

## Writing MCQs at different levels

Most MCQs test factual recall of information. MCQs can be developed, however, to test higher order thinking such as application and evaluation of knowledge.

### **Factual Recall of Knowledge**

The candidate repeats previously learned material by recalling facts, terms, and basic concepts.

Keywords: who, what, why, when, where, which, choose, select, how, match.

#### Examples of lead-in questions

- What is the nerve supply....
- What is the blood supply....
- What muscle is innervated by....

#### Example of a MCQ

*A branch of which cranial nerve supplies the vocal cords?*

- A. *Optic*
- B. *Trochlear*
- C. *Abducens*
- D. *Vestibulocochlear*
- E. *Vagus*

### **Application**

The candidate solves problems by applying acquired knowledge, facts, techniques and rules in a clinically relevant situation.

Keywords: apply, choose, make use of, organise, plan, select, solve, utilise, identify.

#### Examples of lead-in questions

- What is the cause of clinical features in this particular case?
- Which of the following is the likely location of the patient's lesion?
- Which of the following is the likely pathogen?
- What is the diagnosis?

## Example of a MCQ

**Stem:** *A 65 year-old man has difficulty rising from a seated position and straightening his trunk, but he has no difficulty flexing his leg.*

**Lead in:** *Which of the following muscles is affected?*

**Options:**

- A. *Gluteus maximus*
- B. *Gluteus minimus*
- C. *Hamstrings*
- D. *Iliopsoas*
- E. *Obturator internus*

## **Evaluation**

The candidate makes judgements about information, the most likely pathology or quality of work based on a set of criteria.

**Keywords:** award, choose, conclude, decide, defend, determine, evaluate, judge, justify, rate, recommend, agree, interpret, prioritise, opinion, support, importance, criteria, assess, influence, value, estimate.

## Examples of lead-in questions

- What is the most appropriate management of this patient?
- Which of the following is most likely to have caused this condition?
- Which of the following is the most likely diagnosis?
- Which of the following is the most appropriate investigation to do next?
- Which of the following is most likely to confirm the diagnosis?
- Which of the following is the most effective management?
- What is the first priority in caring for this patient?

## Example of a MCQ

**Stem:** *A 68 year old lifelong smoker complains initially of gradual onset of progressive hoarseness. He then develops unilateral throat and ear pain and subsequently complains of breathing and swallowing problems.*

**Lead in:** *What is the likeliest cause of his hoarseness?*

**Options:**

- A. *Recurrent laryngeal palsy*
- B. *Laryngeal cancer*
- C. *Hypopharyngeal cancer*
- D. *Post-cricoid cancer*
- E. *Oesophageal cancer*

# Guidelines for writing MCQs (one-from-five) at different levels

## Before writing

- The MCQ must assess the knowledge outcomes of the course or important related concepts, rather than trivial subject matter.
- Identify the cognitive level that the MCQ intends to assess, e.g. factual recall, application or evaluation.
- Think of the **topic** being tested and the **content area**.
  - **Topic:** e.g. anatomy, physiology.
  - **Content area:** This is identified in your curriculum document. For Cardiothoracic surgery, for example, it is heart failure, data interpretation, chest wall and diaphragm, and so on.

The content area and topic are identified in your assessment blueprint.

## Writing the stem or case

- The stem is **not** usually used in factual recall questions. It is needed in questions testing application and evaluation.
- Usually a clinical case commonly encountered in day to day practice will form the basis of a good stem.
- Describe the details of a patient's complaint in simple language.
- Include as much information as possible in the stem, i.e. stems should be long and the options should be short.
- Avoid technical item flaws.
  - The stem having a phrase or term repeated in the option(s).
  - Tricky or unnecessarily complicated stems.
  - Clues to the answer in the stem.
- The stem should be clear, concise and simple.
- Do not include any questions in the stem; this is the next step.

## Writing the lead-in

- The lead-in should clearly indicate how to answer the question.
- Refer back to the topic and content area when formulating the question, e.g. for the topic 'anatomy' and content area 'cardiac trauma', an appropriate lead-in will be - 'what anatomical structure is most likely to be damaged in this case of blunt trauma to the chest wall?'
- Whenever possible try to present a 'task' for the candidate. A **question** is preferable to an open-ended phrase or an incomplete sentence, e.g. 'What is the patient's diagnosis?' is better than 'Regarding the patient's diagnosis.'
- The lead-in, together with the stem/case should give enough information to answer the MCQ **without looking at the options**.
- Avoid technical items flaws, such as:
  - Absolute terms – 'always, never'.
  - Frequency terms – 'often, rarely'.
  - Using options from different categories, e.g. including a treatment option with diagnostic options (i.e. heterogeneous options). Such options are commonly found

in MCQs with the lead-in: 'which of the following statements is correct?'

- Negative questions, e.g. which one is NOT a beta-blocker?  
If this cannot be avoided, negative words should be: highlighted *or* in bold *or* in upper case (capitals).

- The lead-in should be clear, concise and simple.
- Avoid constructing a "test within a test", e.g. 'how many permutations are possible in a bridge hand?' - This question is designed to test elementary statistics. The candidate will be unable to answer it without knowledge of bridge, which is not the intention of the question (source: Designing and managing MCQs).

## Writing the options

- The list of options (usually five) should have only one clearly correct answer. When 'the best' or 'the most likely' answer is sought this should be clearly stated in the lead-in.
- The distractors, though clearly incorrect, should be equally plausible to a weak candidate. When constructing distractors try to think of how an inexperienced trainee would respond to the clinical situation described in the stem (Wood & Cole, 2001).
- All the options should be homogeneous, i.e. belonging to the same category such as, diagnosis, treatment methods, list of nerves, list of muscles. Heterogeneous or internally inconsistent options are poor distractors, e.g. if the four options are about 'investigations' while the fifth is about 'treatment', the testwise candidate will easily exclude the fifth option.
- Options should be short and uncomplicated.
- List the options in a logical order, e.g. if there are numbers the options should be arranged in ascending or descending order. If there is no logical order alphabetical order is preferred.
- If a negative lead-in is unavoidable the options should be the shortest possible, preferably single words.
- Try to ensure that all the options are of the same length.
- The position of the correct answer in the option list should vary among MCQs.
- Use coherent, consistent terminology and inform the candidates of the meaning of the commonly used terms.
  - "*Recognised*" means "an accepted feature of the disease".
  - "*Pathognomonic*" means "a feature specific to the disease, but to no other".
  - "*Characteristic*" means "a feature without which the diagnosis is in question". This term must therefore be used with care.
  - "*Typical*" is synonymous with "*characteristic*".
  - "*The majority*" or "*most*" means over 50%. However, these are vague terms that should be avoided, if possible.
  - *Percentages* as a specific figure are unacceptable, and should be given as a range e.g. 30-40%.
  - *Eponyms* should be defined unless in common use, e.g. Crohn's Disease.

- Avoid technical item flaws.
  - Issues related to testwiseness
  - Grammatical cues - one or more distractors don't follow grammatically from the stem.
  - Logical cues – a sub-set of options is collectively exhaustive. A few options contain all possible answers. A testwise student will recognise these and will consider only these options. The non-testwise student will consider all five options (see Case & Swanson (2001) a fuller explanation).
  - Absolute terms – terms such as 'always/never'.
  - Long correct answer – correct answer is longer, more specific, or more complete than other options.
  - Word repeats – a word or phrase is included in the stem/lead-in and in the correct answer.
  - Convergence strategy – the correct answer includes the most elements in common with the other options, e.g.

In which form are local anaesthetics most effective?

- A. Anionic form, acting from inside the nerve membrane.
- B. Cationic form, acting from inside the nerve membrane.
- C. Cationic form, acting from outside the nerve membrane.
- D. Unchanged form, acting from inside the nerve membrane.
- E. Unchanged form, acting from outside the nerve membrane.

(adapted from Case & Swanson, 2001).

The testwise candidate will exclude 'anionic form' and 'outside the nerve membrane' as the frequency of their appearing as answers are less. Hence, the candidate will have to only decide between options B and D. The reason for this type of flaw is that the examiners write distracters as modifications of the correct answer.

- Unnecessary complications (irrelevant difficulty)
- Numerical data not being stated consistently.
- Vague terms, e.g. frequency and absolute terms (as described above), usually, may, can.
- Overlapping options, e.g. one option being 'analgesics' while another being 'paracetamol'.
- Double options, e.g. do A and B; do X because of Y. The exception may be if all the options have similar double options (which is very unlikely).
- Language in the options is long-winded and difficult to understand, making it difficult and time consuming to sort out the correct option.
- 'None of the above' or 'all the above' as an option.
- Answer to an item is 'hinged' to the answer of a related item, i.e. the candidate can answer the question based on information given in the stem of a previous MCQ.

Examples of the above issues can be found in NBME book MCQs (Case & Swanson, 2001).

## After writing

Subject the MCQ to the five "tests" below (adapted from Case & Swanson, 2001).

1. Does the MCQ address an important concept related to a learning outcome?
2. Does the MCQ assess factual recall of knowledge, application or evaluation?
3. Can the MCQ be answered by *only* reading the stem and lead-in, *without* reading the options?
4. Are all the distractors homogeneous?
5. Is the MCQ devoid of technical item flaws that benefit the testwise candidate or that post irrelevant difficulty?

## Summary

MCQ writing step	Do	Don't
Before writing	<ul style="list-style-type: none"> <li>• MCQs should assess learning outcomes or important concepts</li> <li>• Identify the cognitive level at which the MCQ should be pitched, e.g. factual recall, application of knowledge or evaluation</li> <li>• Decide on the topic and content area</li> </ul>	<ul style="list-style-type: none"> <li>• Do not assess trivial, insignificant facts</li> </ul>
Writing the stem	<ul style="list-style-type: none"> <li>• Should be a <i>common</i> clinical case</li> <li>• Include as much information as required to arrive at the correct answer, i.e. a long stem (with short options)</li> </ul>	<ul style="list-style-type: none"> <li>• Do not synthesise for the candidate, i.e. give details of the patient's complaint in simple language</li> <li>• Avoid technical item flaws, such as               <ul style="list-style-type: none"> <li>❖ A word in the stem repeated in the option(s)</li> <li>❖ Tricky/complicated stems</li> <li>❖ Clues to the answer in the stem</li> </ul> </li> <li>• Do not include any question (task for the candidate) in the stem</li> </ul>
Writing the lead-in	<ul style="list-style-type: none"> <li>• Should clearly indicate how to answer the MCQ</li> <li>• Should preferably be a question</li> <li>• Refer back to the topic &amp; content area, when constructing the lead-in</li> <li>• Try to present a task to the candidate, e.g. what is the diagnosis?</li> </ul>	<ul style="list-style-type: none"> <li>• Use questions and avoid phrases e.g. Regarding epilepsy:</li> <li>• Avoid technical item flaws, such as:               <ul style="list-style-type: none"> <li>❖ Absolute terms, e.g. always, never</li> <li>❖ Frequency terms, e.g. rarely</li> <li>❖ 'Which of the following statements is correct?' This type of lead-in may lead to heterogeneous options</li> <li>❖ Negative questions</li> </ul> </li> </ul>
Checking the stem and lead-in	<ul style="list-style-type: none"> <li>• Lead-in and stem must give enough information to answer the MCQ, <i>without/before</i> reading the options</li> <li>• Both should be clear, precise and simple</li> </ul>	<ul style="list-style-type: none"> <li>• Do not create a 'test within a test'.</li> </ul>
Writing the options	<ul style="list-style-type: none"> <li>• Should have <i>only</i> one clear answer</li> <li>• Distractors should be clearly incorrect, but plausible</li> <li>• Should be short and uncomplicated</li> <li>• All options should be homogeneous, i.e. like needs to be compared with like, e.g. all options being clinical signs</li> <li>• List in a logical order</li> <li>• The positions of the correct option should vary with other MCQs</li> <li>• All options are of similar length</li> <li>• Use coherent, consistent terminology, e.g. pathognomonic, typical, or recognised feature</li> </ul>	<ul style="list-style-type: none"> <li>• Avoid technical item flaws, such as:               <ul style="list-style-type: none"> <li>❖ Related to testwiseness                   <ul style="list-style-type: none"> <li>- Grammatical cues</li> <li>- Logical cues</li> <li>- Absolute terms</li> <li>- Long correct answer</li> <li>- Word repeats</li> <li>- Convergence strategy</li> </ul> </li> <li>❖ Related to irrelevant difficulty                   <ul style="list-style-type: none"> <li>- Inconsistent numerical data</li> <li>- Vague terms, e.g. may</li> <li>- Overlapping questions</li> <li>- Double options, e.g. do X and Y</li> <li>- Language not parallel to others</li> <li>- 'None of the above/all of the above'</li> <li>- Answer is 'hinged' to another MCQ</li> </ul> </li> </ul> </li> </ul>
After writing	<ol style="list-style-type: none"> <li>1. Does the MCQ assess an important concept?</li> <li>2. Does the MCQ test factual recall of knowledge, application or evaluation?</li> <li>3. Can the MCQ be answered by <i>only</i> reading the stem &amp; lead-in?</li> <li>4. Are all the options homogeneous?</li> <li>5. Is the MCQ (stem, lead-in and options) devoid of technical item flaws?</li> </ol>	

## References

Bloom, B.S. (Ed). (1956). Taxonomy of Educational Objectives Handbook 1: Cognitive Domain. Longman, Green & Co., New York.

Case, S.M. & Swanson, D.B. (2001). Constructing written test questions for basic and clinical sciences. 3<sup>rd</sup> Edn. National Board of Medical Examiners (NBME), Philadelphia, USA.  
<http://www.nbme.org/PDF/2001iwg.pdf> (retrieved on 5/5/2004).

Designing MCQs – do's and don'ts. Appendix B, in: Designing and managing MCQs.  
<http://web.uct.ac.za/projects/cbe/mcqman/mcqcont.html> (retrieved on 5/5/2004).

The Society of Radiologists in Training. Multiple Choice Questions: advice on MCQ examination technique. <http://www.thesrt.org.uk/forum/messages/444/157.html?0> (retrieved on 5/5/2004).

Universities medical assessment partnership. MCQ Writing guide, in: Write an examination question. <http://www.umap.man.ac.uk/Documents/MCQWritingguide.pdf> (retrieved on 5/5/2004).

Wood, T. & Cole, G. (2001). Developing multiple choice questions for the RCPSC certification examinations. The Royal College of Physicians and Surgeons of Canada.  
[http://cls.uibk.ac.at/personalentwicklung/RCPSC\\_multiplechoice\\_e.pdf](http://cls.uibk.ac.at/personalentwicklung/RCPSC_multiplechoice_e.pdf) (retrieved on 5/5/2004).

## References

### General

Case, S.M. (1994). The use of imprecise terms in examination questions: how frequent is frequently? *Academic medicine*, 69 (10), pp. S4-S<sup>^</sup>.

Downing, S.M. (2002). Threats to the validity of locally developed multiple-choice tests in medical education: construct-irrelevant variance and construct under representation. *Advances in Health Sciences Education*, 7, pp. 235-241.

Schuwirth, L.W.T. & Van der Vleuten, C.P.M. (2003). Written assessment. *British Medical Journal*, 326, pp. 643-645.

### MCQs (one-from-five format)

Anderson, J. (2004). Multiple-choice questions revisited. *Medical Teacher*, 26 (2), pp. 110-113.

Case, S.M. & Swanson, D.B. (2001). Writing one-best-answer questions for the basic and clinical sciences, Sec. II, in: *Constructing written test questions for the basic and clinical sciences*. 3<sup>rd</sup> edn. National Board of Medical Examiners (NBME), Philadelphia, USA. Pp. 31-66.  
<http://www.nbme.org/PDF/2001iwg.pdf> (retrieved on 5/5/2004).

Shakun, E.N., Maguire, T.O. & Cook, D.A. (1994). Strategy choices in multiple-choice items. *Academic Medicine*, 69 (10), pp. S7-S9.

Swanson, D.B. & Case S.M. (1997). Assessment in basic science instruction: direction for practice and research. *Advances in Health Science Education: Theory and Practice*, 2, pp. 71-84.

Wood, T. & Cole, G. (2001). Developing multiple choice questions for the RCPSC certification examinations. The Royal College of Physicians and Surgeons of Canada.  
[http://cls.uibk.ac.at/personalentwicklung/RCPSC\\_multiplechoice\\_e.pdf](http://cls.uibk.ac.at/personalentwicklung/RCPSC_multiplechoice_e.pdf) (retrieved on 5/5/2004).

### Assessment blueprinting

Boursicot, K. & Roberts, T. (retrieved on 5/5/2004). Blueprinting. ITSN. [http://www.itsn-01.ac.uk/docs/workshop/osce\\_blueprint.pdf](http://www.itsn-01.ac.uk/docs/workshop/osce_blueprint.pdf)

Crossley, J., Humphris, G. & Jolly, B. (2002). Assessing health professionals. *Medical Education*, 36, pp. 800-804.

Fielding, D.W. et al. (1992). Assuring continuing competence: identification and validation of a practice-based assessment blueprint. *American Journal of Pharmaceutical Education*, 56(1), pp. 21-29.

Fowell, S.L., Southgate, L.J. & Bligh, J.G. (1999). Evaluating assessment: the missing link? *Medical Education*, 33 (4), pp. 276-281.

Fuller, J. (2003). Improving the validity of assessment: blueprinting assessment.  
[http://www.clinicalteaching.nhs.uk/site/DOCS/UTL%203%20\(toolbox\)%20assessment%20blueprint.pdf](http://www.clinicalteaching.nhs.uk/site/DOCS/UTL%203%20(toolbox)%20assessment%20blueprint.pdf) (retrieved on 5/5/2004).

Newble, D., Dawson, B., Dauphinee, D., Gordon, P., Macdonald, M., Swanson, D., Mulholland, H., Thomson, A. & Van der Vleuten, C. (1994). Guidelines for assessing clinical competence. *Teaching and Learning Medicine*, 6 (3), pp. 213-220.

Swanson, D.B. (1987). A measurement framework for performance-based assessment. In: Hart, I.R. & Harden, R.M., Eds. *Further developments in assessing clinical competence*. Can-Heal, Montreal.

Tombleson, P., Fix, R.A. & Dacre, J.A. (2000). Defining the content for the objective structured clinical examination component of the Professional and Linguistic Assessments Board examination: development of a blueprint. *Medical Education*, 34 (7), pp. 566-572.

Wass, V., Van der Vleuten, C., Shatzer, J. & Jones, R. (2001). Assessment of clinical competence. *The Lancet*, 357, pp. 945-949.

# PACKRAT Test Writing Workshop

PAEA Annual Education Forum  
October 25, 2007



## PACKRAT Members



- Jim Van Rhee, MS, PA-C
- Linda Allison, MD, MPH
- Mark Archambault, MHS, PA-C
- Petar Breitinger, MPAS, PA-C
- Christine Bruce, MHSA, PA-C
- Kim Cavanagh, MPAS, PA-C
- Ralph Rice, MPAS, PA-C
- Eric Vangsnes, PhD, PA-C
- Donna Yeisley, MEd, PA-C

## Objectives



- Develop a basic format for test item development
- Identify basic technical flaws in test item creation

## Why do we test?



- Communicate to students what material is important
- Motivate students to study
- Identify areas of deficiency that need remediation or further learning
- Determine final grades or make promotion decisions
- Identify areas where course/curriculum is weak

## Why are we here?



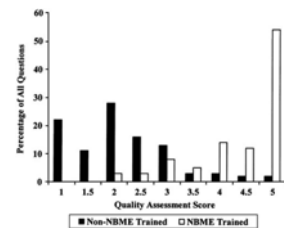
- Improve test writer skills
- Improve quality of examinations

It all starts with the course or lecture objectives.

## Improve Exam Quality



- Jozefowicz et al (2002)
  - Analyzed 555 MCQ
  - Mean scores (Scale 1-5)
    - NBME trained:  $4.28 \pm 0.85$
    - No NBME training:  $2.03 \pm 0.90$
- Experienced test writers prepare better exams



## Student Effect

- Downing (2005)
- Reviewed four exams
  - 36-65% of questions flawed on the 4 exams
  - Students scored lower on flawed items
- As high as 10-15% of students are incorrectly classified as failed when they may have passed had items not been flawed.

Table IV. Pass-fail agreement analysis, all examinations, all students  $N = 749$

	Flawed items		Total
	Fail	Pass	
Standard items			
Fail	73	30	103
Pass	102	544	646
Total	175	574	749

## The NBME MCQ Format (A-Type)

A Multiple Choice Question (MCQ) has the following components:

Stem – the information for the question

Lead-in – the question and how to answer

Answer options:

1. Key – the correct answer
2. Distracters -- the incorrect answers (3-4)

## MCQ- One Best Answer

**Stem:**

A 32-year-old man has a 4-day history of progressive weakness in his extremities. He has been healthy except for an upper respiratory tract infection 10 days ago. His temperature is 37.8 C (100 F), blood pressure is 130/80 mm Hg, pulse is 94/min, and respirations are 42/min and shallow. He has symmetric weakness of both sides of the face and the proximal and distal muscles of the extremities. Sensation is intact. No deep tendon reflexes can be elicited, the plantar responses are flexor.

**Lead-in:**

Which of the following is the most likely diagnosis?

**Options:**

- A. Acute disseminated encephalomyelitis
- B. Guillain-Barré syndrome
- C. Myasthenia gravis
- D. Poliomyelitis
- E. Polyomyositis

## AOTA and NOTA

- All of the Above (AOTA)
  - Avoid this option
  - If present most students will pick
- None of the Above (NOTA)
  - Use carefully
    - Use only with items that can logically include only one correct option
    - Can be used if few incorrect responses are available
    - May improve item discrimination and difficulty
  - Should be avoided by novice item writers

## Negative Questions

- Some researchers found no difference in item difficulty when used negatively worded stems
- Tamir (1993) found at a higher cognitive level negatively worded stems to be more difficult
  - Others have found the opposite
- Use with caution
  - Measure a relevant objective (What to avoid or what is not true)
  - If negative, use CAPITAL LETTERS and **boldface**

## Question Format

- Follow a standard format for stem
  - Identifying data, History, Physical examination, Labs, Lead in question
    - End in question format or sentence-completion format
    - Question format preferred due to directness in getting to central idea of the item
- Distractors
  - Three options are typically sufficient
  - More effective the distractors the more discriminating the item

## Classic Format

Patient vignettes should include some or all of the following components in the order indicated:

- Age, Gender (eg. A 45-year-old man)
- Site of Care (eg. comes to the emergency department)
- Presenting Complaint (eg. because of a headache)
- Duration (eg. that has continued for 2 days)
- Patient History (with Family History ?)
- Physical Findings
- +/- Results of Diagnostic Studies
- +/- Initial Treatment, Subsequent Findings, etc.

## TECHNICAL FLAWS- TESTWISENESS

## Grammatical Cues

- One or more distractors does not follow grammatically from the stem

A 50-year-old male presents with chest pain which radiates to his left arm. He also notes diaphoresis and nausea. Vitals are stable and physical examination is unremarkable. The next best step in the evaluation of this patient is an

- electrocardiogram.
- cardiac catheterization.
- cardiac enzymes.
- admit for observation.

Only one works grammatically.

## Logical Cues

- A subset of objects are collectively exhaustive

An arterial blood gas obtained on a COPD patient reveals the following data.  
pH = 7.32  
PaCO<sub>2</sub> = 60 torr.  
PaO<sub>2</sub> = 33 torr.  
What is the most appropriate interpretation of the oxygenation status?

- hypoxic
- hyperoxic
- normal
- hypercarbic

A and B are opposites, so one or the other is correct.  
Or both are wrong.

## Absolute Terms

- Terms such as "always" or "never" are used in options
- Use verbs in the lead-in, not the options

Which of the following is true of Crohns disease?  
a)Never involves the rectum  
b)Cobblestoning and string sign are common  
c)Always have an elevated sed rate  
d)Treatment may consist of steroids and 5-ASA

Two possible correct answers.

## Undirected Stem

- Insufficient information and lack of direction.
- Should be able to answer without looking at choices.

Insulin secretagogues  
a)Desensitize beta cells to glucose  
b)Include sulfonylureas and metformin  
c)Open the potassium channels in the beta cells, thereby decreasing calcium ions  
d)Sensitize beta cells to stimulate insulin  
e)Will not be effective in response to a glucose challenge

## Word Repeats

- Word or phrase is included in stem and correct answer

Which of the following is a genetic cause of hemolytic anemia?

- a) Hereditary spherocytosis
- b) Hypersplenism
- c) Cold agglutinins
- d) Malaria

**Note genetic and hereditary.**

## Convergence Strategy

- Correct answer includes the most elements in common with the other options

*Local anesthetics are most effective in the*

- A. *anionic form, acting from inside the nerve membrane*
- B. *cationic form, acting from inside the nerve membrane*
- C. *cationic form, acting from outside the nerve membrane*
- D. *uncharged form, acting from inside the nerve membrane*
- E. *uncharged form, acting from outside the nerve membrane*

Exclude anionic and outside the nerve membrane as the frequency of their appearing as answers is less.

## Technical Flaws- Irrelevant Difficulty

- Options are long, complicated, or double
- Numeric data is not stated consistently
- Frequency terms in the options are vague (usually)
- Stems are tricky or unnecessarily complicated

## 5-Step Review: Does the question...

- focus on one important concept?
- assess application of knowledge, not isolated facts?
- pose a clear question, can you answer it with the options covered?
- have all distractors homogeneous?
- contain technical flaws that aid test taker?

### General Guidelines for Item Construction

- Make sure the item can be answered without looking at the options OR that the options are 100% true or false.
- Include as much of the item as possible in the stem; the stems should be long and the options short.
- Avoid superfluous information.
- Avoid "tricky" and overly complex items.
- Write options that are grammatically consistent and logically compatible with the stem; list them in logical or alphabetical order. Write distractors that are plausible and the same relative length as the answer.
- Avoid using absolutes such as *always*, *never*, and *all* in the options; also avoid using vague terms such as *usually* and *frequently*.
- Avoid negatively phrased items (eg. those with *except* or *not* in the lead-in). If you must use a negative stem, use only short (preferably single word) options.

And most important of all:

**Focus on important concepts;  
don't waste time testing trivial facts.**

### Group Work Session (90 minutes):

Get in groups of 6-8 per table  
One PACKRAT member per table  
Work on questions, obtain feedback from group and group leader

- Review for cues just discussed

Submit questions to group leader  
Questions will be collected and e-mailed to all participants

### Debriefing (10 minutes)



## Debriefing

- Write, review, rewrite
  - Analyze
  - Review and rewrite
- 
- Form test writing groups to aid in reviews



## References

- NBME. (2002). *Constructing Written Test Questions for Basic and Clinical Sciences*, 3<sup>rd</sup> edition.
- Downing, S.M. (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*, 10:133-143.
- 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):309-334.
- Jozefowicz, R.F., Koeppen, B.M., Case, S., et al. (2002). The quality of in-house medical school examinations. *Academic Medicine*, 77(2):156-161.

