



Assessing the Cognitive Domain through MCQs: Critical to Quality Assurance in Higher Education

Khairunnisa Aziz Dhamani* & Zeenatkhanu Kanji**

Abstract

While establishing assessment methods as a critical element of quality assurance, one must not lose sight of the intent to measure both expected learning outcomes and intended programme objectives. The achievement of quality assessment is rooted in ensuring that the content and approaches to delivery produce graduates and programmes which are high quality. One of the most common strategies to assessing higher order cognitive domain is via Multiple Choice Questions (MCQs). In this paper we will reflect on a quality assurance effort to ensure that this strategy was appropriate (doing the right thing), necessary (for the right reasons) and sufficient (achieving the desired results) in meeting the requirements. An exemplar of a continuing professional development approach to the constructing, measuring and evaluation of MCQs is provided, which emphasised the imperative of both faculty (peer) and organisational commitment to the process and achieving the desired product of competent and successful graduates.

Résumé

En établissant les méthodes d'évaluation comme élément clé de l'assurance qualité, l'on ne doit perdre de vue le but de mesurer les résultats d'apprentissage attendus et les objectifs visés pour le programme. La réalisation de l'évaluation de la qualité cherche à garantir la production de diplômés et de programmes de haute qualité à travers les contenus et les approches d'exécution. L'une des

* School of Nursing and Midwifery, Aga Khan University, Dar es Salaam, Tanzania.
Email: khairunnisadhamani@gmail.com

** School of Nursing and Midwifery, Aga Khan University, Dar es Salaam, Tanzania.
Email: zeenatkhanukanji70@gmail.com

stratégies les plus fréquemment utilisées dans l'évaluation du domaine cognitif de niveau supérieur est celle via les questions à choix multiples (QCM). Dans le présent article, notre réflexion sera axée sur l'effort de l'assurance qualité pour avoir une stratégie appropriée (bien faire), nécessaire (pour les bonnes raisons) et suffisante (obtenir les résultats escomptés) en répondant aux exigences. Un modèle de l'approche de développement professionnel continu est fourni pour construire, mesurer et évaluer les QCM, ce qui a mis en avant l'obligation de considérer l'engagement à la fois de la faculté (pair) et organisationnel dans le processus de réalisation du produit de diplômés compétents et qui réussissent.

Introduction

Assessment methods are designed to measure the expected learning outcomes and intended programme objectives (IUCEA 2010). They also form one of the critical elements of quality assurance mechanisms for the curriculum. Of importance to quality assurance is the role of assessment methods to improve results and/or catalyse students' learning outcomes. The expected learning outcomes of the course or the programme guide the faculty to develop teaching activities and to inform the design and implementation of appropriate assessment and evaluation strategies (McDonald 2007). Evaluation serves as one of the basic means of assuring quality in the teaching-learning process (Bourke and Ihrke 2012). Hence, there is an imperative for well-designed, appropriate and diverse instructional methods and assessment strategies in order to assess and produce competent graduates.

A variety of approaches are used to assess the learning outcomes in relation to various domains. This paper focuses on the authors' experiences of assessing a higher order cognitive domain by a well-constructed and analysed high stake multiple choice questions (MCQs). Additionally, the paper highlights the important role played by educators in preparing this kind of questions.

Quality Assurance Guidelines on Assessment

According to resources from the Inter-University Council for East Africa (IUCEA), when looking for quality in any aspect of higher education, there are three main guiding questions. First, 'Are we doing the right things?'; second, 'Are we doing the right things in the right way?' and third, 'Do we achieve our goals?' (IUCEA 2010:6).

In terms of utilising MCQs these questions lead the educators to scrutinise this strategy's potential to assess higher order cognitive and affective domains of students. In a recent effort to consider the effectiveness and efficacy (as part of the QA efforts) of MCQs, the faculty from the Aga Khan University in a Post-Registered Bachelor of Science in Nursing programme undertook

to ensure that the chosen strategy was appropriate in testing the required learning outcomes of the course.

Notionally it is important for educators and evaluators to be confident that the strategy being implemented for assessment is optimal – essentially the right method of assessment. However, choosing the right assessment method is not enough. Implicit and explicit to each method is a parallel aspect of building the measurement; hence, in the case of MCQs, knowing the technical aspects of constructing the right questions (Er, Ramamurthy and Pook 2014) including distractors and using a blueprint or Table of Specification (ToS). This knowledge and framework not only means increasing the likelihood of doing the right things in the right way every time, but also increases accountability for the decision path to such formulation and administration of MCQs. The cycle of assessment is not limited to the examination construction and administration (exam taking), but includes the faculty's capacity to analyse and interpret results of item analysis. This latter component contributes to the validity and reliability of the test instrument in the effort to ensure that the goal or the outcome was achieved. Finally, learners should receive timely feedback on their performance in the test as an element of the quality assurance programme. If we falter, omit or fail in any of these processes, quality is potentially compromised.

Doing the Right Thing: MCQs as a Tool of Assessment for Nursing Students

One of the most popular written assessment formats used is the standard four distractors MCQ tests with a forced choice, single best answer or response. If MCQs simply assess students' ability to recall information and/or comprehension, then their contribution to future work life and capacities may be limited and somewhat suspect. This has been a criticism of MCQs as pure recall content testing is simplistic and limited. Thus, there is a risk that an underdeveloped or minimalised assessment strategy will lead to superficial learning and miss the opportunity to engage students in higher order thinking such as application, analysis, synthesis and evaluation of knowledge acquired in their profession. It is, therefore, important that the assessment method is designed carefully to promote quality learning and teaching tools (Ramsden 2003).

There is lack of agreement among faculty members regarding the optimal number of distractors (i.e, four or five distractors per MCQ) as a means to decrease the percentage of random guessing the correct answer by the students. Vegada *et al.* (2016) carried out a random study with three groups of third-year medical students comparing three, four and five distractors and found that the three distractors per MCQ could be preferred above the four and five

distractors per MCQ. The quality parameters of MCQ test analysis showed no significant difference in reliability, validity and discriminatory index between the three groups (Vegada *et al.* 2016). Although there was normal distribution of scores and two functional distractors per question on the MCQ test for all three groups in this study, the mean score was higher while the average time taken for completing the exam was less for the group who had three distractors in the MCQ test. Similarly, Tarrant and Ware (2010) compared four distractors of MCQ test by re-writing the same questions as three distractors questions. Item analysis was done to remove the distractor with the least response. Comparisons of the three and four distractors given to two student cohorts over two academic years showed minimal difference in the item analysis of difficulty and discriminatory index (Tarrant and Ware 2008). The findings indicated that both three and four distractors of MCQs work equally well. Moreover, MCQs with three distractors take less time to construct and administer. It also means that more content could be tested by giving the students more MCQs in a single examination. Hence, there appears to be both rationale and merit in using three distractors per MCQ (Rodriguez, 2005).

This criticism and scrutiny of MCQs has been taken very seriously in professional education, such as nursing, where faculty and programme directors recognise that there is an imperative to use such measurements to test higher levels of cognitive development. MCQs are the foundation of a number of national nursing examination platforms such as the NC-LEX®, which implies a high level of confidence in this assessment strategy to measure beyond content. The strengths of MCQs include the ability for the faculty to directly assess content and course objectives in a succinct and direct manner – more so than other written assessment methods. MCQ tools are easy to administer and score, which enables mass assessment of students. And, perhaps of greatest importance, is the ease of item analysis which enables a rapid and measurable quality assessment of each item (Vegada *et al.* 2016).

For the Right Reasons? MCQs and Higher Order Understanding in Nursing Students

In recent years, in nursing programmes, we have seen a capacity growth in the use, development and assessment of MCQs. The existence of evidence-based guidelines for building, administering and assessing MCQs further provides the rationales for this approach (Bandaranayake 2008; Considine, Botti and Thomas 2005; Haladyn, Downing and Rodriguez 2002). One of the mechanisms has been the use of MCQs constructed in parallel with clinical vignettes (often actual scenarios from the practice settings), which allow appropriate assessment of nursing students' theoretical knowledge and application to nursing practice.

Achieving our Goals: MCQs as Capacity and Quality Assurance Measures

Nurse educators have both the legal and ethical responsibility to ensure the evaluation of students is valid, effective and reflects their potential to practise safely and knowledgeably (National League for Nursing 2005; Tarrant, Knierim, Hayes and Ware 2006). However, in many cases, the members of nursing faculties are recruited based on clinical and academic achievements, with little consideration of educator capacities, often leading to a disjoint between what they know about teaching/assessment and what they think they know (Ramsden 2003). According to McDonald (2007), the quality of a measurement instrument, such as test/exam, depends on assessment competency of the faculty. However, when inadequacy is experienced in the faculty, the instruments developed tend to provide invalid and unreliable results (Downing 2005; Tarrant *et al.* 2006). Hence, one way to improve tests/exams is to enhance nurse educators' skills in identifying learning outcomes of the course being assessed; developing ToS; constructing exam/test items; and analysing results of the exam (McDonald 2007).

A well-developed test/instrument using a ToS promotes validity and enhances the likelihood of items in the test matching outcomes and content (McDonald 2007). However, it requires time and proper planning on behalf of the faculty. Historically, many faculties fail to recognise the linkage which often yields test items which do not test the required higher level cognitive skills (Morrison and Free 2001).

An Exemplar

In 2012, the primary author of this paper conducted continuing professional development (CPD) sessions for the nursing faculty on developing a ToS or a blueprint in order to address this perceived deficit within a small nursing faculty at a private university in East Africa. The CPD was considered incomplete and insufficient as the faculty continued to face the challenge of constructing appropriate test/exam items to measure the desired content and outcomes. It was realised that regardless of the integrity and intactness of the ToS, without proper construction of test/exam items, the exam was invalid. Therefore, in October 2014, a workshop on developing MCQs test items was conducted. During the workshop, a sample of previously used MCQs was provided for the faculty to assess the quality of questions. The faculty realised that many questions were ambiguous, had grammatical and spelling errors and lacked the focus of what was being tested. Of greater importance, was the recognition that the faculty lacked skills in interpreting

test scores. As a result of this foundational gap, the second author of this paper started offering CPD to a few faculty members on item analysis, with the intention of a formal roll out of this education programme in faculties of two other countries of East Africa. Statistically analysing test data assures that tests are functioning as intended (McDonald 2007).

CPD on item analysis was given to a few faculties and this exercise was done for three semesters. This meant covering all the nursing courses in the undergraduate nursing programme. The faculty carried out item analysis of the MCQs of their respective courses. As each faculty became familiar with the process, they realised the importance of item analysis. For example, they realised the errors with their answer key, having ambiguous distractors, increasing chances of guesswork by having correct distractor longer or more complete than others, a word or phrase included in the stem and in the correct answer, and the need to improve/revise test items for future use.

As a step towards quality assurance, the nursing department of the university started a peer review process of exam papers developed by the content expert faculty member who taught the course. This effort included reviewing the MCQs as well as short and long essay questions to identify any gaps in the exam paper comparing with the outline of the ToS and for any flaws in the construction of the questions before examining the students. Questions were revised according to the gaps identified. Presently this exercise occurs in every semester prior to the exam time; however, item analysis of MCQs continues to be done by only a few faculties. If every faculty engages in this two-pronged approach, it will ensure the reliability and validity of the test questions (Sadaf, Khan and Ali 2012). The future plan is to roll out CPD sessions on item analysis for all the nursing faculties. This will provide an opportunity to write good MCQs, as well as improve teaching, and student learning outcomes (Talebi, Ghaffari, Eskandarzadeh and Oskouei 2013).

Conclusion

Assessment strategies need to receive the attention they deserve and it is critical to enhance educators' skills in identifying learning outcomes being assessed, developing a table of specification and constructing and analysing high quality test items to assess the desired learning outcomes. Life-long learning activities, including CPDs, will facilitate and expand the expertise in teaching and education, thereby providing faculty members with capacities and experiences to improve their skills and knowledge regarding MCQs. In addition, there is a clear requirement for organisational commitment to the process, such as peer review opportunities of evaluative tools, in order to achieve the desired product of competent and successful graduates.

References

- Bandaranayake, R., 2008, 'Setting and maintaining standards in multiple choice examinations: AMEE guide No.37', *Medical Teacher*, No. 30, pp. 836–845. Available online at <https://doi.org/10.1080/01421590802402247>.
- Bourke, M. P., and Ihrke, B. A., 2012, 'The evaluation process: an overview', in D. M. Billings, and J. A. Halstead, *Teaching in nursing: a guide for faculty*, 5th ed., pp. 422–440, St. Louis, MO: Elsevier Saunders.
- Considine, J., Botti, M. and Thomas, S., 2005, 'Design, format, validity, and reliability of multiple choice questions for use in nursing research and education', *Collegian*, No. 12, pp. 19–24. Available online at [https://doi.org/10.1016/S1322-7696\(08\)60478-3](https://doi.org/10.1016/S1322-7696(08)60478-3).
- Downing, S., 2005, 'The effects of violating standard item writing principles on tests and students: The consequences of using flawed test items on achievement examinations in medical education', *Advances in Health Sciences Education*, No. 10, pp. 133–143. Available online at <https://doi.org/10.1007/s10459-004-4019-5>.
- Er, H. M., Ramamurthy, S., and Pook, P. C. K., 2014, 'Can learning outcomes in cognitive domain be assessed effectively using multiple choice questions? A study in an undergraduate pharmacy curriculum'. *International E-Journal of Science Medical and Education (IeJSME)*, Vol. 8, No. 3, pp. 9–18. Available online at <https://doi.org/10.1007/s10459004-4019-5>.
- Haladyna, T., Downing, S., and Rodriguez, M., 2002, 'A review of multiple-choice item writing guidelines for classroom assessment', *Applied Measurement in Education*, Vol. 15, No. 3, pp. 309–334. Available online at https://doi.org/10.1207/S15324818AME1503_5.
- Inter-University Council for East Africa (IUCEA), 2010, *A Road Map to Quality: Handbook for Quality Assurance in Higher Education, Volume 1: Guidelines for Self-assessment at Program Level*, The Inter-University Council for East Africa/DAAD. Available online at <http://iucea1.org/qawp/wp-content/uploads/2013/12/Handbook-Volume-11.pdf>.
- McDonald, M. E., 2007, *The Nurse Educator's Guide to Assessing Learning Outcomes*, Sudbury, MA: Jones & Bartlett.
- Morrison, S., and Walsh Free, K. (2001). 'Writing multiple choice test items that promote and measure critical thinking', *Journal of Nursing Education*, 40(1), 17–24.
- National League for Nursing, 2005, 'Core competencies of nurse educators'. (<http://www.nln.org/profdev/corecompetencies.pdf>).
- Ramsden, P., 2003, *Learning to Teach in Higher Education*, Routledge New York: Falmer.
- Rodriguez, M. C., 2005, 'Three options are optimal for multiple choice items: a meta-analysis of 80 years of research', *Educational Measures Issues and Practices*, No. 24, pp. 3–13. Available online at <https://doi.org/10.1111/j.1745-3992.2005.00006.x>.
- Sadaf, S., Khan, S., and Ali, S. K., 2012, 'Tips for developing a valid and reliable bank of multiple choice questions (MCQs)', *Education for Health*, Vol. 25, No. 3, pp. 195–197. Available online at <https://doi.org/10.4103/1357-6283.109786>.

- Talebi, G. A., Ghaffari, R., Eskandarzadeh, E., and Oskouei, O. E., 2013, 'Item analysis an effective tool for assessing exam quality, designing appropriate exam and determining weakness in teaching', *Research and Development in Medical Education*, Vol. 2, No. 2, pp. 69–72. Available online at <https://doi.org/10.5681/rdme.2013.016>.
- Tarrant, M., Knierim, A., Hayes, S.K., and Ware, J., 2006, 'The frequency of item writing flaws in multiple choice questions used in high stakes nursing assessments', *Nurse Education Today*, Vol. 26, No. 8, pp. 662–671. Available online at <https://doi.org/10.1016/j.nedt.2006.07.006>.
- Tarrant, M., and Ware, J., 2008, 'Impact of item-writing flaws in multiple choice questions on student achievement in high-stakes nursing assessments', *Medical Education*, No. 42, pp. 198–206. Available online at <https://doi.org/10.1111/j.1365-2923.2007.02957.x>.
- Tarrant, M., and Ware, J., 2010, 'A comparison of the psychometric properties of three – and four-option multiple-choice questions in nursing assessments', *Nurse Education Today*, No. 30, pp. 539–543. Available online at <https://doi.org/10.1016/j.nedt.2009.11.002>.
- Vegada, B., Shukla, A., Khilnani, A., Charan, J., and Desai, C., 2016, 'Comparison between three option, four option and five option multiple choice question tests for quality parameters: A randomized study', *Indian Journal of Pharmacology*, Vol. 48, No. 5, pp. 571–575. Available at <https://doi.org/10.4103/0253-7613.190757>.

