Assessment Strategies: Multiple Choice Item Writing

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied measurement in education*, *15*(3), 309-333.(Published in 2002, still considered the gold standard)

Abstract: A taxonomy of 31 multiple-choice item-writing guidelines was validated through a logical process that included two sources of evidence: the consensus achieved from reviewing what was found in 27 textbooks on educational testing and the results of 27 research studies and reviews published since 1990. This taxonomy is mainly intended for classroom assessment. Because textbooks have potential to educate teachers and future teachers, textbook writers are encouraged to consider these findings in future editions of their textbooks. This taxonomy may also have usefulness for developing test items for large-scale assessments. Finally, research on multiple-choice item writing is discussed both from substantive and methodological viewpoints.

Masters, J. C., Hulsmeyer, B. S., Pike, M. E., Leichty, K., Miller, M. T., & Verst, A. L. (2001). Assessment of multiple-choice questions in selected test banks accompanying text books used in nursing education. *Journal of Nursing Education*, *40*(1), 25-32.

Abstract: The purpose of this study was to assess multiple-choice questions used in test-banks accompanying selected nursing textbooks. A random sample of 2913 questions was selected from a convenience sample of 17 test banks. Questions were evaluated on (a) adherence to generally accepted guidelines for writing multiple-choice questions; (b) cognitive level as defined by Bloom's (1961) taxonomy; and (c) distribution of correct answers as A, B, C, or D. The results were 2233 violations of item-writing guidelines, most of which were minor but some were serious. A large number of questions (47.3%) were written at the knowledge level and only 6.5% were written at the analysis level. The correct answers were evenly distributed: C2S ranged from 0.00 to 4.84; chi square value needed to reach .05 probability was 26.30. Faculty are encouraged to evaluate multiple-choice questions from test banks carefully before using them for exams.

McDonald, M. (2007). *The nurse educator's guide to assessing learning outcomes*. Jones & Bartlett Learning. (BOOK)

Abstract: As multiple choice exams are a fundamental component of the assessment of learning outcomes in the classroom setting, this text consolidates all of the available information for the development of classroom exams, and focuses specifically on the systematic development of multiple-choice exams in nursing education. The second edition includes the following topics: assessment of critical thinking, learning objectives and outcomes, development of tests, creation of multiple choice items, and analysis of test reliability.

Morrison, S., & Free, K. W. (2001). Writing multiple-choice test items that promote and measure critical thinking. *Journal of Nursing Education*, *40*(1), 17-24.

Abstract: Faculties are concerned about measurement of critical thinking especially since the National League for Nursing Accrediting Commission cited such measurement as a requirement for accreditation (NLNAC, 1997). Some writers and researchers (Alfaro-LeFevre, 1995; Blat, 1989; McPeck, 1981, 1990) describe the need to measure critical thinking within the context of a specific discipline. Based on McPeck's position that critical thinking is discipline specific, guidelines for developing multiple-choice test items as a means of measuring critical thinking within the discipline of nursing are discussed. Specifically, criteria described by Morrison, Smith, and Britt (1996) for writing Critical-thinking multiple-choice test items are reviewed and explained for promoting and measuring critical thinking.

Oermann, M. H., & Gaberson, K. B. (2016). *Evaluation and testing in nursing education*. Springer Publishing Company.

Abstract: Considered the "gold standard" for evaluation and testing in nursing education, this fifth edition of the classic text helps educators to assess the level of learning achieved in the classroom, in clinical settings, and online with expanded coverage of essential concepts in assessment, evaluation, and testing in a wider variety of learning environments. It presents new content on evaluation in online programs and testing and features a new chapter on using simulation for assessment and high-stakes evaluations. Also included is updated information on clinical evaluation and program evaluation along with current research featuring new examples and tools. The fifth edition expands content on standardized tests, including how to write test items for licensure and certification exam prep, and provides new information on developing rubrics for assessing written assignments.

The fifth edition is distinguished from other texts by its focus on developing a framework that integrates all types of evaluation in a nursing program. It addresses how to develop a test blueprint and assemble, administer, write, and analyze tests. It provides rubrics for scoring tests and written assignments along with examples. Its guidelines for preventing cheating and conducting productive post-test discussions are especially helpful to educators. Additionally, the book explores important social, ethical, and legal issues associated with testing and evaluation.

Sutherland, K., Schwartz, J., & Dickison, P. (2012). Best practices for writing test items. *Journal of Nursing Regulation*, *3*(2), 35-39.

Abstract: This article enhances the current literature’s guidance for those interested in developing, assessing, or utilizing items to test competency in nursing. It does so by underlining the purpose of a test item and deriving four item-writing principles from that central purpose. The article then relates these four principles to the task of developing multiple-choice nursing test items. These principles include the strong alignment of a test item with the measurement of

the examinee's grasp of knowledge, the inclusion of important rather than nonessential concepts, the differentiation of competent from incompetent examinees by items of appropriate difficulty, and the creation of a fair examination consisting of appropriate and applicable items. Within the framework provided by these four principles, this article suggests some effective methods for devising nursing test items that remain faithful to the central purpose of assessing proficiency.

6.18.18