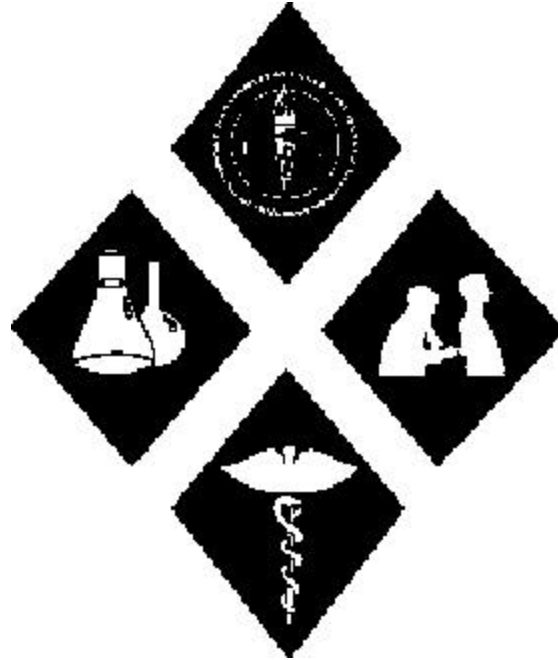

Constructing Written Test Questions For the Basic and Clinical Sciences

Third Edition



Susan M. Case

David B. Swanson

National Board of Medical Examiners

3750 Market Street

Philadelphia, PA 19104

Chapter 6

Extended-Matching (R-Type) Items



Extended Matching items are multiple-choice items organized into sets that use one list of options for all items in the set. A well-constructed Extended-Matching set includes four components:

1. a theme;
2. an option list;
3. a lead-in statement; and
4. at least two item stems, as illustrated below.

Theme: **Fatigue**

Options:

A. Acute leukemia	H. Hereditary spherocytosis
B. Anemia of chronic disease	I. Hypothyroidism
C. Congestive heart failure	J. Iron deficiency
D. Depression	K. Lyme disease
E. Epstein-Barr virus infection	L. Microangiopathic hemolytic anemia
F. Folate deficiency	M. Miliary tuberculosis
G. Glucose 6-phosphate dehydrogenase deficiency	N. Vitamin B ₁₂ (cyanocobalamin) deficiency

Lead-in: For each patient with fatigue, select the most likely diagnosis.

Stems:

1. A 19-year-old woman has had fatigue, fever, and sore throat for the past week. She has a temperature of 38.3 C (101 F), cervical lymphadenopathy, and splenomegaly. Initial laboratory studies show a leukocyte count of 5000/mm³ (80% lymphocytes, with many lymphocytes exhibiting atypical features). Serum aspartate aminotransferase (AST, GOT) activity is 200 U/L. Serum bilirubin concentration and serum alkaline phosphatase activity are within normal limits.

Ans: E

2. A 15-year-old girl has a two-week history of fatigue and back pain. She has widespread bruising, pallor, and tenderness over the vertebrae and both femurs. Complete blood count shows hemoglobin concentration of 7.0 g/dL, leukocyte count of 2000/mm³, and platelet count of 15,000/mm³.

Ans: A

Extended-matching items are written differently than traditional one-best-answer items. Most often, the theme, lead-in, and options are written first; the item stems are written last. For example, if you want to write some questions related to the diagnosis of fatigue, you would begin by listing the diagnoses that might cause fatigue. You would then write a vignette for each (or many) of the options in the list. The example above includes vignettes for Epstein-Barr virus infection and for acute leukemia. Additional items might be written for some of the remaining diagnoses; for common, treatable diagnoses, more than one item might be prepared. The sample vignettes are moderate in length; shorter, more-focused vignettes could also be used. Alternatively, examinees could be challenged to identify key diagnostic information, intermingled with incidental findings, by using longer vignettes.

Avoiding Flaws When You Write Extended-Matching Items for Your Own Examination

The four components (theme, options, lead-in, and stems) are all essential for the construction of a good quality extended-matching set. Sets without lead-ins (or with nonspecific lead-ins, such as “Match each item with the best option”) should NOT be used, because they generally pose inconsistent or ambiguous tasks for examinees. The following set is flawed. The options are heterogeneous; there is no lead-in; the stems cannot be answered without reading the options. Rules for extended-matching items are completely analogous to those for one-best answer items.

Sample Extended-Matching Set - Flawed

- | | |
|---|--|
| A. is motion sickness | I. are completely controlled |
| B. have no effects on people | J. cause plant and eye damage |
| C. indirectly increase CO ₂ | K. are negligible |
| D. cause death | L. increase risk of skin cancer |
| E. increased odor sensitivity | M. cannot be controlled |
| F. is a reduction in visibility | N. excess acute respiratory illness among children |
| G. esthetics, economics, health | O. contrary to public opinion |
| H. products of fossil fuel combustion | |
| 1. Factors that people consider when evaluating air quality | |
| 2. A principal effect of particulate matter in air | |
| 3. The products of photochemical smog | |

After reading the stem in Item #1, examinees have only the vaguest idea what the question is about. In an attempt to determine the “best” answer, the examinees have to decide whether “is motion sickness” is more or less true than “have no effects on people.” The task is not do-able. Under these circumstances, unless an option is absolutely 100% true or false, it cannot be rank-ordered with the other options. The stem of item #1 by itself is not clear; the item cannot be answered without looking at the options.

As with one-best-answer items, the stems should be long; the options should be short. There **MUST** be a lead-in that establishes the relationship between the items and the options. There should be **NO** verbs in the options. The “cover-the-options” rule is as relevant for extended-matching items as it is for one-best-answer items.

Sample Lead-ins and Topics for Option Lists

Patient vignettes provide an excellent structure for stems, not only in the clinical sciences, but also to assess knowledge in the basic sciences. Lead-ins generally begin with a phrase such as “For each of the following patients.” Often sets are organized around chief complaints or some other factor that allows a more specific introductory phrase such as “For each of the following patients with fatigue,” or “For each of the following patients with an enzyme deficiency.” The second part of the lead-in describes the task and the option set: “select the most likely diagnosis”; “select the protein that is most likely to be defective.”

The following are some additional sample lead-ins and some suggested topics for option lists.

- For each of the following patients, select the [eg, nerve] that is most likely to be [abnormal/defective/deficient/non-functioning].
Options sets could include list of nerves; list of muscles; list of enzymes; list of hormones; list of proteins; list of types of cells; list of neurotransmitters; list of pathologic processes.
- For each of the following patients, select the [finding] that would be expected.
Options sets could include list of laboratory results; list of additional physical signs; autopsy results; results of microscopic examination of fluids, muscle or joint tissue; DNA analysis results; hormone levels.
- For each of the following patients, select the most likely [cause].
Options sets could include list of underlying mechanisms of the disease; medications that might cause side effects; list of drugs or drug classes; toxic agents; hemodynamic mechanisms.
- For each of the following patients, select the [eg, drug] that should be administered.
Options sets could include list of drugs, vitamins, amino acids, enzymes, hormones.
- For each of the following patients with [chief complaint], select the most likely diagnosis.
Options sets could include list of diagnoses, most often organized around a chief complaint such as diseases that cause chest pain or diseases that cause fever.
- For each of the following patients, select the most appropriate next step in patient care.
Options sets could include list of pharmacologic therapies, list of laboratory studies, disposition alternatives, or the options could include a mixed set of treatments and additional studies to assess whether the student knows when sufficient data have been gathered.

More on Options for R-Sets

Generally, anything that can be listed can form the basis for options in an R-set. Below are some topics that have been used as the basis for option lists.

Arteries	Connective Tissue Types
Nerves	Anatomic Structures
Muscles	Endocrine Structures
Amino Acids	Neurotransmitters
Peptides	Metabolic Defects
Hormones	Immune Disorders
Enzymes	Motor System Components
Cell Components	Cardiac Structures
Cell Types	Organelles
Blood Components	Congenital Anomalies
Molecules	Segments of the Spinal Cord
Karyotypes	Central Nervous System Components
Proteins	Secretory Products
Lipids	Extracellular Matrix Components
Pathogens/Bacteria/Fungi	Management Alternatives
Viruses	Drugs/Drug Classes
Cytokines	Pathologic Processes
Toxins	Pathophysiologic States
Vitamins/Minerals	Electrolyte Abnormalities
Diagnoses	Diagnostic Tests

The list of options should be single words or very short phrases. They must be homogeneous (all diagnoses, all management options, all anatomical sites, all vitamins, etc). They can be labeled areas in a graph or in pictorial material. Options, especially those involving laboratory values, are often expressed in tabular form (see physiology example). Include all relevant options that are appropriate for the examinees; subtle distinctions and uncommon diagnoses may be inappropriate. For some topics, as few as three options might be appropriate; for others, a list of 26 (one for each letter in the alphabet) might be required.

Writing the Item Stems

Patient vignettes provide an excellent structure for stems both in the basic and the clinical sciences. In the clinical sciences, the vignette commonly provides the patient's age, gender, chief complaint, and site of care, followed by personal history, family history (if relevant), then physical examination information, then laboratory data (if provided). Depending upon the purpose of the set, vignettes can be brief, prototypic presentations or fuller descriptions that challenge examinees to identify key information. Generally, these items would include at least the patient's age, gender, chief complaint, and related history. Items to assess knowledge of the basic sciences, particularly for courses taught in the first year of medical school, might include less detailed prototypical descriptions.

Each patient description should be similar in structure to the others in the set. For example, if race, ethnicity, or occupation is included in one item, it should be included in all items; if laboratory data are included for one item, include them in all items. It is advisable not to mix adults and pediatric cases in the same set — too often the age alone provides too much cueing and eliminates large numbers of options from consideration.

One advantage of the use of patient vignettes is that it helps to assure that the content assesses application of knowledge. These items should not resemble crossword puzzles, where both the options and the stems are single words or short phrases. Avoid reconstructing those items you were faced with in junior high school where you had to draw a line from something in column A to the matching option in column B.

It is particularly important that the items be straightforward. There is no reason to make them tricky; the extended option list makes them difficult enough to allow you to distinguish the knowledgeable student from the other students without resorting to trickery. As with well-constructed A-type questions, the “cover-the-options” rule is paramount. Knowledgeable students should be able to generate an answer to the question and then find that answer in the alphabetical list of options.

An item should be prepared for most of the options; for common or important options, more than one item can be written. In constructing an examination assessing general competence, to avoid overemphasizing a topic, all of the options, but only two or three of the items, would be used; the remaining items are retained for subsequent exams. On the other hand, if you want to assess knowledge in greater depth on a smaller number of topics, 10 to 20 items can be included for each set, with a sub-score calculated for each topic.

In reviewing the items, check to make sure that there is only a single “best” answer for each question. Also make sure that there are at least four reasonable distractors for each item. As a final check, it is recommended that you ask a colleague to review the items (without the correct answer indicated). If the colleague has difficulty determining the correct answer, modify the option list or the item to eliminate the ambiguity.

Sample Good and Bad Item Stems Using the Same Option List

The following is a good microbiology set. The options are a homogeneous list of pathogens; including both viruses and bacteria makes sense. There is a lead-in that presents a clear task for the examinee. There are two item stems that require students to apply their basic science knowledge of microbiology to arrive at the most likely cause of each patient's illness.

- | | |
|---------------------------------------|--------------------------------------|
| A. Adenovirus | L. <i>Haemophilus influenzae</i> |
| B. <i>Aspergillus fumigatus</i> | M. <i>Histoplasma capsulatum</i> |
| C. <i>Bacillus anthracis</i> | N. <i>Mycobacterium tuberculosis</i> |
| D. <i>Candida albicans</i> | O. <i>Mycoplasma pneumoniae</i> |
| E. <i>Chlamydia psittaci</i> | P. <i>Neisseria gonorrhoeae</i> |
| F. <i>Coccidioides immitis</i> | Q. <i>Neisseria meningitidis</i> |
| G. Coronavirus | R. <i>Pneumocystis carinii</i> |
| H. <i>Corynebacterium diphtheriae</i> | S. Rhinovirus |
| I. <i>Coxiella burnetii</i> | T. <i>Streptococcus pneumoniae</i> |
| J. Coxsackievirus | U. <i>Streptococcus pyogenes</i> |
| K. Epstein-Barr virus | (group A) |

For each patient with fever, select the pathogen most likely to have caused his/her illness.

1. A 7-year-old girl has a high fever and a sore throat. There is pharyngeal redness, a swollen right tonsil with creamy exudate, and painful right submandibular lymphadenopathy. Throat culture on blood agar yields numerous small β -hemolytic colonies that are inhibited by bacitracin.

Ans: U

2. For the past week, an 18-year-old man has had fever, sore throat, and malaise with bilaterally enlarged tonsils, tonsillar exudate, diffuse cervical lymphadenopathy, and splenomegaly. There is lymphocytosis with atypical lymphocytes. The patient tests positive for heterophil antibodies.

Ans: K

The following stem, developed for the same set, assesses recall of isolated facts rather than application of knowledge. It looks more like a crossword puzzle question than a question for a medical school examination.

3. An encapsulated gram-positive organism that usually grows in pairs or short chains.

Ans: T

The following set includes two item stems. The first requires that the examinee synthesize information in order to determine a diagnosis; the second requires only recall of an isolated fact.

- | | |
|---------------------------|--------------|
| A. Vitamin A | I. Biotin |
| B. Vitamin B ₁ | J. Copper |
| C. Vitamin B ₂ | K. Folate |
| D. Vitamin B ₆ | L. Iodine |
| E. Vitamin C | M. Iron |
| F. Vitamin D | N. Magnesium |
| G. Vitamin E | O. Niacin |
| H. Vitamin K | P. Zinc |

For each patient with clinical features caused by metabolic abnormalities, select the vitamin or mineral that is most likely to be involved.

1. A 70-year-old widower has ecchymoses, perifollicular petechiae, and swelling of the gingiva. His diet consists mostly of cola and hot dogs.

Ans: E

2. Involved in clotting factor synthesis.

Ans: H

The following set includes two item stems. The first item stem requires that the examinee synthesize information to determine the diagnosis of trisomy 21; the second item provides this information. When you develop item stems, you need to decide the level of synthesis you will provide.

- | | |
|---|------------------------------|
| A. Atrial septal defect | E. Patent ductus arteriosus |
| B. Coarctation of the aorta | F. Pulmonic valve stenosis |
| C. Complete transposition of the great arteries | G. Tetralogy of Fallot |
| D. Endocardial cushion defect | H. Ventricular septal defect |

For each patient, select the most likely congenital heart defect.

1. A 3090-g infant is born at term. Physical examination shows the child to be hypotonic with a weak suck. There is no cyanosis. Prominent epicanthal folds, a large tongue, and small incurved 5th digits on the hands are present. There is a loud holosystolic murmur over the entire precordium and a palpable thrill at the upper left sternal border in the 2nd through 4th intercostal spaces.

Ans: D

2. An infant with trisomy 21 has clefts of the right and left atrioventricular valves; there is no cyanosis.

Ans: D

This set poses a task that is clinically backward. The examinee is given a vaccine and asked to select the best patient for this vaccine. A more appropriate task would be for the examinee to be given a patient and asked for the most appropriate next step in patient care (ie, describing a patient in each item stem; using “For each patient, select the most appropriate next step in patient care”; and using things such as vaccines as options). A second problem with the set is that insufficient information is provided about each patient. For example, the examinee would want the immunization history for a patient before deciding which immunizations to provide.

	Birth Year	Gender	Occupation	Pregnant	Childhood Disease	Medical History	Allergy
A.	1980	Male	Student	-	None	None	Egg products
B.	1975	Female	Painter	Yes	None	None	Penicillin
C.	1970	Female	Teacher	Yes	None	None	None
D.	1965	Female	Lawyer	No	None	None	Gelatin products
E.	1960	Male	Painter	-	Measles	None	Tetanus toxoid
F.	1955	Female	Clerk	No	Mumps	Diabetes	None
G.	1950	Female	Nurse	No	Varicella	None	None
H.	1945	Male	Executive	-	Measles	Hypertension	None
I.	1940	Male	Driver	-	Rubella	Splenectomy	None
J.	1935	Female	Homemaker	No	Unknown	Gastritis	Sulfonamides

For each vaccination, select the patient profile that represents its most appropriate use.

1. Measles vaccine
2. Meningococcal vaccine

Overview of the Steps for Writing Extended-Matching Items

1. **Identify the theme for the set.** The theme can be a chief complaint (eg, chest pain, fatigue), a disposition situation (eg, admission/discharge from the emergency department), a drug class (eg, antihypertensive agents, antibiotics).
2. **Write the lead-in for the set** (eg, *For each patient described below, select the most likely diagnosis*). The lead-in indicates the relationship between the stems and options, clarifying the question posed for examinees. It is an essential component of an Extended-Matching set.
3. **Prepare the list of options.** The list of options should be single words or very short phrases. List the options in alphabetical order unless there is a logical order.
4. **Write the items.** The items within a set should be similar in structure. Most often, patient vignettes are appropriate.
5. **Review the items.** Check to make sure that there is only a single “best” answer for each question. Also make sure that there are at least four reasonable distractors for each item. As a final check, it is recommended that you ask a colleague to review the items (without the correct answer indicated). If the colleague has difficulty determining the correct answer, modify the option list or the item to eliminate the ambiguity.

Additional information on writing Extended-Matching items can be found in:

Case SM, Swanson DB. Extended-matching items: a practical alternative to free-response questions. *Teaching and Learning in Medicine*. 1993;5(2):107-115.

Case SM, Swanson DB, Woolliscroft JO. Assessment of diagnostic pattern recognition skills in medicine clerkships using a written test. In: Harden R, Hart I, Mulholland H, eds. *Approaches to Assessment of Clinical Competence*. Norwich, England: Page Brothers; 1992:452-458.

Sample Extended-Matching Sets

Sample Anatomy Set

- | | |
|-----------------------------------|------------------------------------|
| A. Left anterior cerebral artery | E. Left posterior cerebral artery |
| B. Right anterior cerebral artery | F. Right posterior cerebral artery |
| C. Left middle cerebral artery | G. Left lenticulostriate arteries |
| D. Right middle cerebral artery | H. Right lenticulostriate arteries |

For each patient with neurologic abnormalities, select the artery that is most likely to be involved.

1. A 72-year-old right-handed man has weakness and hyperreflexia of the right lower limb, an extensor plantar response on the right, normal strength of the right arm, and normal facial movements. **Ans: A**
2. A 68-year-old right-handed man has right spastic hemiparesis, an extensor plantar response on the right, and paralysis of the lower two-thirds of his face on the right. His speech is fluent, and he has normal comprehension of verbal and written commands. **Ans: G**

Sample Pharmacology Set

- | | |
|-------------------------|-------------------|
| A. Acetaminophen | J. Nalidixic acid |
| B. Amiodarone | K. Nitrofurantoin |
| C. ACE inhibitors | L. Penicillin |
| D. Aspirin | M. Prednisone |
| E. Atenolol | N. Procainamide |
| F. Bleomycin | O. Propranolol |
| G. Cytosine arabinoside | P. Sulfasalazine |
| H. Furosemide | Q. Tetracycline |
| I. Metronidazole | R. Verapamil |

For each patient, select the drug most likely to have caused the adverse effect.

1. A 56-year-old man with recurrent ventricular arrhythmias began taking an antiarrhythmic drug 5 months ago. He now has progressive dyspnea, cough, and low-grade fever. Erythrocyte sedimentation rate is increased. X-ray film of the chest shows a diffuse interstitial pneumonia. Pulmonary function tests show that diffusing capacity for carbon monoxide is decreased. **Ans: B**
2. A 62-year-old man with chronic obstructive pulmonary disease begins therapy with an antihypertensive drug. Two weeks later, he has marked worsening of dyspnea and clearly audible wheezing. **Ans: O**

Sample Physiology Set — Clinical Features (Which additional findings are likely?)

	<u>pH</u>	<u>PO₂ mm Hg</u>	<u>PCO₂ mm Hg</u>	<u>HCO₃ - mEq/L</u>
A.	7.15	98	33	11
B.	7.15	98	24	8
C.	7.30	56	80	38
D.	7.40	100	40	25
E.	7.50	100	33	25
F.	7.50	100	24	18
G.	7.50	56	33	25

For each patient described below, select the most likely arterial blood gas findings.

1. A 22-year-old man with a 3-week history of polyuria and polydipsia has had nausea, vomiting, and decreased responsiveness for the past 12 hours. Urinalysis (dipstick) shows 4+ glucose and 4+ ketones. **Ans: B**

2. A 25-year-old woman is brought to the emergency department 12 hours after a suicide attempt. She took approximately 100 500-mg aspirin tablets. **Ans: F**

Sample Diagnosis Set

- | | |
|----------------------------------|------------------------------|
| A. Ankylosing spondylitis | E. Osteoporosis |
| B. Intervertebral disc infection | F. Spinal stenosis |
| C. Multiple myeloma | G. Spondylolysis |
| D. Myofascial pain | H. Tuberculosis of the spine |

For each patient with back pain, select the most likely diagnosis.

1. A 26-year-old man has insidious onset of low back pain and early morning stiffness. The pain alternates from side to side and occasionally radiates into the buttocks and back of the thighs, but not below the knees. The patient has acute anterior uveitis, diffuse low back and sacroiliac tenderness, and restricted range of motion at the hips. His erythrocyte sedimentation rate is 40 mm/h; latex fixation test is negative; and mild hypoproliferative anemia is present.

Ans: A

2. Twelve hours after being struck from the rear while driving her car, a 28-year-old woman has vague cervical and lumbar pain associated with headache and restricted cervical mobility. She is now very anxious. Rope-like bands of muscle are present in the lumbar area and over the left buttock; the bands are painful

Ans: D

Additional items would cover some of the remaining diagnoses. The sample vignettes are of average length; shorter, more focused vignettes may also be used. Alternatively, examinees could be challenged to identify key diagnostic information by using longer vignettes.

Sample Diagnosis Set

- | | |
|---------------------------------|---------------------------------|
| A. Abdominal aneurysm | K. Kidney stone |
| B. Appendicitis | L. Mesenteric adenitis |
| C. Bowel obstruction | M. Mesenteric artery thrombosis |
| D. Cholecystitis | N. Ovarian cyst — ruptured |
| E. Colon cancer | O. Pancreatitis |
| F. Constipation | P. Pelvic inflammatory disease |
| G. Diverticulitis | Q. Peptic ulcer disease |
| H. Ectopic pregnancy — ruptured | R. Perforated peptic ulcer |
| I. Endometriosis | S. Pyelonephritis |
| J. Hernia | T. Torsion |

For each patient with abdominal pain, select the most likely diagnosis.

1. A 25-year-old woman has sudden onset of persistent right lower abdominal pain that is increasing in severity. She has nausea without vomiting. She had a normal bowel movement just before onset of pain. Examination shows exquisite deep tenderness to palpation in right lower abdomen with guarding but no rebound; bowel sounds are present. Pelvic examination shows a 7-cm, exquisitely tender right-sided mass. Hematocrit is 32%. Leukocyte count is 18,000/mm³. Serum amylase activity is within normal limits. Test of the stool for occult blood is negative.

Ans: B

2. An 84-year-old man in a nursing home has increasing poorly localized lower abdominal pain recurring every 3-4 hours over the past 3 days. He has no nausea or vomiting; the last bowel movement was not recorded. Examination shows a soft abdomen with a palpable, slightly tender, lower left abdominal mass. Hematocrit is 28%. Leukocyte count is 10,000/mm³. Serum amylase activity is within normal limits. Test of the stool for occult blood is positive.

Ans: E

Sample Management Set: Disposition

- A. Observe in emergency department
- B. Admit for surgery
- C. Admit for medical management
- D. Admit for endoscopy
- E. Admit for laparoscopy
- F. Order contrast studies
- G. Order MRI
- H. Order CT scan
- I. Order ultrasonography
- J. Send home with analgesics
- K. Send home for follow-up by personal physician

For each of the following patients, select the most appropriate next step in patient care.

Items might describe patients with appendicitis, ectopic pregnancy, endometriosis, Crohn's disease, diverticulitis, pelvic abscess, sickle cell crisis, renal lithiasis, twisted ovarian cyst, or other problems that commonly present as emergencies. Other disposition sets might focus on "telephone triage," hospital transfer/discharge decisions, etc.

Sample Management Set: Diagnostic Testing

- A. Test of the stool for occult blood
- B. Fasting serum glucose level
- C. Hemoglobin level
- D. Prostate-specific antigen level
- E. Serum cholesterol level
- F. Serum iron level
- G. Thyroid function tests
- H. Exercise tolerance test
- I. Digital prostate examination
- J. ECG
- K. Spirometry
- L. X-ray film of the chest
- M. Sigmoidoscopy

For each patient who comes to the physician for a health maintenance examination, select the most appropriate diagnostic study.

1. A 22-year-old man who weighs 89 kg (196 lb) and is 175 cm (69 in) tall has smoked one pack of cigarettes daily for 8 years; he does not exercise. His last examination was 5 years ago. His father had a myocardial infarction at the age of 48 years. Physical examination shows no abnormalities. **Ans: E**
2. A 28-year-old woman who weighs 70 kg (154 lb) and is 173 cm (68 in) tall has smoked one pack of cigarettes daily for 12 years; she does not exercise. Her last examination was 5 years ago, though she had a Pap smear 9 months ago that showed normal results. Her father had a myocardial infarction at the age of 48 years. Her grandmother was diagnosed with colon cancer at the age of 62 years. Physical examination shows no abnormalities. **Ans: E**

Sample Option List for Electrolyte Abnormalities

- A. Hypocalcemia
- B. Hypokalemia
- C. Hypomagnesemia
- D. Hyponatremia
- E. Hypercalcemia
- F. Hyperkalemia
- G. Hypermagnesemia
- H. Hyponatremia

For each of the following patients, select the electrolyte abnormality most likely to be present.

Sample Behavioral Sciences/Pediatrics Option List

	Cognitive/ Language Skills	Gross Motor Skills	Social Skills
A.	Normal	Normal	Normal
B.	Normal	Normal	Delayed
C.	Normal	Delayed	Normal
D.	Normal	Delayed	Delayed
E.	Delayed	Normal	Normal
F.	Delayed	Normal	Delayed
G.	Delayed	Delayed	Normal
H.	Delayed	Delayed	Delayed

For each child, select the best description of development.

Sample Nutrition/Biochemistry Option List

- | | | | |
|----|-------------------------------------|----|-----------|
| A. | Vitamin A | I. | Folate |
| B. | Vitamin B ₁ (thiamine) | J. | Biotin |
| C. | Vitamin B ₂ (riboflavin) | K. | Niacin |
| D. | Vitamin B ₆ | L. | Iron |
| E. | Vitamin C | M. | Magnesium |
| F. | Vitamin D | N. | Copper |
| G. | Vitamin E | O. | Zinc |
| H. | Vitamin K | P. | Iodine |

For each child with a metabolic abnormality, select the vitamin or mineral that is most likely to be involved.

See Appendix B for additional examples.

Steps for Organizing a Group to Write Clinical R-Sets

The following steps may be followed to utilize a group in writing clinical R-sets. Some groups have met over dinner; followed the steps below; and generated a first draft of a dozen or more items from each attendee. Others have scheduled this as a full-day off-campus retreat, with the goal of generating a pool of near-final items.

The organizer of the “item writing party” should think about how the participant’s time can best be spent. The yield will be substantially greater in terms of both quality and quantity if some work is done in advance. For example, there will be significantly smaller yield if the participants are expected to decide what topics to write on; time will be saved if the topics for each set are defined in advance of the meeting (Step #1 below). Similarly, more items will be developed if a draft of the options for each set is developed in advance (Step #3 below). More items will also be developed if a sample item is written as a model for each set (Step #4 below). There are situations where it is best to allow item writers considerable flexibility in determining what to write. In these circumstances, providing one or more option sets and allowing them to generate one or more options sets on their own might be workable.

Decisions will also need to be made about the composition of the pairs of item writers. In some circumstances, it might be best to allow participants to self-select their partners. In other circumstances, it might be best to assign the pairs. We have had the most success assembling item-writing pairs with similar interests, but including individuals with more diverse expertise as reviewers. If the goal is to generate interdisciplinary items for an exam, the same list of options may be provided to several pairs of item writers. For example, options related to abdominal pain could be given to a pair of gynecologists; a pair of surgeons; and a pair of internists. In constructing the exam, one or more items from each discipline could be merged into the same set, requiring that examinees think across disciplines in determining the most likely diagnoses of the patients.

Significant time can be saved if participants write their items on computers. This saves many hours of trying to decipher the handwriting of the item writers. We typically have two item writers work at one computer (we have found that someone automatically assumes control of the keyboard). Then, several approaches are possible for the review: reviewers may congregate around the computer to read the items on screen; the item author can read the item aloud to the reviewers who do not have a copy of the items; or hard copies can be printed for everyone to use in the review session.

- 1. Define the content domain of the exam.** For example, in developing an exam to test the ability to diagnose common clinical problems, you might define the domain by a list of the chief complaints. Time will be saved if this is done in advance of the item-writing meeting.

Abdominal mass	Dizziness	Lymphadenopathy
Abdominal pain	Easy bruising	Movement abnormalities
Anemia	Fatigue	Nausea
Ascites	Fever	Palpitations
Back pain	Gastrointestinal bleeding	Sexual dysfunction
Chest pain	Headache	Shortness of breath
Confusion	Hematemesis/Melena	Skin lesion
Cough	Itching	Syncope
Developmental delay	Jaundice	Weight change
Diarrhea	Joint pain	Wheezing

- 2. Train a group of faculty members to serve as item writers.** Training should include a brief discussion of the purpose of the exam, some sample items, and the procedures to be followed during item writing.
- 3. Divide the group into pairs to write items.** Each pair is assigned to write on 2-4 chief complaints; they generate (or modify) a list of diagnoses for each assigned complaint and write one or more patient descriptions for the diagnoses they included in their option list. Expect 20 to 60 item stems from each pair (10 to 20 per complaint). Use of computers will save considerable time in the long run.

4. Stress the following guidelines for writing stems.

Each item should describe a patient with one of the diagnoses in the option list, beginning with the patient's age, gender, chief complaint, and site of care, followed by personal history, family history (if relevant), then physical examination information, then laboratory data (if provided).

Depending upon the purpose of the set, vignettes can be brief prototypic presentations or fuller descriptions that challenge examinees to identify key information.

Each patient description should be similar in structure to the others in the set. For example, if race, ethnicity, or occupation is included in one item, include it in all items; if laboratory data are included for one item, include them in all items.

5. Merge the pairs into a larger group to review the items. One approach is to have the author read the item aloud; others attempt to provide the correct answer. The group reviews the option list and modifies the item or the option list to eliminate any ambiguity. Other approaches are outlined above.

6. Type, edit, and subject the items to external review. Items should be reviewed without the correct answer indicated after they are in their final form.

7. Construct the test. Select a sample of items from each complaint; save the remaining items for subsequent exams. Items can be converted into one-best-answer items by adding a lead-in and the best five (or more) options from the option list.

Form for Writing R-Sets

Theme: _____
(eg, a presenting complaint)

Lead-In: _____
(eg, For each patient with fever, select the most likely diagnosis.)

Options

A	N
B	O
C	P
D	Q
E	R
F	S
G	T
H	U
I	V
J	W
K	X
L	Y
M	Z

(Write items on separate pages)

Comparison of Items in Five-Option and Extended-Matching Format

In several studies to investigate the optimal number of options for multiple-choice items, we have consistently found that, other things being equal, more options are better than fewer options. Based on items used on NBME exams, extended-matching items are more discriminating than any other format; 5-option A-types are second best; and the various forms of true/false items are the worst. In controlled studies comparing otherwise equivalent 5-option and extended-matching items, extended-matching items were found to be more discriminating than 5-option items; comparable levels of reproducibility can be achieved with the extended-matching format using one-third fewer items than with 5-option items. Extended-matching items were also found to be more difficult than content-parallel 5-option items: there is a lower probability of guessing the correct answer, and item writers are not always able to select the most functional distractors in reducing the number of options to five.

The following table shows examinee responses to a sample item presented in a 5-option format and in a 15-option format. The item was originally written as a 15-option item; the item writer then reduced the number of options to five by selecting what he thought were the best distractors (B, F, G, J, N). The item was markedly easier in the 5-option format (p value of 81 vs 59), and the discrimination was markedly lower (not shown). There is an increased probability of examinees selecting the correct answer in the 5-option format, especially because item writers do not uniformly identify the most salient distractors (eg, option D). Tests constructed of extended-matching items tend to spread out the lower ability students; the extended list of options gives them more opportunity to show what they don't know.

Examinee Responses

Format	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5-option		81		*		4	0			14			*	1	
15-option	2	59	2	13	1	2	1	0	1	7	2	2	8	0	0

For additional information, see also:

Case SM, Swanson DB, Ripkey DR. Comparison of items in five-option and extended-matching format for assessment of diagnostic skills. *Academic Medicine*. 1994;69(suppl):S1-S3.

Case SM, Swanson DB. Extended-matching items: a practical alternative to free-response questions. *Teaching and Learning in Medicine*. 1993;5(2):107-115.

Swanson DB, Case SM. Trends in written assessment: a strangely biased perspective. In: Harden R, Hart I, Mulholland H, eds. *Approaches to Assessment of Clinical Competence*. Norwich, England: Page Brothers; 1992:38-53.

A's to R's and Back Again

Although research has consistently shown a psychometric advantage of Extended-Matching (R-type) items over five-option A-type items, there may be circumstances where you need to convert items from one format to another. This should be straightforward.

For example, the following item was written as a five-option A-type item:

A patient with the classic phenotypic features of trisomy 21 (Down syndrome) has 46 chromosomes on each of 100 metaphase karyotypes. Which of the following is the most likely explanation for this finding?

- | | |
|----------------------------|------------------------------|
| <i>A. Deletion</i> | <i>*D. Translocation</i> |
| <i>B. Mosaicism</i> | <i>E. Undetected trisomy</i> |
| <i>C. Somatic mutation</i> | |

This can easily be transformed into the following extended-matching item. Once in this format, additional stems can be written to expand the set.

- | | |
|----------------------------|------------------------------|
| <i>A. Deletion</i> | <i>D. Translocation</i> |
| <i>B. Mosaicism</i> | <i>E. Undetected trisomy</i> |
| <i>C. Somatic mutation</i> | |

For each patient with genetic abnormalities, select the genetic pattern that is most likely to be involved.

A patient with the classical phenotypic features of trisomy 21 (Down syndrome) has 46 chromosomes on each of 100 metaphase karyotypes.

Additional options could be written to make the examinee task more challenging.

- | | |
|------------------------------|---------------------------------|
| <i>A. Deletion</i> | <i>F. Somatic mutation</i> |
| <i>B. Genomic imprinting</i> | <i>G. Translocation</i> |
| <i>C. Mosaicism</i> | <i>H. Undetected trisomy</i> |
| <i>D. Pleiotropy</i> | <i>I. Variable expressivity</i> |
| <i>E. Reduced penetrance</i> | |

For each patient with genetic abnormalities, select the genetic pattern that is most likely to be involved.

Items from this set could be converted back to single A-type items with five or more options.

Chapter 7

Pick N Items: An Extension of the Extended-Matching Format



The Pick N format may be similar to either the extended-matching or the A-type format; the primary difference is that the examinee is told to pick 2, 3, 4, or even 5 of the options listed. As with extended-matching sets, the option list may include up to 26 options. The format was developed to replace negative items or items with double options, particularly in areas such as Health Maintenance and Disease Prevention. Items might focus on various patients with different risk factors who come for routine examination; the examinee would be asked to select the laboratory studies or immunizations that should be ordered for each patient. The format could also be used for emergency management items where several procedures would be carried out simultaneously; the examinee would be asked to select a specific number of actions from the option list.

The item-writing rules are the same as for extended-matching sets. The options should be short (usually a single word or very short phrase); the patient vignettes can be long. If the set asks for management decisions, each vignette should contain all relevant history and physical examination data. As with extended-matching items, the format works well for items that appear to be extremely easy; tricky or unnecessarily complex vignettes should be avoided.

The Pick N format is designed to specify exactly how many options to select. The rationale for this decision is derived from the essential difference between true/false and one-best-answer items, where true/false items require the examinee to indicate all responses that are appropriate, and one-best-answer items require the examinee to indicate a specific number of responses. Specifying the number of options to be selected changes the task from a true/false task to a best-answer task.

Research indicates that partial credit scoring is preferred but logistical considerations might preclude this. If you use all-or-nothing scoring, the items may be extremely difficult, and it is best to require examinees to select only two or three options, rather than more.

For additional information, see also:

Ripkey DR, Case SM, Swanson DB. A “new” item format for assessing aspects of clinical competence. *Academic Medicine*. 1996;71 (suppl):S34-S36.

- | | |
|--------------------------------------|-------------------------------------|
| A. <i>Diabetic polyneuropathy</i> | F. <i>Multiple sclerosis</i> |
| B. <i>Huntington's disease</i> | G. <i>Parkinson's disease</i> |
| C. <i>Lateral medullary syndrome</i> | H. <i>Pontine glioma</i> |
| D. <i>Lead encephalopathy</i> | I. <i>Tabes dorsalis</i> |
| E. <i>Medulloblastoma</i> | J. <i>Wernicke's encephalopathy</i> |

A 50-year-old man has the gradual onset of mental confusion, disorientation, and loss of short-term memory. He has a left footdrop. A blood smear shows microcytosis and basophilic stippling of erythrocytes. (Select the two most likely diagnoses)

In the example above, there would be disagreement about what diagnoses are likely, but the task becomes clear if the examinee is told to select the two most likely diagnoses. The options can be diagnosed as follows.

F	C	E	D	A	G	H	B
<hr/>							
<i>Least Likely Diagnosis</i>						<i>Most Likely Diagnosis</i>	

Sample Pick N Set

- | | |
|--------------------------------------|---|
| A. Calcium | G. Vitamin B ₆ |
| B. Fluoride | H. Vitamin B ₁₂ (cyanocobalamin) |
| C. Folic acid | I. Vitamin C |
| D. Iron | J. Vitamin D |
| E. Vitamin A | K. Vitamin E |
| F. Vitamin B ₁ (thiamine) | |

For each child, select the appropriate vitamin or mineral supplements.

1. A 1-month-old infant is brought to the physician for a well-child examination. He has been exclusively breast-fed, and examination shows normal findings. (SELECT 2 SUPPLEMENTS). **Ans: B, J**
2. A 6-year-old girl has cystic fibrosis. She has been taking no medications. (SELECT 3 SUPPLEMENTS). **Ans: E, J, K**

Sample Pick N Set

- | | |
|--|------------------------------|
| A. Analysis and culture of cerebrospinal fluid | F. Urinalysis |
| B. Blood culture | G. Urine culture |
| C. Complete blood count | H. X-ray film of the abdomen |
| D. Examination of the stool for leukocytes | I. X-ray film of the chest |
| E. Measurement of serum electrolyte levels | |

For each child with fever, select the appropriate initial diagnostic studies.

1. A previously healthy 1-year-old girl is brought to the emergency department because of fever for 1 day. Her temperature is 41 C (105.8 F). She is otherwise asymptomatic. Physical examination shows no abnormalities. (SELECT 4 STUDIES). **Ans: B, C, G, I**
2. A previously healthy 10-day-old newborn is brought to the emergency department because of fever for 2 hours. He was born at term after an uncomplicated pregnancy. His temperature is 39 C (102.2 F). Physical examination shows no abnormalities. (SELECT 6 STUDIES). **Ans: A, B, C, E, G, I**
3. A 7-year-old boy with sickle cell disease is brought to the emergency department because of fever for 1 day and chest pain for 1 hour. His temperature is 39.5 C (103.1 F). Breath sounds are slightly decreased in the right lower lung; he is not in respiratory distress. (SELECT 3 STUDIES). **Ans: B, C, I**

PATIENT CHART**Patient History**

Sex: male
Current age: 28 years
Chief complaint: health maintenance examination
Social history:
Marital status: single
Occupation: computer programmer
Alcohol: 2-4 beers/weekend
Smoking: 5-10 cigarettes daily from age 16-24
Exercise prgm: sedentary
Medical history:
Childhood: obese since grade school
Immunizations: all childhood immunizations; last tetanus toxoid age 15;
no immunizations since childhood
Screening: no physician visits since college
Family history:
Parents: father age 57; hypertensive mother age 55; obese and hypertensive
Siblings: none
Children: none
Current medications: none
Allergies: none

Physical Examination

Height: 178 cm (70 in)
Weight: 134 kg (295 lb)
Vital signs:
Blood pressure 148/86 mm Hg
Pulse 90/min
Respirations 16/min
Skin: erythematous rash in groin
Abdominal: obese
Laboratory studies: none ordered

1. For the patient whose chart is shown, select the conditions for which he is at increased risk.
(SELECT 4 CONDITIONS)

- | | |
|----------------------------|----------------------------|
| A. Alcoholism | G. Hypertension |
| B. Colon cancer | H. Hypothyroidism |
| C. Coronary artery disease | I. Osteoarthritis |
| D. Diabetes mellitus | J. Skin cancer |
| E. Hemochromatosis | K. Thyroid cancer |
| F. HIV | L. Urinary tract infection |

2. For the patient whose chart is shown, select the most appropriate health maintenance interventions.
(SELECT 4 INTERVENTIONS)

- | | |
|--------------------------------|--|
| A. Complete blood count | H. Serum urea nitrogen (BUN) and creatinine level measurements |
| B. Dietetic counseling | I. Tetanus toxoid vaccine |
| C. Exercise counseling | J. Thyroid function tests |
| D. Fasting serum lipid profile | K. Urinalysis |
| E. Hepatitis B immunization | L. X-ray film of the chest |
| F. HIV testing | |
| G. Influenza immunization | |