Home > Journal > Social Sciences & Humanities > CE

Indexing   View Papers   Aims & Scope   Editorial Board   Guideline   Article Processing Charges

OPEN ACCESS

Validity, Reliability and Equivalence of Parallel Examinations in a University Setting

Author(s)

Bunmi S. Malau-Aduli, Justin Walls, Craig Zimitat

ABSTRACT

A key issue to address in the design and implementation of any assessment system is ensuring its reliability and validity. University assessment policies often require staff to prepare parallel examinations for students who are unable to sit the initial examination. There is little published literature to give confidence to staff or students that these examinations are indeed reliable or equivalent. This study was conducted to determine the validity, reliability and equivalence of two parallel examinations that have been developed under highly defined quality assurance (QA) processes in a university setting. Collated assessment results for all the 76 participants who sat the parallel examinations were subjected to statistical and correlational analysis to test for significant differences between mean scores and their associated standard deviations. Item analysis was conducted for each assessment by computing the difficulty index (DIF), discrimination index (DI) and Kuder-Richardson 20 (KR-20) reliability using classical test theory. Results indicated comparative proportions of difficulty, functional distractors and internal consistency of the assessment items on both examinations. Comparison of student performances in both examinations revealed that there was no significant difference in mean scores. However, a highly positive and significant correlation ($r = 0.82$) between student total scores in both examinations was evident. Approximately two thirds (62.5 %) of students with low scores in the first examination also achieved low scores in the second examination. Furthermore, two thirds of the students were ranked in the same order based on performance in both examinations. The established QA processes for assessment in the school provided a strong basis for the generation of multiple sources of data to support arguments for the validity of examinations. It is possible to develop valid, reliable and equivalent parallel tests in university settings with the presence of well-defined QA processes.

KEYWORDS

Parallel Examinations; Quality Assurance; Assessment

Cite this paper

References

[1]   Downing, S. M. (2002). Threats to the validity of locally developed multiple-choice tests in medical education: Construct-irrelevance variance and construct under-representation. Advances in Health Sciences Education, 7, 235-241. doi:10.1023/A:1021112514626

[2]   Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. Medical Education, 37, 830-837. doi:10.1046/j.1365-2923.2003.01594.x

[3]   Downing, S. M. (2004). Reliability: On the reproducibility of assessment data. Medical Education, 38, 1006-1012. doi:10.1111/j.1365-2929.2004.01932.x

[4]   Downing, S. M., & Haladyna, T. M. (1997). Test item development: Validity evidence from quality assurance processes. Applied Measurement in Education, 10, 61-82. doi:10.1207/s15324818ame1001_4

[5]   Downing, S. M., & Haladyna, T. M. (2009). Validity and its threats. In S. M. Downing, & R. Yudkowsky

(Eds.), Assessment in health professions education (pp. 21-55). London: Routledge.

[6]   Fowell, S. L., Southgate, L. J., & Bligh, J. G. (1999). Evaluating assessment: The missing link? Medical Education, 33, 276-281. doi:10.1046/j.1365-2923.1999.00405.x

[7]   Hamdy, H. (2006). Blueprinting for the assessment of health professsionals. The Clinical Teacher, 3, 175-179. doi:10.1111/j.1743-498X.2006.00101.x

[8]   Hays, R. (2008). Assessment in medical education: Roles for clinical medical educators. The Clinical Teacher, 5, 23 27. doi:10.1111/j.1743-498X.2007.00165.x

[9]   Jozefowicz, R. F., Koeppen, B. M., Case, S. M., Galbraith, R., Swanson, D. B., & Glew, R. H. (2002). The quality of in-house medical school examinations. Academic Medicine, 77, 156-161. doi:10.1097/00001888-200202000-00016

[10]   Kane, M. (2006). Content-related validity evidence in test development. In S. M. Downing, & T. M. Haladyna (Eds.), Handbook of test development (pp. 131-153). Mahwah, NJ: Lawrence Erlbaum Associates.

[11]   Malau-Aduli, B. S., Zimitat, C., & Malau-Aduli, A. E. O. (2011). Quality assured assessment processes: Evaluating staff response to change. Journal of Higher Education Management & Policy, 23, 1-23.

[12]   Malau-Aduli, B. S., & Zimitat, C. (2011). Peer review improves the quality of MCQ examinations. Assessment & Evaluation in Higher Education, 34, 1-13. doi:10.1080/02602938.2011.586991

[13]   Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-104). New York, NY: American Council on education and Macmillan.

[14]   Norcini, J., Anderson, B., Bollela, V., Burch, V., Costa, M. J., Duvivier, R., Galbraith, R., Hays, R., Kent, A., Perrott, V., & Roberts, T. (2011). Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 Conference. Medical Teacher, 33, 206-214. doi:10.3109/0142159X.2011.551559