Available online at www.sciencedirect.com

ScienceDirect

Currents in Pharmacy Teaching and Learning 8 (2016) 160–163

Currents in Pharmacy Teaching & Learning

http://www.pharmacyteaching.com

Research

The impact of a faculty development seminar on the quality of multiple-choice questions

Eric C. Nemec, PharmD, BCPS*, Beth Welch, PharmD

Department of Pharmacy Practice, Western New England University College of Pharmacy, Springfield, MA

Abstract

Introduction: Multiple-choice questions (MCQ) are almost ubiquitous in higher education. The composition of quality MCQ is an art that takes practice and multiple revisions. Internal evaluations within other health profession disciplines indicate that the quality of faculty developed MCQ are relatively poor.

Materials and methods: A 1.5 hour faculty development seminar was developed regarding the construction and the assessment of MCQ. Faculty submitted MCQ both pre-seminar and post-seminar and these were assessed against a quality checklist and scoring system. Faculty were surveyed pre-session and post-session electronically regarding their confidence in constructing MCQ.

Results: The mean pre-session MCQ quality score was 16.2 (SD = 2.4), which after revision improved to 18.1 (SD = 1.5, $p < 0.05$). Results of the individual learner strengths and needs assessment indicated that confidence of faculty identifying good and bad questions in their own work improved.

Conclusion: Faculty development seminars can improve the quality of in-house MCQ and improve faculty confidence in constructing new MCQ.

Keywords: Multiple-choice questions; Assessment; Faculty development

Introduction

The use of multiple-choice questions (MCQ) is a frequently used method of assessment in pharmacy education. From the Pharmacy College Admission Test (PCAT), to pharmacy program course exams, culminating in the North American Pharmacist Licensure Examination (NAPLEX), and perhaps a Board of Pharmaceutical Specialties (BPS) examination, multiple-choice questions are a common thread in these assessments. In comparison to more qualitative assessments such as short answer and essay assignment, MCQs are reliable, less labor intensive to grade, and allow faculty to trend areas of understanding or difficulty.

However, there are differences between the process of vetting questions for a standardized assessment such as the PCAT or NAPLEX and internal pharmacy program course exams. Standardized exams have a more rigid vetting process in comparison to in-house exams. For example, the NAPLEX exam includes 185 questions, 35 of which are "trial balloon" questions, which the National Association of Boards of Pharmacy are considering for possible inclusion in future NAPLEX exams. These questions are thoroughly evaluated for content accuracy, style, fairness, and psychometric properties before being included in a high-stakes assessment, such as the licensing examination. For in-house course exams, internal review processes have been effective in improving item quality. However, these reviews are labor

* Corresponding author: Eric C. Nemec, PharmD, BCPS, Department of Pharmacy Practice, Western New England University College of Pharmacy, 1215 Wilbraham Rd., Springfield, MA 01119.

E-mail: eric.nemec@wne.edu

http://dx.doi.org/10.1016/j.cptl.2015.12.008

1877-1297/

intensive and can limit their utility.[1] To compound this problem, pharmacy faculty are typically not trained through formal educational pathways in writing MCQ. While teaching certificates exist for pharmacy residents, there is no information on how many new faculty obtain a teaching certificate and whether these certificate programs include exam-writing instruction.[2]

A review of the health profession education literature did not yield any relevant references specific to pharmacy education; however, there are examples of the impact of poorly written MCQ in both medical and nursing education. In 1998, Jozefowicz compared the quality of in-house medical education examination questions written by faculty to United States Medical Licensing Examination (USMLE) questions written by experienced faculty. The in-house developed questions were of relatively low quality as compared to those written by faculty for USMLE.[3] Some research suggests that many in-house medical examination MCQs violate general item-writing guidelines.[4–6] For example, Kahn et al.[4] retrospectively analyzed MCQs submitted by faculty for "testwiseness" (cues for the keyed answer) and irrelevant difficulty flaws at a medical college over a three-year period (2009–2011). They reported that approximately 37% of final examination MCQs ($N = 4500$) were flawed.[4] The annual flaw rates ranged from 21% in 2011 to 67% in 2009. While this research showed that these flaws had little effect on test score reliability, they did introduce systematic error and disadvantaged some medical students. Additionally, these flaws can introduce cues to students that make it easier for some students to answer the question correctly based upon their test taking skills and not their knowledge base. Examples of this include grammatical cues, the use of absolute terms, heterogeneity between choices, and other logical cues.[4]

Within nursing education, it has been reported that multiple-choice questions within a hospital assessment system were fraught with item-writing flaws. Additionally, it was noted that many questions were written to assess low-level cognitive processes and were not appropriately linked to learning objectives.[7] Furthermore, analysis of MCQ over a five-year period from one nursing department identified a number of problems. Almost half (46.2%) of these questions violated item-writing guidelines, most (90%) were written at low cognitive levels, and few were original, with 36.2% of items taken from test banks.[8] Overall, there is a negative effect on students pursuant to violating item-writing guidelines with flawed items being inappropriately more difficult than standard items.[9] In comparison to non-flawed items, MCQ that are flawed are more difficult for students to answer. A non-experimental study by Downing[9] that examined item flaws concluded that flaws introduce systematic error and reduce the validity of assessment. This construct irrelevant variance led to 10–15% of students failing a test based solely on item flaws. Conversely, when Tarrant et al.[10] analyzed flawed items, they found that there was no significant difference in difficulty. Furthermore, it was described that high-achieving students were more likely to be penalized by flawed items than borderline students.[10] The difficulty in answering may be that the poor quality of an item confuses student understanding of what the actual question is. It is not that the question is a more difficult item in problem solving or cognitive knowledge assessment, rather a lack of clarity causes students to guess what the actual question is.

While results are mixed regarding which population of students is affected more by item-writing flaws, it is clear that there is a negative impact on students and therefore faculty should work to improve item writing. Within medical education, multi-day faculty development seminars for item-writing skills are effective in improving faculty ability to adhere to item-writing guidelines.[6] However, multi-day sessions are not feasible for many institutions.[6,11] The authors of this project hypothesized that a 1.5-hour faculty development seminar on MCQ-writing principles would improve the MCQ-writing skills of pharmacy faculty.

Methods

A faculty development seminar was designed to present tenets of A-Type MCQ construction and assess the impact of training on MCQ quality. The hypothesis of the project was that if faculty could adhere to established question-writing guidelines, the quality of their own questions would improve. All College of Pharmacy Faculty were invited via an email survey (SurveyMonkey®; Palo Alto, CA) to attend the 1.5-hour faculty development seminar on writing MCQ. Of the 24 faculty who indicated that they would attend the first session, approximately 68% were assistant professors, 20% were associate professors, and the remaining 12% were full professors.

In order to provide adequate background knowledge, the faculty development seminar presented a brief overview of assessment theory. This included information on writing and mapping course goals to lecture objectives and their link to formal assessment. The majority of the session was dedicated to creating items that utilized quality stems and appropriate distractors in a vignette format suitable for clinical education. The item metrics of the quality checklist (Table 1) were woven throughout this part of the seminar. The session concluded by addressing how MCQ can be used to assess above the Knowledge level within Bloom's taxonomy of the cognitive domain. While questions that assess domains above Knowledge can be more difficult to construct, they are possible. For example, asking a student to identify unspecified assumptions within a clinical case vignette can assess Bloom's Analysis domain.

An individual learning needs assessment was administered to participants both pre-session and post-session via SurveyMonkey (Palo Alta, CA). This survey evaluated a faculty member's confidence in identifying quality and poor MCQ. Attendees were instructed to bring three of what they considered to their best, de-identified MCQ that were