EDUCATIONAL ARTICLE

Received: 12 MAY 2025 | Accepted: 19 MAY 2025 | **DOI:** https://doi.org/10.57849/ulisboa.fm.jscml.0000015.2025

JORNAL DA SOCIEDADE DAS CIÊNCIAS MÉDICAS DE LISBOA

Three Misconceptions in Medical Education: Student Evaluation of Teaching, Curricular Repetition, and Assessment Objectivity

Thomas Hänscheid [1,2], **Emília Valadas** [3], **João Costa** [1]

1 Universidade de Lisboa, Faculdade de Medicina, Departamento da Educação Médica, Avenida Prof. Egas Moniz, 1649-028, Lisboa, Portugal.

2 Universidade de Lisboa, Faculdade de Medicina, Instituto de Microbiologia, Avenida Prof. Egas Moniz, 1649-028, Lisboa, Portugal.

3 Universidade de Lisboa, Faculdade de Medicina, Clínica Universitária de Doenças Infecciosas, Avenida Prof. Egas Moniz, 1649-028, Lisboa, Portugal.

✉ **Corresponding author:**
Thomas Hänscheid
Universidade de Lisboa, Faculdade de Medicina, Instituto de Microbiologia
Avenida Prof. Egas Moniz
1649-028, Lisboa, Portugal
Email: *t.hanscheid@medicina.ulisboa.pt*

ABSTRACT: Three widely held assumptions in medical education may inadvertently hinder effective learning. First, student evaluations of teaching (SETs) are often treated as proxies for educational quality, yet they are shaped by factors unrelated to actual learning and provide limited actionable insight. Second, repetition in the curriculum is frequently dismissed as poor design rather than recognised as essential for mastery, reinforcing knowledge, developing fluency, and enabling long-term retention. Third, the overemphasis on objectivity in assessment, exemplified by multiple-choice questions (MCQs) and objective structured clinical examinations (OSCEs), may obscure important dimensions of clinical reasoning and are weak predictors of real-world clinical competence. Reconsidering these three assumptions may help realign educational strategies with their central goal: preparing students to become competent, reflective, and effective physicians.

KEY WORDS: Medical education, myths, repetition, student evaluation of teaching, objectivity

INTRODUCTION

The number three carries a powerful resonance in human cognition, education, and culture. It ranges from triads in diagnostic reasoning and vaccine schedules in medicine to the three-act structure in storytelling. The so-called "Rule of Three" (Latin: omne trium perfectum) reflects our preference for completeness, pattern recognition, and cognitive efficiency[1-3] and even in medical education, the power of three apparently has an effect [4]. In this spirit, three prevailing tenets of contemporary medical education merit critical re-examination: (i) the outsized role of students' evaluations of teaching; (ii) the drive to eliminate (curricular) repetition; and (iii) the overvalued ideal of objectivity in assessments.

1. Are students' evaluations of teaching (SETs) really that informative and relevant?

Do SETs, widely accepted and central to faculty appraisal systems, truly reflect educational quality, or are they merely popularity metrics? Despite their prominence, SETs are not strongly supported by current research [5]. A meta-analysis of over 100 different courses showed that students do not learn more from professors with higher SET since there is no significant correlation between SETs and actual learning [6], while a systematic review concluded that SETs are "questionable for high-stakes decisions," being heavily influenced by student bias and grade expectations [7]. In fact, a 2025 study across 160 veterinary courses found for the first time in veterinary medicine, a small but negative and statistically significant relationship between SET and an independent measure of learning [8]. Rather than reflecting educational value, SETs seem to mirror a broader trend toward instant ratings, reducing evaluations to a superficial popularity contest. While SETs may have limited utility, such as flagging an unprepared lecturer or unprofessional behaviour, they should be recognised for what they are: inherently subjective, often inconsistent, and frequently unrepresentative. They have little to do with what ultimately matters: whether students are learning and developing into competent physicians. Furthermore, it is at least unclear what the consequences are of a poor SET in terms of curriculum and education methodology changes, at least in medical schools.

2. In Praise of Repetition: Not a Design Flaw, but the Foundation of Mastery

The second misconception treats curricular repetition as a design fault rather than a pedagogical necessity. Students often describe repeated content as boring or redundant, an attitude perhaps intensified in today's mobile phone and social media generation, where novelty is constant, attention is fragmented, and initial enthusiasm fades quickly (novelty effect) [9]. In this mindset, long-term retention of knowledge is often sacrificed in the pursuit of the new. This perspective also feeds a deeper misconception: that medicine is about grasping complex concepts once (e.g. the Embden–Meyerhof pathway), and that once understood, repetition becomes unnecessary. While this may apply to isolated theories, most of medicine is different. It is more like learning a language, where progress depends on memorizing vast amounts of grammar and vocabulary, and fluency requires repeated reinforcement. If "appropriate antibiotic use" were a language, then its terminology and internal structure would need to be repeatedly driven deep, across pharmacology, microbiology, infectious diseases, and beyond, until students can speak it with fluency and precision.

Like in many areas of life, mastery in medicine rarely comes from novelty; it is built through repetition, feedback, and refinement. Rote memorisation, as used in language learning, is after all a form of repetition. While direct evidence on this is limited, studies consistently show that spaced repetition enhances learning and performance. Trials in microsurgical training, paediatrics rotations, and digital education all demonstrate that spaced repetition enhances knowledge, skills, and clinical outcomes [10-12]. A related idea is captured in Harden's spiral curriculum, which reinforces learning by systematically revisiting key topics at increasing levels of complexity [13]. A touch of novelty may still spark attention [14], but it is repetition that builds mastery.

3. Does our obsession with objectivity risk neglecting what really matters in assessment?

The third misconception is that only objective formats such as multiple-choice questions (MCQs) or objective structured clinical examinations (OSCEs) are the most valid ways to assess competence. This belief may be appealing, especially as it aims to address students' (often prioritized) demands for fairness and transparency. But it can also be misleading.

Curiously, nobody insists that a driving test be fully standardised or stripped of subjective judgement. While certain tasks, like parallel parking or emergency stops, are standardised, we ultimately care whether the driver in the next lane can actually drive, not just whether they passed a checklist of manoeuvres. That judgement depends on the examiner's expert ability to assess how well a person integrates knowledge and skills in real situations, something oral examinations in medicine are uniquely suited to capture. In contrast, this is difficult to assess in OSCEs, if not overlooked entirely, where assessment is fragmented into narrowly scripted tasks and scored against rigid checklists.

Standardization has a role in ensuring fairness, comparability, and clarity. Although OSCEs were a well-intentioned attempt to address inconsistencies in traditional assessments, over-reliance on their rigid format can limit what we are able to assess. They should be part of a broader assessment strategy, not the sole

measure, just as a driving examiner uses a checklist alongside real-time observation and judgement. A 2023 systematic review found structured oral exams (vivas) to be both reliable and well-accepted [15]. Meanwhile, a scoping review of OSCEs found that in 78% of studies, they correlated only weakly with written tests and real-world performance, particularly with poorly trained examiners [16]. Instead of overemphasising standardisation in the name of objectivity, should we not be asking the more important question: does the assessment, even at the cost of some subjectivity, help us judge whether a student is becoming a competent physician?

CONCLUSION

Modern educational practices often promise progress, but their impact on the real goal, producing competent physicians, is often uncertain. Some rest more on assumption than evidence, and not everything old is obsolete, especially if it worked. Student evaluations often reflect popularity more than learning, repetition remains the foundation of mastery, and rigid objectivity may obscure true competence. Ultimately, what matters is that our teaching and assessments truly help students become good doctors, equipped to deliver high-quality, patient-centred care.

DISCLOSURES: *The views and reflections expressed in this article are solely those of the authors and do not necessarily represent the official position of the School of Medicine with which they are affiliated.*
The authors used artificial intelligence (AI) tools – including GPT-4, Google Gemini, and DeepSeek (version/date of use: May 2025) – exclusively for language editing and clarity enhancement. No AI-generated content was used for original analysis, interpretation, or development of arguments. AI support was limited to grammar and stylistic refinement at the sentence level and did not involve the generation of original or substantive content.
No identifiable or sensitive data were shared with GPT-4, Google Gemini, or DeepSeek during the editing process. All usage complied with institutional privacy and security guidelines. The final manuscript was reviewed and approved by all authors to ensure accuracy and adherence to ethical standards for research and publication.

All authors declare no conflicts of interest. No funding was received for this work. Ethical approval was not required.

REFERENCES

1. Lease EB. The number three, mysterious, mystic, magic. Class Philol. 1919;14(1):56–73.
2. Vaux DL, Fidler F, Cumming G. Replicates and repeats—what is the difference and is it significant? EMBO Rep. 2012 Apr;13(4):291–6.
3. Microsoft Corporation. What is the Rule of Three in writing and communication? [Internet]. Redmond (WA): Microsoft; [date unknown] [cited 2024 May 16]. Available from: https://www.microsoft.com/en-us/microsoft-365-life-hacks/writing/what-is-rule-of-three.
4. Chu DI. The power of three [Internet]. Association for Academic Surgery; 2017 Jan 10 [cited 2024 May 16]. Available from: https://www.aasurg.org/blog/the-power-of-three/.
5. Constantinou C, Wijnen-Meijer M. Student evaluations of teaching and the development of a comprehensive measure of teaching effectiveness for medical schools. BMC Med Educ. 2022;22:113. doi:10.1186/s12909-022-03148-6
6. Uttl B, White CA, Wong Gonzalez DW. Meta analysis of faculty teaching effectiveness: SET ratings and student learning are not related. Stud Educ Eval. 2017;54:22–42.
7. Quansah F, Cobbinah A, Asamoah Gyimah K, Hagan JE. Validity of student evaluation of teaching in higher education: a systematic review. Front Educ. 2024;9:1329734.
8. Gilbert RO, Gilbert DR. Student evaluations of teaching do not reflect student learning: an observational study. BMC Med Educ. 2025;25:313. doi:10.1186/s12909-025-06896-3.
9. Elston DM. The novelty effect. J Am Acad Dermatol. 2021;85(3):565–566. doi:10.1016/j.jaad.2021.06.846
10. Teo WZW, Dong X, Mohd Yusoff SK, Das De S, Chong AKS. Randomized controlled trial comparing the effectiveness of mass and spaced learning in microsurgical procedures using computer aided assessment. Sci Rep. 2021;11(1):2810. doi:10.1038/s41598-021-82419-6.
11. Durrani SF, Yousuf N, Ali R, Musharraf FF, Hameed A, Raza HA. Effectiveness of spaced repetition for clinical problem solving amongst undergraduate medical students studying paediatrics in Pakistan. BMC Med Educ. 2024;24:676. doi:10.1186/s12909-024-05261-5.
12. Martinengo L, Ng MSP, Ng TDR, Ang YI, Jabir AI, Kyaw BM, Tudor Car L. Spaced Digital Education for Health Professionals: Systematic Review and Meta-Analysis. J Med Internet Res. 2024;26:e57760. doi:10.2196/57760.
13. Harden RM. What is a spiral curriculum? Med Teach. 1999;21(2):141–3.
14. Houillon A, Lorenz RC, Boehmer W, Rapp MA, Heinz A, Gallinat J, Obermayer K. The effect of novelty on reinforcement learning. In: Dreher JC, Tremblay L, eds. Progress in Brain Research. Vol. 202. Elsevier; 2013:415–439. doi:10.1016/B978-0-444-62604-2.00021-6
15. Abuzied AIH, Nabag WOM. Structured viva validity, reliability, and acceptability as an assessment tool in health professions education: a systematic review and meta-analysis. BMC Med Educ. 2023;23:531. doi:10.1186/s12909-023-04524-6.
16. Chang O, Holbrook AM, Lohit S, et al. Comparability of OSCEs and written tests for assessing medical student competencies: scoping review. Eval Health Prof. 2023;46:213–24.