Designing good Multiple-choice Questions

Prof. Paul Lam

Centre for Learning Enhancement And Research

The Chinese University of Hong Kong

Assessment-driven learning

 "Assessment drives student learning. Student assessment can be designed to foster the development of elaborated knowledge structure by making relationships and understanding rather than isolated facts—the objects of assessment." (Bordage, 1994)

Assessment purposes

- Measure student's competence in course?
- Develop competent physicians?
 - Application of knowledge
 - Clinical decision-making

Advantages of using MCQs

- Accurate and efficient scoring; expedite the scoring of exams that are machinery graded (Brady, 2005; McDonald, 2002)
- Objective measurement (Brady, 2005; McDonald, 2002)
- Easier to control the difficultly level (McDonald, 2002)



Limitations of using MCQs

- Time consuming to construct a set of MCQ items (Brady, 2005)
- Being criticized as only able to test students' retention, superficial knowledge (McDonald, 2002; Scouller, 1998)



5

Aim higher

- Assess wide range of learning outcomes at cognitive level (evaluation, comprehension, application and analyzing skills) (Brady, 2005; McDonald, 2002)
- Able to challenge students' ability of critical thinking skills by using plausible alternatives (Brady, 2005)



Example I: Recall Question

- The dose of intravenous heparin should be adjusted to maintain the client's activated partial thromboplastin time (APTT) at how many times the control?
 - A. Less than 1.5
 - B. Between 1.5 and 2.5 *
 - C. Greater than 2.5 and less than 3.5
 - D. Between 3.5 and 4.5

- A client who is receiving intravenous heparin has an activated thromboplastin time (APTT) of 2.5 times the control. In addition to documenting the finding, which of these actions would be appropriate for a nurse to take?
 - A. Call the lab for a stat repeat of the test.
 - B. Discontinue the client's heparin infusion immediately.
 - C. Continue to monitor the client.*
 - D. Alert the blood bank to have a unit of packed cells available.

9

Example 2: Recall Question

 A nurse is aware that a screening test has high specificity. This means that the test

- A. provides precise findings.
- B. correctly indentifies those who have a disease.*
- C. accurately indentifies those who do not have a disease.
- D. has a high correlation with severity of disease.

(McDonald, 2002) 10

- A nurse who is planning a health screening program identifies that a particular screening test has a specificity of 90%. The nurse should recognize that this indicates that the test accurately indentifies
 - A. 10% of those who actually have the disease.
 - B. 10% of those who do not have the disease.
 - C. 90% of those who actually have the disease.*
 - D. 90% of those who do not have the disease.

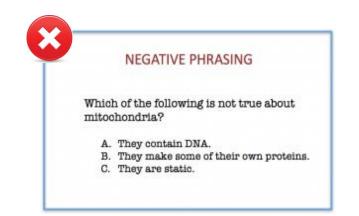
11

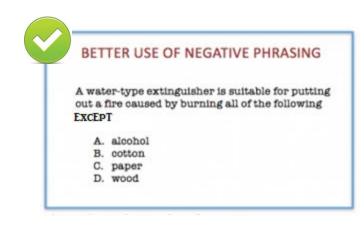
- A nurse plans all of these measures for a client who was rescued from a fire and has deep burn injury of the chest and arms. To which of these measures should the nurse assign priority during the emergent phrase of burn management?
 - A. Monitoring the client's respiration.*
 - B. Assessing the client's peripheral circulation.
 - C. Measuring the client's urine output.
 - D. Preventing the client from developing infection.

- An elderly client is about to have a minor surgical procedure. The client says to a nurse, "I really don't know why it is so important for me to have this surgery." The nurse notes that the client has signed a consent for the surgery. Before administering the client's preoperative medication, which of these actions would be appropriate for the nurse to take?
- A. Discuss nonsurgical treatments with the client and document the discussion in the client's medical record.
- B. Reassure the client that this is minor surgery and the surgeon has an impeccable reputation for performing only surgery that is beneficial for clients.
- C. Contact the client's adult children to determine if they understand the need for the surgical procedure.
- D. Inform the surgeon that the client does not understand the need for the surgery.*

Constructing an Effective Stem

- The stem should be negatively stated <u>only when</u> significant learning outcomes require it.
 - Students often have difficulty understanding items with negative phrasing.
 - If a significant learning outcome requires negative phrasing, such as identification of dangerous laboratory, the negative element should be emphasized with *italics* or CAPITALIZATION.



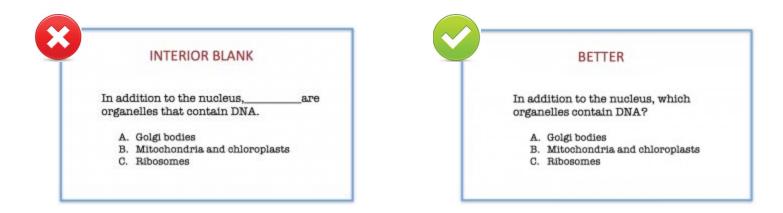


Source: http://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/

Constructing an Effective Stem

2) A question stem is preferable.

It allows the student to focus on answering the question rather than holding the partial sentence in working memory and sequentially completing it with each alternative.



Source: http://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/

Constructing Effective Alternatives

3) Alternatives should be stated clearly and concisely.

Items that are excessively wordy assess students' reading ability rather than their attainment of the learning objective.

WORDY ALTERNATIVES

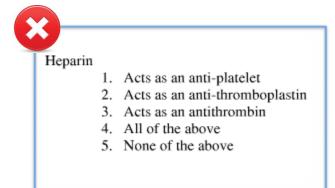
The term hypothesis, as used in research, is defined as

- A. A conception or proposition formed by speculation or deduction or by abstraction and generalization from facts, explaining or relating an observed set of facts, given probability by experimental evidence or by factual or conceptual analysis but not conclusively established or accepted.
- B. A statement of an order or relation of phenomena that so far as is known is invariable under the given conditions, formulated on the basis of conclusive evidence or tests and universally accepted, that has been tested and proven to conform to facts.
- C. A proposition tentatively assumed in order to draw out its logical or empirical consequences and so test its accord with facts that are known or may be determined, of such a nature as to be either proved or disproved by comparison with observed facts.

Source: Steven J. Burton, Richard R. Sudweeks, Paul F. Merrill, and Bud Wood. How to Prepare Better Multiple Choice Test Items: Guidelines for University Faculty. 1991.

Constructing Effective Alternatives

- 4) The alternatives "all of the above" and "none of the above" should NOT be used.
 - When "all of the above" is used, test-takers who can identify more than one alternative as correct can select the correct answer even if unsure about other alternative(s).
 - When "none of the above" is used, test-takers who can eliminate a single option can thereby eliminate a second option. In either case, students can use partial knowledge to arrive at a correct answer.

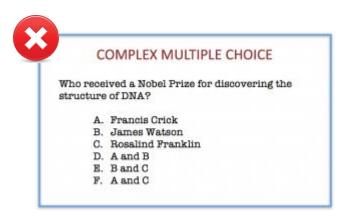


Source: http://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/

Constructing Effective Alternatives

5) Avoid complex multiple choice items.

- > Alternatives consist of different combinations of options.
- > As with "all of the above" answers, a sophisticated test-taker can use partial knowledge to achieve a correct answer.



Source: http://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/

Making judgment

- Subjective methods
- Item analysis: performance of MCQ in an exam

Difficulty factor

- Proportion of respondents selecting the right answer to that item
- $\mathbf{D} = \mathbf{c} / \mathbf{n}$
- **D** = difficulty factor
- c = number of correct answers
- n = number of respondents
- Range 0 -1
- The HIGHER the difficulty factor the **easier** the question is

Discrimination index

- How well the item discriminates between students who performed highest on the exam (e.g. top 25%) and students who performed lowest on the exam (bottom 25%).
 - DI = (a-b) / n
 - a=correct responses of the High group
 - b=correct responses of the Low group
 - n-number of respondents in each group

Interpretation

- The most effective questions will have moderate difficulty and high discrimination values.
- Questions having low or negative values of discrimination need to be reviewed very carefully.

http://www.schreyerinstitute.psu.edu/Tools/ItemAnalysis/

Example

Q	n	D.I.	-	bottom 25%	D.F.	Α	В	С	D	Ε
1	105	81	100	59	0.43	0	19	85	1	0
2	93	81	96	63	0.31	14	4	75	0	0
3	93	70	96	54	0.40	13	65	10	1	4
4	93	75	92	67	0.21	1	1	19	2	70

Adapted from "Transforming Multiple Choice Questions to Effectively Assess Application of Knowledge", STReME Series, August 11, 2011 Brenda Roman, MD, Professor of Psychiatry, BSOM Paul Koles, MD, Associate Professor of Pathology and Surgery, BSOM

Reference

- Brady, A. M. (2005). Assessment of learning with multiple-choice questions. Nurse education in practice, 5(4), 238-242.
- McDonald, M. (2002). Writing Critical Thinking Multiple-Choice Items. Systematic assessment of learning outcomes : developing multiple-choice exams (pp. 121-131). Boston : Jones and Bartlett Publishers.
- Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, 35(4), 453-472.
- http://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions
- Bordage G: Elaborated Knowledge: A Key to Successful Diagnostic Thinking. Acad Med 69:883-885, 1994