

EXAMINING THE EXAM: NEW METRICS FOR IMPROVING THE PERFORMANCE OF MULTIPLE-CHOICE QUESTIONS

Rashika Agarwal^a, Di Warren^b

Presenting Authors: Rashika Agarwal (rashika.agarwal@sydney.edu.au); Di Warren (diana.warren@sydney.edu.au)

^aSchool of Mathematics and Statistics, University of Sydney NSW 2001, Australia

^bSchool of Mathematics and Statistics, University of Sydney NSW 2001, Australia

KEYWORDS: constructive alignment, belonging, multiple-choice exams

SUBTHEME: Assessment; Empowering Educators

ABSTRACT

In any unit, the credible assessment of learning outcomes is vital, and needs constant vigilance. Any misalignment between what is defined in the unit of study outline, what is taught in class, and what is assessed, has many negative consequences. Firstly, misalignment in assessment undermines tertiary standards and complicates transition into subsequent units. Secondly, misalignment in assessment weakens students' sense of agency and belonging, which can affect ongoing academic confidence and performance.

Our study focuses on the evaluation of multiple-choice questions (MCQs) in a final exam, using a case study from a large cohort (N=1234) in a first-year data science unit. First: we use multiple known metrics to gauge the quality of the exam questions, including the discrimination index (DI), difficulty index (DIFF), distractor efficiency (DE), and the Cronbach's alpha (CA). We discuss the strengths and weaknesses of each metric, and how together they can highlight questions that may indicate a potential gap in constructive alignment in the curriculum. Second: we propose new metrics to more deeply analyse each question, including a targeted distractor analysis, which reveals how distractors (the wrong multiple-choice answers) are functioning in practise, both across the cohort, and across the cohort divided into grades (P, CR, DI, HD). In total, our methods allow a detailed, transparent analysis of the performance of each multiple-choice question, which leads to concrete strategies empowering educators to improve a multiple-choice question bank for its next iteration.

REFERENCES

- Belay, L. M., Sendekie, T. Y., & Eyowas, F. A. (2022). Quality of multiple-choice questions in medical internship qualification examination determined by item response theory at Debre Tabor University, Ethiopia. *BMC Medical Education*, 22(1). <https://doi.org/10.1186/s12909-022-03687-y>
- Christian, D. S., Prajapati, A. C., Rana, B. M., & Dave, V. R. (2017). Evaluation of multiple choice questions using item analysis tool: a study from a medical institute of Ahmedabad, Gujarat. *International Journal of Community Medicine and Public Health*, 4(6), 1876. <https://doi.org/10.18203/2394-6040.ijcmph20172004>
- Date, A. P., Borkar, A. S., Badwaik, R. T., Siddiqui, R. A., Shende, T. R., & Dashputra, A. V. (2019). Item analysis as tool to validate multiple choice question bank in pharmacology. *International Journal of Basic & Clinical Pharmacology*, 8(9), 1999. <https://doi.org/10.18203/2319-2003.ijbcp20194106>
- Kilgour, J. M., & Tayyaba, S. (2015). An investigation into the optimal number of distractors in single-best answer exams. *Advances in Health Sciences Education*, 21(3), 571–585. <https://doi.org/10.1007/s10459-015-9652-7>

Proceedings of the Australian Conference on Science and Mathematics Education, The University of Canberra, 18 – 20 September 2024, page 8, ISSN 2653-0481.