. A qualitative exploration of students' perspectives regarding gamification has revealed certain common themes in the literature where students have found it to be 'interesting', 'activating', 'helpful', and 'fun' [3].

However, it is essential to note that the effect of gamification on test scores may depend on various factors such as the design of the gamified assessments, student preferences, and the specific learning objectives [4].

While gamification can be a valuable tool to enhance learning experiences, its direct impact on test scores may vary across different educational settings and student populations.

With this background, the objective of the present study was to explore the role of gamification in improving the academic performance of medical students in Pakistan.

METHODOLOGY

The impact of gamification on medical students' formative assessment test scores was investigated using a mixed-method sequential explanatory design. The study was conducted at Shifa Tameer-e-Millat University and Bahria University Medical and Dental College from March to May 2022; after obtaining approval from the two institutional review boards (IRB # 048-22, ERC 51/2022).

To determine the appropriate sample size for this study, we utilized a prevalence rate of test anxiety among students, reported in the literature to range between 25% and 40% [5]. For this calculation, a prevalence rate of 30% (CI: 99%, Error Margin: 5%) was selected with the help of Open Epi. Software. This mid-range value provided a balanced estimate, ensuring that our sample

*****Corresponding author: Madiha Ata, Indus University of Health Sciences, Karachi, Pakistan, Email: madihaata@hotmail.com*
Received: July 26, 2024; Revised: October 04, 2024; November 12, 2024
DOI: https://doi.org/10.37184/lnjpc.2707-3521.7.30

  ISSN: 2708-9134 (Online) Liaquat National Journal of Primary Care 2025; 7(2): 130-134

size adequately reflects the anticipated variability in the population. The initial sample size calculation yielded 276 participants.

We adjusted the sample size and recruited 300 participants to account for potential non-responses or incomplete data. This adjustment ensured that the final sample size remains robust and representative, even if there is some attrition or data loss. The final sample consisted of 300 male and female students belonging to Year 1 and Year 2 of MBBS, coming from diverse backgrounds. They were selected based on their consent and willingness to participate in the study. All students participating in the study were comfortable with the use of mobile phones and laptops. The participants were assigned to one of four groups through the lottery method by using their enrollment numbers. Two groups, one each from Year 1 and Year 2 completed a formative assessment using gamified application Quizziz, while the other two groups, one each from Year 1 and Year 2 completed the same formative assessment using the quiz feature of the non-gamified app, Google Forms. The assessment was carried out in the classroom with 3 instructors proctoring the students physically. The students were divided into groups and took a 20-question multiple-choice quiz using Quizziz and Google Forms. Quizziz features like leaderboards, feedback, bonuses, music, and memes were enabled. The formative assessment scores were based on the number of correct answers on the first attempt, without considering any bonuses earned through the gamification features.

In the qualitative phase, eight students were interviewed to explore their perceptions of gamification. The students were selected through purposive sampling to include both high achievers and low achievers, to be representative of participants who took the assessment with and without gamification. The interviews were semi-structured, using a guide developed on the theoretical framework provided by 'Self Determination Theory' [6]. The students were approached through e-mail and invited to participate in interviews voluntarily. Face-to-face interviews were recorded with a smartphone and transcribed using Otter. The interviewer also took notes of the contextual elements such as body language, environment, *etc.* during the interviewees for thick description and reflexivity. After six interviews it was observed that new themes could not be identified. So, the interviews were stopped at a total of 8 students due to saturation of data.

The quantitative data was analyzed using SPSS (Statistical Package for the Social Sciences) version 26.0. The Shapiro-Wilk test was used to determine the normality of quantitative data, and a p-value of 0.05 was chosen as the threshold for statistical significance. The Mean ± Standard Deviation was used to represent the continuous variables. The student's t-test was used to compare numerical variables, and the Chi-square or Fisher's exact test, whichever was appropriate, was used to investigate associations.

MaxQDA 2020 was used to assess, code, and analyze the themes of the qualitative data. Each line of the transcript was coded using a mixture of line-by-line and in-vivo coding to explore the data in detail. Comparison of data sets of respondents was done continuously to find common themes. From these initial codes, similar ones were categorized to identify themes and create sub-themes. After identifying initial themes, the transcripts were revisited to reorganize the codes, eliminate redundant themes, and combine similar ones. These themes and sub-themes were finalized using the SDT as the theoretical lens. The themes were categorized under the broad headings of autonomy, competence, and relatedness.

RESULTS

Of the 251 participants in the study, 140 (55.8%) belonged to first-year and 111 (44.2%) to second-year MBBS. The group comprised of 115 (45.8%) females and 136 (54.2%) males. The first- and second-year participants had mean ages of 19.26 ± 1.14 and 20.16 ± 0.96, respectively. Most of the participants [214 (85.3%)] were hostel dwellers, with day scholars making up a lesser percentage [37 (14.7%)]. Out of all the participants, 130 (51.80%) attempted the formative assessment with gamification, and 121 (48.20%) without gamification.

The mean formative score of Year 1 students was 80.69 ± 15.97 and that of Year 2 students was 73.02 ± 17.51. The comparison revealed a statistically significant difference *i.e.* p<0.001. The comparison of mean formative score with gamification (Quizziz) for Year 1 students (82.46 ± 15.43); and Year 2 students (74.28 ± 20.78) also depicted a significant difference *i.e.* p=0.011. A similar comparison of mean formative scores without gamification of Year 1 (n=71) students and Year 2 students was 78.96 ± 16.40 and 71.48 ± 12.45 respectively. The comparison yet again revealed a significant statistical difference *i.e.* p=0.005 (**Table 1**).

The mean formative score of Year 1 students with gamification (82.46 ± 15.44) and without gamification (78.96 ± 16.40) did not reveal a statistically significant

Table 1: Comparison of formative scores of year 1 and year 2.

Variables	Mean ± SD	p-value
Mean formative score Year 1 (n=140)	80.69 ± 15.97	<0.001*
Mean formative score Year 2 (n=111)	73.02 ± 17.51	
Mean formative score with gamification Year 1 (n=69)	82.46 ± 15.43	0.011*
Mean formative score with gamification Year 2 (n=61)	74.28 ± 20.78	
Mean formative score without gamification Year 1 (n=71)	78.96 ± 16.40	0.005*
Mean formative score without gamification Year 2 (n=50)	71.48 ± 12.45	

*= statistically significant

Table 2: Comparison of scores of year 2 formative assessments with gamification and formative assessments without gamification.

Variables	Mean ± SD	p-value
Mean formative score with gamification (n=61)	74.28 ± 20.78	0.383
Mean formative score without gamification (n=50)	71.48 ± 12.45	
Mean formative score of male students with gamification (n=32)	75.66 ± 19.85	0.824
Mean formative score of male students without gamification (n=28)	74.75 ± 8.76	
Mean formative score of female students with gamification (n=29)	72.76 ± 22.02	0.326
Mean formative score of female students without gamification (n=22)	67.32 ± 15.18	
Mean formative score of male students with gamification (n=32)	75.66 ± 19.85	0.591
Mean formative score of female students with gamification (n=29)	72.76 ± 22.02	
Mean formative score of male students without gamification (n=28)	74.75 ± 8.76	0.045*
Mean formative score of female students without gamification (n=22)	67.32 ± 15.19	

*=statistically significant

difference *i.e.* p = 0.195. However, the gender-wise analysis showed that within Year 1 male students' formative scores with gamification were 85.37 ± 13.30, which was significantly higher than those tested without gamification *i.e.* 78.24 ± 16.56 (p = 0.045). A similar comparison between females in Year 1 with gamification (79.47 ± 17.05) and without gamification (79.93 ± 16.40) was not significant.

The comparison of Year 2 students' mean formative scores with and without gamification and their gender-wise analysis is shown in Table **2**.

The qualitative data was analyzed through the lens of self-determination theory (SDT). The main themes that were identified included autonomy, competence, relatedness, motivation, and experience. The students reported that the 'leaderboard' element of the gamified assessment served as a source of motivation to improve their performance. The same was reported to have an impact on their perceived competence. Confidence and achievement were linked to the 'immediate feedback' feature offered by both technology-enhanced tools, by the students. Immediate feedback associated with automated marking and display of results was viewed as an added benefit, contributing to improved teaching and learning experiences. Moreover, the leaderboard was reported as a source of 'relatedness' and 'motivation'. Lastly, 'memes' and 'music' were mentioned as reasons for improving the overall experience of gamification.

DISCUSSION

Gamification of assessment can be explored as a tool to increase student engagement, motivation, and thus learning outcomes. This study explored the impact of gamification on the academic performance of medical students in Pakistan.

Our study found that the mean formative score of male students in both year 1 and year 2 was higher than female students. This may be attributed to higher levels of test anxiety in females which have been documented in literature [7]. This increased test anxiety has been found to have a negative correlation with academic performance [8, 9]. Another possible reason for these results could be that these online assessments were not very strictly proctored and male students being more comfortable with cheating, may have sought help from each other to score better [10]. Cheating as a confounder is supported by earlier research that has shown that students achieve lower scores in proctored online assessments when compared with non-proctored assessments [11]. This emphasizes the need to implement stricter proctoring methods onsite, and online (such as vProctor), to increase the validity and reliability of online and gamified assessments [12].

In year 1, all comparisons between scores of formative assessments with and without gamification showed that formative scores of all groups were statistically similar except for one exception in which male students with gamification had significantly higher scores in assessments in comparison to scores of male students in formative assessments without gamification. This one exception could be due to the increased level of engagement offered by formative assessments with gamification which could have resulted in higher scores (85.37 ± 13.30 *vs.* 78.24 ± 16.57; p=0.04). This possible explanation is supported by the findings of a recent innovative study which after monitoring and comparing physiological variables such as pulse rate concluded that formative assessments with gamification tend to increase student engagement and may improve the performance of students in assessments [13].

When comparing the formative scores of Year 2 students in various groups, we found that the scores were generally similar, except for the comparison between male students without gamification and female students without gamification. In the latter case, male students without gamification scored significantly higher than female students without gamification (74.75 ± 8.76 *vs.* 67.32 ± 15.19, p=0.035). A possible reason for this discrepancy in results could be the fact that female students experience higher levels of test anxiety which can affect their performance [14-16]. The formative scores of Year 1 were consistently higher than Year 2 formative assessment scores in all year-wise comparisons regardless of the presence or absence of gamification in assessments. The probable reason for these consistent patterns could be the difficulty level of the two courses.

Technology-enhanced assessment has been received with mixed emotions by students and faculty worldwide. Particularly its compulsive use during the pandemic was viewed as a challenge [17-19]. However, when the use of TEA is not a necessity but a choice, it is received with

more enthusiasm as reflected by our findings and the existing literature [20-22].

In our study, the students related the feedback given by both the gamified and non-gamified tools as a source of instant gratification or admonition, which invariably resulted in motivation. The same effect of effective feedback on motivation has been reported in other studies [23, 24]. The leaderboard feature was unanimously reported as a source of connectedness and motivation by the participants of our study. Prior studies have also found that the leaderboard feature in gamified learning can influence student motivation, with both high and low rankings having an effect [25]. The exploration of students' overall experiences regarding gamification revealed that they viewed it as 'fun' and 'colorful' and 'entertaining', which is similar to the findings reported in the literature [26].

The positive impact of gamification, reflected in our study is in agreement with a recently published meta-analysis which revealed a moderately positive effect of gamification on academic performance (Hedges's $g = 0.782$, $p < 0.05$) across different factors such as geographical regions, education levels, and subjects [27]. Another systematic review highlighted gamification's positive impact on motivation, knowledge assimilation, and skill improvement, emphasizing the importance of creativity and adaptability in successful implementation [28].

CONCLUSION

The gamification of assessment holds promises as a strategy to enhance student academic performance through its impact on student engagement and motivation. This can be achieved through a careful alignment of game mechanisms and learning tasks, which requires further exploration in the context of medical education. Our study revealed disparities in academic performance between male and female medical students when using technology-enhanced assessment techniques, possibly attributed to test anxiety and the environment of online assessments. Notably, the integration of gamification demonstrated potential in improving scores for male students in a specific class, highlighting the need for further investigation with a larger sample size and varied demographics. While gamification has been met with mixed feelings globally, our findings suggest that when utilized as a choice rather than a necessity, it can generate enthusiasm among both students and faculty.

ETHICS APPROVAL

The study was approved by the Institutional Review Board of Shifa Tameer-e-Millat University and the Ethical Review Committee of Bahria Univeristy Medical & Dental College (IRB # 048-22, ERC 51/2022). All procedures performed in studies involving human participants followed the ethical standards of the institutional and/ or national research committee and the Helsinki Declaration.

CONSENT FOR PUBLICATION

Written informed consent was taken from the participants.

AVAILABILITY OF DATA

Available on demand.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

AUTHORS' CONTRIBUTION

M.A. Conceived the idea, designed the study, collected and analyzed the data, and drafted the manuscript

S.I. Designed the study, analyzed the data, and reviewed the final manuscript.

F.A. Designed the study, analyzed and interpreted the data, and drafted the manuscript.

A.S. Designed the study, analyzed and interpreted the data, and drafted the manuscript.

S.R. Designed the study and reviewed the final manuscript

H.A. Designed the study and reviewed the final manuscript

M.W.R. Analyzed and interpreted the qualitative data.

REFERENCES

1. Rusman E, Boon J, Martínez-Monés A, Jesús Rodríguez-Triana M, Retalis S. Towards the use of new methods for formative e-assessment of 21st century skills in schools. Available from: https://ceur-ws.org/Vol-1147/paper1.pdf
2. Rutledge C, Walsh CM, Swinger N, Auerbach M, Castro D, Dewan M, *et al.* Gamification in action: theoretical and practical considerations for medical educators. Acad Med 2018; 93(7): 1014-20.
DOI: https://doi.org/10.1097/ACM.0000000000002183
3. Alawadhi A, Abu-Ayyash EA. Students' perceptions of Kahoot!: An exploratory mixed-method study in EFL undergraduate classrooms in the UAE. Educ Inf Technol 2021; 26(4): 3629-58.
DOI: https://doi.org/10.1007/s10639-020-10425-8
4. Lumsden J, Edwards E, Lawrence N, Coyle D, Munafò M. Gamification of cognitive assessment and cognitive training: A systematic review of applications and efficacy. JMIR Serious Games 2016; 4(2): e11.
DOI: https://doi.org/10.2196/games.5888
5. Rehman F, Naiyar I, Umar M. Pattern and associated factors of exam anxiety among students of private medical college Pakistan. Pak J Med Health Sci 2020; 14(3): 669-71.
6. Ryan RM, Deci EL. Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. Contemp Educ Psychol 2020; 61: 101860.
DOI: https://doi.org/10.1016/j.cedpsych.2020.101860
7. Shaheen A, Azam F, Rabbani MW, Kazmi N. Correlation of computer-based test anxiety with medical students' performance before, during and after assessments. Pak J Med Sci 2022; 38(3): 476-80.
DOI: https://doi.org/10.12669/pjms.38.3.4989

8. Dawood Nawaz DU, Aqdas R, Shahzadi M, Iqbal S. Impact of test anxiety and mindfulness on academic performance among university students. Psychol Educ 2021; 58(2): 10662-73. DOI: https://doi.org/10.1080/0048590X.2021.1897169
9. Jerrim J. Test anxiety: Is it associated with performance in high-stakes examinations? Oxf Rev Educ 2023; 49(3): 321-41. DOI: https://doi.org/10.1093/oxrev/ogac003
10. Iqbal Z, Anees M, Khan R, Ara I, Begum S, Rashid A, Farooq H. Cheating during examinations: Prevalence, consequences, contributing factors and prevention. Int J Innov Creat Chang 2021; 15(6): 601-9. DOI: https://doi.org/10.13180/ijicc.2021.15.06.06
11. Vazquez JJ, Chiang EP, Sarmiento-Barbieri I. Can we stay one step ahead of cheaters? A field experiment in proctoring online open book exams. J Behav Exp Econ 2021; 90: 101653. DOI: https://doi.org/10.1016/j.jbe.2021.101653
12. Rezazadeh M, Tavakoli M. Investigating the relationship among test anxiety, gender, academic achievement and years of study: A Case of Iranian EFL university students. Engl Lang Teach 2009; 2(4): 68-74.
13. Hinds JA, Sanchez ER. The role of the hypothalamus–pituitary–adrenal (HPA) axis in test-induced anxiety: assessments, physiological responses, and molecular details. Stresses 2022; 2(1): 146-55. DOI: https://doi.org/10.3390/stresses2010011
14. Tsegay L, Shumet S, Damene W, Gebreegziabhier G, Ayano G. Prevalence and determinants of test anxiety among medical students in Addis Ababa Ethiopia. BMC Med Educ 2019; 19(1): 423. DOI: https://doi.org/10.1186/s12909-019-1859-5
15. Farooqi YN, Ghani R, Spielberger CD. Gender differences in test anxiety and academic performance of medical students Int J Psychol Behav Sci 2012; 2(2): 38-43. DOI: https://doi.org/10.6000/ijpbs.2012.2.2.3
16. Al-Ataby A. Technology-enhanced learning and teaching in COVID-19 era: Challenges and recommendations. Int J Innov Educ Res 2020; 8(10): 317-31. DOI: https://doi.org/10.34218/ijier.8.10.317
17. Kukulska-Hulme A. How should the higher education workforce adapt to advancements in technology for teaching and learning? Int High Educ 2012; 15(4): 247-54. DOI: https://doi.org/10.1016/j.iheduc.2012.06.003
18. Khan RA, Jawaid M. Technology enhanced assessment (TEA) in COVID 19 pandemic. Pak J Med Sci 2020; 36(COVID19-S4): S108. DOI: https://doi.org/10.12669/pjms.36.COVID19-S4.2795
19. Van Gaalen AE, Brouwer J, Schönrock-Adema J, Bouwkamp-Timmer T, Jaarsma AD, Georgiadis JR. Gamification of health professions education: a systematic review. Adv Health Sci Educ 2021; 26(2): 683-711. DOI: https://doi.org/10.1007/s40694-021-00203-1
20. Warfvinge P, Löfgreen J, Andersson K, Roxå T, Åkerman C. The rapid transition from campus to online teaching–how are students' perception of learning experiences affected? Eur J Eng Educ 2022; 47(2): 211-29. DOI: https://doi.org/10.1080/03043947.2022.2037780
21. Keane T, Linden T, Hernandez-Martinez P, Molnar A. University students' experiences and reflections of technology in their transition to online learning during the global pandemic. Educ Sci 2022; 12(7): 453. DOI: https://doi.org/10.3390/educsci12070453
22. Deeley SJ. Using technology to facilitate effective assessment for learning and feedback in higher education. Assess Eval High Educ 2018; 43(3): 439-48. DOI: https://doi.org/10.1080/02602938.2018.1459744
23. Alt D, Naamati-Schneider L, Weishut DJ. Competency-based learning and formative assessment feedback as precursors of college students' soft skills acquisition. Stud High Educ 2023; 48(12): 1901-17. DOI: https://doi.org/doi/abs/10.1080/03075079.2023.2217203
24. Na K, Han K. How leaderboard positions shape our motivation: the impact of competence satisfaction and competence frustration on motivation in a gamified crowdsourcing task. Internet Res 2023; 33(7): 1-8. DOI: https://doi.org/10.1108/IR-02-2023-0049
25. Iman N, Ramli M, Saridewi N. Kahoot as an assessment tools: Students' perception of game-based learning platform. J Penelit Pembelajar 2021; 7(2): 245-59. DOI: https://doi.org/10.30870/jppi.v7i2.8304
26. Zeng J, Sun D, Looi CK, Fan AC. Exploring the impact of gamification on students' academic performance: A comprehensive meta-analysis of studies from the year 2008 to 2023. Brit J Educ Technol 2024. DOI: https://doi.org/10.1111/bjet.12977
27. García-López IM, Acosta-Gonzaga E, Ruiz-Ledesma EF. Investigating the impact of gamification on student motivation, engagement, and performance. Educ Sci 2023; 13(8): 813. DOI: https://doi.org/10.3390/educsci13080813
28. Wang WT, Kartika Sari M. Examining the effect of the task-technology fit of game mechanisms on learning outcomes in online gamification platforms. J Educ Comput Res 2024; 61(8): 32-59. DOI: https://doi.org/10.1177/07356331231187285