Preface

The mission of the NBOME is:

To protect the public by providing the means to assess competencies for osteopathic medicine and related health care professions.

Of the many endeavors in which you can participate, the one for which the NBOME is requesting your assistance is special in that it provides an outlet to serve the profession and the general public. At the same time you will gain a collegial experience and the opportunity to increase your scope of medical knowledge.

Many clinicians and academicians convey knowledge to osteopathic medical students in the course of teaching basic and clinical sciences. However, when called upon to create effective, reliable, and psychometrically sound exam items, intended to assess the knowledge and skills that have been imparted to students, these attempts often fall short of the mark.

In order to prepare all item-writers who contribute exam items for incorporation into COMLEX-USA, this Item-Writing Guide was created to assist in setting standards for exam item-writing and to facilitate the efforts of content experts in creating good exam items that will comprise an effective tool by which medical boards may make their licensure decisions.

The NBOME is greatly appreciative of the time and effort that is given to this task, and is assured that you will find the experience of creating quality exam items both challenging and rewarding.

Thank you for your participation!
<table>
<thead>
<tr>
<th>Table of Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item-Writing Overview (1.0) .................................................................1</td>
</tr>
<tr>
<td>Examination Levels (2.0) ...........................................................................1</td>
</tr>
<tr>
<td>Examination Content (3.0) ..................................................................... 1</td>
</tr>
<tr>
<td>Dimension I – Patient Presentation – All Levels (3.1) .......................... 2</td>
</tr>
<tr>
<td>Dimension II – Physician Task (3.2) .................................................. 3</td>
</tr>
<tr>
<td>Item Types (4.0) .................................................................................... 3</td>
</tr>
<tr>
<td>Approved Item Types (4.1) .................................................................... 3</td>
</tr>
<tr>
<td>Non-Approved Item Types (4.2) ......................................................... 3</td>
</tr>
<tr>
<td>Item Assignment Creation (5.0) ......................................................... 4</td>
</tr>
<tr>
<td>Sample Level 1 Assignment (5.1) ...................................................... 5</td>
</tr>
<tr>
<td>Steps in Writing Items (6.0) ................................................................ 7</td>
</tr>
<tr>
<td>Item Creation (6.1) .............................................................................. 7</td>
</tr>
<tr>
<td>Completed Item (6.2) ............................................................................. 9</td>
</tr>
<tr>
<td>Researching and Referencing Content for an Item (7.0) .......................10</td>
</tr>
<tr>
<td>Constructing Clinical Scenarios (8.0) ..................................................13</td>
</tr>
<tr>
<td>Interrogatories (9.0) ............................................................................14</td>
</tr>
<tr>
<td>Health Promotion &amp; Disease Prevention (9.1) ......................................14</td>
</tr>
<tr>
<td>History &amp; Physical Examination (9.2) ................................................14</td>
</tr>
<tr>
<td>Diagnostic Technologies (9.3) ......................................................... 15</td>
</tr>
<tr>
<td>Management (9.4) .................................................................................15</td>
</tr>
<tr>
<td>Scientific Understanding of Health &amp; Disease Mechanisms (9.5) ..........15</td>
</tr>
<tr>
<td>Health Care Delivery Issues (9.6) ...................................................... 16</td>
</tr>
<tr>
<td>Clinical Scenario Patterns (10.0) ....................................................... 16</td>
</tr>
<tr>
<td>Item Format Basics (11.0) ................................................................. 17</td>
</tr>
<tr>
<td>Item Flaws (12.0) ................................................................................17</td>
</tr>
<tr>
<td>Pseudocases (12.1) .............................................................................17</td>
</tr>
<tr>
<td>Unfocused Items (12.2) .................................................................... 18</td>
</tr>
<tr>
<td>Items With Flaws That Favor Testwise Candidates (12.3) ..................18</td>
</tr>
<tr>
<td>Items With Flaws Due To Irrelevant Difficulty (12.4) .......................18</td>
</tr>
<tr>
<td>Items With Teaching in the Stem (12.5) ..............................................19</td>
</tr>
<tr>
<td>Item Quality Checklist (13.0) ..............................................................19</td>
</tr>
<tr>
<td>Constructing Matching Items (14.0) ..................................................20</td>
</tr>
<tr>
<td>Constructing Distractors (15.0) .........................................................21</td>
</tr>
<tr>
<td>OPP/OMM Integration (16.0) ............................................................ 22</td>
</tr>
<tr>
<td>Using Bloom's Taxonomy (17.0) ....................................................... 23</td>
</tr>
<tr>
<td>Evidence-Based Medicine (EBM) Principles in Item-Writing (18.0) ....25</td>
</tr>
<tr>
<td>Definitions (18.1) ..............................................................................26</td>
</tr>
<tr>
<td>General Principles (18.2) ................................................................ 27</td>
</tr>
<tr>
<td>Writing the Exam Item (18.3) ........................................................... 29</td>
</tr>
<tr>
<td>Core Competencies (19.0) .................................................................. 31</td>
</tr>
<tr>
<td>Writing Multimedia and Advanced Type Exam Items (20.0) ...............31</td>
</tr>
<tr>
<td>Media Types (20.1) ............................................................................32</td>
</tr>
<tr>
<td>Guidelines of Media Use (20.2) ....................................................... 34</td>
</tr>
<tr>
<td>Specifications of Submitted Media (20.3) .........................................34</td>
</tr>
<tr>
<td>Creating A Multimedia Exam Item (20.4) ..........................................35</td>
</tr>
<tr>
<td>Exam Development Process (21.0) .................................................... 36</td>
</tr>
<tr>
<td>Item-Writing Practice Session (22.0) ................................................38</td>
</tr>
<tr>
<td>Critiquing Written Items (22.1) ....................................................... 38</td>
</tr>
<tr>
<td>Building a New Exam Item (22.2) .................................................... 40</td>
</tr>
<tr>
<td>Resources (23.0) .................................................................................42</td>
</tr>
</tbody>
</table>
Item-Writing Overview (1.0)
You have probably taken standardized tests since you were very young. Some of the questions you have seen were straight-forward (“What is the product of 12 and 2?”), but as your education increased so did the complexity of the questions you were asked.

The National Board of Osteopathic Medical Examiners creates examinations for the purpose of assisting state medical boards in making determinations regarding medical licensure. By definition, then, the examinations we create are high-stakes exams. Because of this, the type of questions (or “items”) we use in our COMLEX-USA (Comprehensive Osteopathic Medical Licensing Examination) series must be highly controlled and of the best possible quality. These items are designed to meet specific criteria, and it is important that item-writers, as content experts and authors, abide by the guidelines set forth by the NBOME.

Item-writing involves:
- Content delineation
- Familiarization with item format
- Solid writing techniques
- Research
- Referencing supported by medical evidence

Our intention is to have item-writers write the most exquisite items feasible for osteopathic assessment. In doing this, an important principle to keep in mind is that we do not want to assess the recall of isolated facts. We want to assess skill in interpreting data and making decisions.

Examination Levels (2.0)
The content and difficulty of each item are to be tailored to the educational level of the candidate. There is no objective measurement for this. The curriculum generally covered in each year of osteopathic medical school or by current acceptable clinical practices and protocols should match the content included by item-writers in the exam items intended for that level of testing. Examinations are generally taken as follows:

<table>
<thead>
<tr>
<th>Level</th>
<th>After two years of osteopathic medical school</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2-CE</td>
<td>After three years of osteopathic medical school</td>
</tr>
<tr>
<td>Level 3</td>
<td>During postdoctoral training</td>
</tr>
</tbody>
</table>

Item-writers are cautioned to be aware of the level of the candidate for whom their items are intended.

Examination Content (3.0)
When formulating an exam item, it is helpful for item-writers to understand how their contribution relates to the examination as a whole.

A COMLEX-USA examination is made up of 400 items, and each COMLEX-USA examination administration has a unique combination of items. There are at least 40 administrations each year for each of the three levels—for a total of over 120 examination days.
Examinations are built from two dimensions: patient presentation and physician task. Each item is classified as to the reason the patient is being seen by a physician, and it is also classified based on what type of medical knowledge or skill must be applied to the patient’s situation. A percentage of the examination is assigned to each presentation (or group of presentations) based on the significance and frequency of that condition in the realm of general osteopathic medical practice.

The percentage of an examination based on each physician task is determined by the examination level. Level 1 is largely devoted to Scientific Understanding of Health & Disease Mechanisms, whereas Levels 2-CE and 3 have varying degrees of emphasis from all six physician task categories. Osteopathic principles are integrated across both dimensions, and no specific percentage allocation of such items is specified in the blueprint. Item-writers are to consider incorporating OPP/OMT wherever such principles and practices are appropriate and applicable. The breakdowns for the two exam dimensions are as follows:

### Dimension I – Patient Presentation – All Levels (3.1)

<table>
<thead>
<tr>
<th>Main Topic</th>
<th>Percentage</th>
<th>Patient Presentation Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population Health Concepts and Patients with Presentations Related to Health Promotion, Chronic Disease Management, and Human Development</td>
<td>8-16%</td>
<td>Asymptomatic/General</td>
</tr>
<tr>
<td>Patients with Presentations Related to Digestion &amp; Metabolism</td>
<td>4-10%</td>
<td>Digestive Difficulties</td>
</tr>
<tr>
<td>Patients with Presentations Related to Cognition, Behavior, Sensory &amp; Central Nervous Systems, Substance Abuse, and Visceral and Sensory Pain</td>
<td>28-38%</td>
<td>Cognitive Difficulties, Consciousness Alterations, Fatigue &amp; Weakness, Sensory &amp; CNS Difficulties, Sensory Pain, Substance Abuse</td>
</tr>
<tr>
<td>Patients with Presentations Related to the Musculoskeletal System, including Somatic Pain</td>
<td>6-12%</td>
<td>Musculoskeletal Difficulties &amp; Muscular Pain</td>
</tr>
<tr>
<td>Patients With Presentations Related to the Genitourinary System and Human Sexuality</td>
<td>3-8%</td>
<td>Genitourinary Disorders/Issues</td>
</tr>
<tr>
<td>Patients With Presentations Related to Circulation and the Respiratory System</td>
<td>8-16%</td>
<td>Bleeding, Respiratory Difficulties</td>
</tr>
<tr>
<td>Patients With Presentations Related to Thermoregulation</td>
<td>2-6%</td>
<td>Fever &amp; Hypothermia</td>
</tr>
<tr>
<td>Patients With Presentations Related to Trauma, Masses, Edema, Discharge, and the Skin, Hair and Nails</td>
<td>8-16%</td>
<td>Discharge, Masses &amp; Edema, Skin, Nail, Hair &amp; Tooth Disorders, Trauma</td>
</tr>
<tr>
<td>Patients With Presentations Related to Pregnancy, the Peripartum, and the Neonatal Period</td>
<td>3-8%</td>
<td>Pregnancy/Childbirth/Postpartum/Neonatal Assessment</td>
</tr>
</tbody>
</table>
### Dimension II – Physician Task (3.2)

<table>
<thead>
<tr>
<th>Physician Skill Topic</th>
<th>Level 1 Percentage</th>
<th>Level 2-CE Percentage</th>
<th>Level 3 Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Promotion &amp; Disease Prevention</td>
<td>1-5%</td>
<td>15-20%</td>
<td>15-20%</td>
</tr>
<tr>
<td>History &amp; Physical Examination</td>
<td>5-15%</td>
<td>30-40%</td>
<td>10-20%</td>
</tr>
<tr>
<td>Diagnostic Technologies</td>
<td>1-5%</td>
<td>10-20%</td>
<td>15-25%</td>
</tr>
<tr>
<td>Management</td>
<td>2-7%</td>
<td>10-20%</td>
<td>25-40%</td>
</tr>
<tr>
<td>Scientific Understanding of Health &amp; Disease Mechanisms</td>
<td>70-85%</td>
<td>5-15%</td>
<td>5-10%</td>
</tr>
<tr>
<td>Health Care Delivery Issues</td>
<td>1-3%</td>
<td>5-10%</td>
<td>5-10%</td>
</tr>
</tbody>
</table>

An emphasis is placed on high-frequency and/or high-impact clinical and basic science situations.

The COMLEX-USA examination blueprint is constructed by reviewing data that reflects the actual practice patterns of osteopathic physicians on a nationwide basis. Material incorporated into the curricula of the osteopathic medical schools was also considered in the examination design. Expert opinion gathered from academicians and clinicians from all geographic locations of the country was taken into account as the final design was completed. The examination structure is periodically reviewed to ensure that it is applicable and useful in making osteopathic licensure decisions.

### Item Types (4.0)

Although the NBOME only uses four basic item types, item-writers should be familiar with other item types used for other purposes. Only approved item types are considered for inclusion in COMLEX-USA examinations.

#### Approved Item Types (4.1)

<table>
<thead>
<tr>
<th>Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Type</td>
<td>One multiple-choice item based on a clinical scenario</td>
</tr>
<tr>
<td>B-Type</td>
<td>A list of five lettered headings followed by a series of clinical scenarios (each a separate exam item). Each exam item is most closely associated with only one of the headings</td>
</tr>
<tr>
<td>S-Type</td>
<td>Two or more multiple-choice items linked to a single clinical scenario</td>
</tr>
<tr>
<td>Extended A-Type</td>
<td>Similar to A-Type in appearance but with greater than five answer choices</td>
</tr>
<tr>
<td>X-Type</td>
<td>Similar to B-Type but with greater than five lettered headings (6 to 25)</td>
</tr>
</tbody>
</table>

#### Non-Approved Item Types (4.2)

<table>
<thead>
<tr>
<th>Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-Type</td>
<td>Similar to B-Type in appearance; the candidate chooses “A” if (A) is true, “B” if (B) is true, “C” if both are true, or “D” if neither is true</td>
</tr>
<tr>
<td>D-Type</td>
<td>A complex matching set of three lettered headings of functional disturbances and five numbered situations; the candidate selects the functional disturbance that four of the situations are related to and indicates the situation that does not belong</td>
</tr>
</tbody>
</table>
E-Type | A sentence with two main parts: an assertion and a reason for that assertion; the candidate selects whether the assertion and reason are both true, both false, or a combination, and whether the reason explains the assertion
---|---
H-Type | An item that consists of a pair of statements describing two entities to be compared in a quantitative sense; the candidate selects “A” if (A) is greater than (B); “B” if (B) is greater than (A); and “C” if the two are approximately equal
I-Type | Similar to H-Type but with pairs of phrases that describe conditions or quantities that might vary in relation to each other; the candidate selects what type of relationship these variables have to each other
K-Type | A multiple true/false item consisting of a stem followed by four options, one or more of which is correct

The majority of items in a COMLEX-USA examination consist of A-Type items. These items contain a clinical scenario or case vignette, a question or incomplete statement, and five answer choices, only one of which is correct. (Four answer choices are only allowed under certain circumstances when there are only four plausible options. The NBOME requires that all item-writers submit five answer choices.)

Some items are grouped into case sets (S-Types) with two (or more) items based on one clinical scenario. Each item can address a different physician task: one might ask about the best diagnostic test to order, while another might ask about how this disorder could have been prevented. Recognizing the psychometric impact, a relatively small percentage of our exams consists of cases. When writing them, item-writers should attempt to utilize independent interrogatories for each item, so that if a candidate answers one incorrectly, it does not necessarily affect his or her responses to the others.

Matching sets (B- and X-Types) are occasionally used in COMLEX-USA examinations. Such sets have a list of options (usually 5-9), followed by two or more clinical scenarios.

**Item Assignment Creation (5.0)**

Items are pretested before being used in the scoring of an examination. Pretesting, along with subsequent testing, provides statistical information that reflects the performance of an item in the examination(s) in which it was used. Good statistics signify that an item is useful in comparing the performance of a candidate against the standard, expected performance of an individual based on a criterion-referenced group of candidates. Therefore, items used in COMLEX-USA examinations must have good test performance.

Information summarizing the number of items available for use in the coming exam cycle—items with good statistics—is reviewed annually. The data is then evaluated in terms of several variables (listed below). Item assignments are then made, assigning items to meet test design specifications.
The following variables are considered when creating item assignments:

- Content (both dimensions)
- Discipline (specialty/subject area)
- Item format (e.g., A-Type, B-Type, S-Type, X-Type)
- Items with visual or multimedia enhancements

An author may be given a visual or other digital file (e.g., audio, motion) and be assigned to write a corresponding item. The NBOME also encourages writers to submit high-quality visuals and other digital files, which can then be used, where appropriate, as the foundation for items.

The assignment an item-writer receives will be specific as to the patient presentation (presenting complaint) and physician task to be addressed. It will also indicate the Level for which the item is to be written. It will not, however, specify an age, gender, setting, or other details regarding the clinical scenario. Whether or not to include such information is left to the judgment of the content expert writing the item.

A sample Level 1 item-writing assignment appears below:

**Sample Level 1 Assignment (5.1)**

<table>
<thead>
<tr>
<th>Name:</th>
<th>[Name]</th>
<th>Information Update:</th>
</tr>
</thead>
<tbody>
<tr>
<td>AOA number:</td>
<td>[AOA number]</td>
<td>____________________</td>
</tr>
<tr>
<td>Institution:</td>
<td>[name of school, if applicable]</td>
<td>____________________</td>
</tr>
<tr>
<td>Department:</td>
<td>[name of dept., if applicable]</td>
<td>____________________</td>
</tr>
<tr>
<td>Address:</td>
<td>[street address]</td>
<td>____________________</td>
</tr>
<tr>
<td></td>
<td>[city, state  zip]</td>
<td>____________________</td>
</tr>
<tr>
<td>Phone Number:</td>
<td>[phone number]</td>
<td>____________________</td>
</tr>
<tr>
<td>E-mail Address:</td>
<td>[e-mail address]</td>
<td>____________________</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item Type</th>
<th>Topic</th>
<th>Subtopic</th>
<th>Etiology</th>
<th>Physician Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Type</td>
<td>Cognitive</td>
<td>behavioral disturbances</td>
<td>Alzheimer disease</td>
<td>Scientific Understanding of Health &amp; Disease Mechanisms</td>
</tr>
<tr>
<td></td>
<td>Difficulties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-Type</td>
<td>Cognitive</td>
<td>behavioral disturbances</td>
<td>mood disorders</td>
<td>Scientific Understanding of Health &amp; Disease Mechanisms</td>
</tr>
<tr>
<td></td>
<td>Difficulties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-Type</td>
<td>Cognitive</td>
<td>depression</td>
<td>drug-induced</td>
<td>Scientific Understanding of Health &amp; Disease Mechanisms</td>
</tr>
<tr>
<td></td>
<td>Difficulties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-Type</td>
<td>Sensory &amp; CNS Difficulties</td>
<td>seizures</td>
<td>drug-induced</td>
<td>Scientific Understanding of Health &amp; Disease Mechanisms</td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------</td>
<td>----------------</td>
<td>--------------</td>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>A-Type</td>
<td>Substance Abuse</td>
<td>alcohol abuse</td>
<td>behavioral abnormalities</td>
<td>Scientific Understanding of Health &amp; Disease Mechanisms</td>
</tr>
<tr>
<td>A-Type</td>
<td>Substance Abuse</td>
<td>controlled substance abuse</td>
<td>benzodiazepine abuse</td>
<td>Scientific Understanding of Health &amp; Disease Mechanisms</td>
</tr>
<tr>
<td>A-Type</td>
<td>Substance Abuse</td>
<td>controlled substance abuse</td>
<td>hallucinogen abuse</td>
<td>Scientific Understanding of Health &amp; Disease Mechanisms</td>
</tr>
</tbody>
</table>

**Total Number of Items Assigned: 7**

Before writing each item, an item-writer is encouraged to outline additional pertinent parameters to include, such as:

- Patient age
- Patient gender
- Setting of care
- Medical history or habits, including drug allergies or interactions
- Family or occupational history
- Laboratory or medical imaging findings (pertinent)
- Type of condition (acute vs. chronic)
- Ethnicity (when essential to the item)

The item-writer will then incorporate these factors into the item. Applying these additional parameters will help provide structure and depth for the exam item.

When outlining the parameters listed, please keep balance in mind over the course of the assignment. It is best to vary the age and type of condition, for example, but the type of patients with which the item-writer works most often may be emphasized. For example, if an item-writer is an emergency department physician, the majority of the exam items should be set in the emergency department.

Given the time allotted for reading an exam item and the physical constraint posed by the space allowed on the computer monitor during CBT testing, not all parameters or factors can be used in each constructed exam item. Some “irrelevant” information may be included if the intent is to determine whether the candidate can eliminate nonessential information in order to arrive at the solution to the problem. Item-writers should strive for concise, clearly-written clinical scenarios or case vignettes.
Steps in Writing Items (6.0)

First of all, items written for any other purpose CANNOT be submitted to the NBOME for use in COMLEX-USA exams. Please note that the NBOME cannot accept items that have been created or used for any other reason—in a course, for another publication, for a review session, etc. Items submitted to the NBOME must be original, unpublished, and unseen by osteopathic medical students. They also cannot be used in the future for any purpose other than for COMLEX-USA examinations. All submitted items become the property of the NBOME.

Visuals may be submitted, but they must be obtained privately (not through any publication or the Internet). Permission for use must be obtained and submitted from any identifiable patient shown.

It is important to determine what candidates must know and understand (and be able to apply) in terms of the content of the assignment. Item-writers should focus on high-impact and/or high-frequency issues for this examination.

Item Creation (6.1)

Suppose the assignment is:

**Patient presentation:** dyspnea due to congestive heart failure

**Physician task:** diagnostic technologies

The item-writer will start by considering associated tasks that are essential to a physician’s application of diagnostic technologies to a patient presenting with dyspnea due to congestive heart failure. Now be careful; the item should assess skill in interpreting data and making decisions—not the ability to recall facts.

1. Create a clinical scenario in which a patient has dyspnea. Do not simply relay the fact that it is due to congestive heart failure to the examinee.

   A 76-year-old male presents to the emergency department with the complaint of progressive shortness of breath of eight hours’ duration. He has never had dyspnea like this before.

2. Give enough additional history in the scenario to help the examinee with information regarding past medical history that may be impacting the patient.

   A 76-year-old male presents to the emergency department with the complaint of progressive shortness of breath of eight hours’ duration. He has never had dyspnea like this before. He has a history of hypertension and admits to a lack of compliance with his low-salt diet. He has also neglected to take his medications, diltiazem and hydrochlorothiazide, for two weeks.
3. If a chest radiograph is available showing frank Kerley lines or other signs of congestive heart failure, it can be used as a visual with the item testing on the next step in the work-up of this patient. Alternatively, the results of the chest radiograph can be described.

A 76-year-old male presents to the emergency department with the complaint of progressive shortness of breath of eight hours’ duration. He has never had dyspnea like this before. He has a history of hypertension and admits to a lack of compliance with his low-salt diet. He has also neglected to take his medications, diltiazem and hydrochlorothiazide, for two weeks. A chest radiograph taken in the emergency department reveals a cardiac:thoracic ratio of 0.6, Kerley lines, and small bilateral pleural effusions. The most appropriate test to determine the etiology of this patient’s disorder is

4. Create four distractors to add to the correct answer, and mark the correct answer with an asterisk.

A 76-year-old male presents to the emergency department with the complaint of progressive shortness of breath of 8 hours’ duration. He has never had dyspnea like this before. He has a history of hypertension and admits lack of compliance with his low-salt diet, and has neglected to take his medications, diltiazem and hydrochlorothiazide, for two weeks. A chest radiograph taken in the emergency department shows a cardiac:thoracic ratio of 0.6, Kerley lines, and small bilateral pleural effusions. The most appropriate test to determine the etiology of this patient’s disorder is

(A) arterial blood gas analysis
(B) beta-natriuretic peptide level **
(C) C-reactive protein level
(D) pulmonary function testing
(E) ventilation-perfusion scan

5. Indicate a current acceptable text that validates the item content.

Ref: Braunwald’s Heart Disease – A Textbook of Cardiovascular Medicine, 9th edition, 2011, Chapter 26: Clinical Assessment of Heart Failure

Now, suppose the assignment is:

**Patient Presentation:** drug-induced nausea & vomiting

**Physician Task:** Scientific Understanding of Health & Disease Mechanisms

The item-writer will start by considering associated tasks that are essential to a physician’s application of Scientific Understanding of Health & Disease Mechanisms to a patient presenting with nausea due to drug use. Again, be careful: this item should not assess the understanding of a fact regarding this situation. The item should assess skill in interpreting data and making decisions—not the ability to recall facts.
1. Create a clinical scenario:

A 56-year-old female presents with a two-day history of nausea. Past medical history reveals an elbow strain three weeks earlier. She has been taking ibuprofen to relieve the elbow pain.

2. Create a question based on the physician task:

Viscerosomatic changes related to this patient’s nausea would most likely be found at

3. Create distractors based on the correct answer:

(A) C3-C5
(B) T1-T4
(C) T5-T9 **
(D) T10-L1
(E) L2-L4

4. Reference the item with a current acceptable text:

Ref: Digiovanna, Osteopathic Approach to Diagnosis and Treatment, 3rd edition, 2004, P47

Completed Item (6.2)

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Digestive difficulties</th>
<th>Nausea &amp; Vomiting, drug-induced</th>
<th>Scientific Understanding of Health &amp; Disease Mechanisms</th>
</tr>
</thead>
</table>

A 56-year-old female presents with a two-day history of nausea. Past medical history reveals an elbow strain three weeks earlier. She has been taking ibuprofen to relieve the elbow pain. Viscerosomatic changes related to this patient’s nausea would most likely be found at

(A) C3-C5
(B) T1-T4
(C) T5-T9 **
(D) T10-L1
(E) L2-L4

Ref: Digiovanna, Osteopathic Approach to Diagnosis and Treatment, 3rd ed., 2004, p 47

Other items can be created based on the same basic stem but with different physician tasks. You may need to change the level to match the emphasis of different level exams.

<table>
<thead>
<tr>
<th>Level 3</th>
<th>Digestive difficulties</th>
<th>Nausea &amp; Vomiting, drug-induced</th>
<th>Management</th>
</tr>
</thead>
</table>

A 56-year-old female presents with a two-day history of nausea. Past medical history reveals an elbow strain three weeks earlier. She has been taking ibuprofen to relieve the elbow pain. The most appropriate treatment for this patient’s nausea is
Level 2-CE | Digestive difficulties | Nausea & Vomiting, drug-induced | History and Physical

A 36-year-old female complains of nausea after using naproxen for the past week to treat a cervical strain. The most likely diagnosis is

Level 3 | Digestive difficulties | Nausea & Vomiting, drug-induced | Health Promotion and Disease Prevention

A 66-year-old female with chronic low back pain asks about using naproxen daily for pain relief. In the past she has had nausea with the medication, but otherwise it seems to help her symptoms. Which of the following medications can she take with this medicine to decrease gastrointestinal symptoms?

Researching and Referencing Content for an Item (7.0)

Once the topic and parameter information for an item are designated, the next step for the item-writer is to conduct some research. The writer is to find the content in a standard reference source and provide a detailed description of where the information can be found.

All references used should be:

1. Accepted as authoritative in the domain being tested
2. Current in concept, classification, and/or criteria
3. Universally available to the population being tested
4. Reflective of the highest level of medical evidence practicable
5. Quoted accurately as to content
6. Generally accepted (rather than an author’s opinion)

Once again, EVERY item written for the NBOME must be referenced properly, or it cannot be accepted for use.

The guidelines to follow are explained in detail below.

LEVEL 1 (EXCLUDING HPDP, HCD, & OPP/OMM)

TEXTBOOKS

Only textbooks are acceptable to use. Please use the most recent edition.

- Hard copies of current, common texts (include title, author, edition, year & page number)
- Texts available on MDConsult or Access Medicine (include title, author, edition, year, chapter number & section title)
Textbooks are the most appropriate source for many questions in which the basic underlying principles have not changed. Please use the most recent edition.

- Hard copies of current, common texts (include title, author, edition, year & page number)
- Texts available on MDConsult or Access Medicine (include title, author, edition, year, chapter number & section title)

Peer-reviewed, current journal articles are appropriate for use (include journal name, volume, issue, page numbers, author & article title).

Appropriate websites can be used to reference specific information. Providing only a general web address that does not link to a specific article is not acceptable.

**ALL INTERNET REFERENCES MUST HAVE BEEN REVIEWED WITHIN THE PAST FIVE YEARS**

<table>
<thead>
<tr>
<th>Accepted Sites</th>
<th>Not Accepted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government sites (ending in .gov)</td>
<td>Abstracts</td>
</tr>
<tr>
<td>Nationally-recognized sites providing clinical guidelines (e.g., CDC, USPSTF, AHA)</td>
<td>Conference proceedings</td>
</tr>
<tr>
<td>ACP Journal Club</td>
<td>Kaiser Family Foundation</td>
</tr>
<tr>
<td>Cochrane Collaboration</td>
<td>WebMD</td>
</tr>
<tr>
<td>Dynamed</td>
<td></td>
</tr>
<tr>
<td>Emedicine</td>
<td></td>
</tr>
<tr>
<td>UpToDate</td>
<td></td>
</tr>
<tr>
<td>Meta-analyses in major journals</td>
<td></td>
</tr>
</tbody>
</table>

If you wish to use a site not listed as approved, you may submit a request to the NBOME staff liaison or the Coordinator, who will make a determination as to whether the site is acceptable for use.

**Information to provide**

1) The URL (complete link) for the site and content
   **When using UpToDate, please provide the article title and date of the last update**
2) Saved PDF file or printout
3) Date the content was last reviewed (often it will be listed on the site)
4) Date the content was accessed
For items regarding the diagnosis and treatment of somatic dysfunction, only texts designated on the current OPP/OMM reference listing are acceptable.

Please refer to the listing below of primary and secondary OPP/OMM references.

### PRIMARY Reference Texts For OMM Items

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Edition</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>DiGiovanna and Schiowitz</td>
<td>Osteopathic Approach to Diagnosis and Treatment</td>
<td>3rd edition</td>
<td>2004</td>
</tr>
</tbody>
</table>

### SECONDARY Reference Texts For OMM Items

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Edition</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danto, J.B.</td>
<td>Normalization of Muscle Function</td>
<td>1st edition</td>
<td>2005</td>
</tr>
<tr>
<td>Steele, K., Comeaux, Z., Lempley</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnston, Friedman,</td>
<td>Functional Methods</td>
<td>2nd edition</td>
<td>2005</td>
</tr>
<tr>
<td>Eland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kimberly, P., Halma, K.</td>
<td>Outline of Osteopathic Manipulative Procedures:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2008</td>
<td></td>
</tr>
<tr>
<td>Magoun, H. I.</td>
<td>Osteopathy in the Cranial Field</td>
<td>3rd edition</td>
<td>1976</td>
</tr>
<tr>
<td></td>
<td>Consult</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rennie, P., Glover, P.</td>
<td>Counterstrain and Exercise (CD/Text)</td>
<td>1st edition</td>
<td>2004</td>
</tr>
<tr>
<td></td>
<td>Adolescents: A Practical Handbook</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Constructing Clinical Scenarios (8.0)

All items must refer to a realistic clinical scenario. This should give enough information to rule in or out all of the answer choices, and to provide a complete but not excessive description of the patient's condition and history. A good example of a clinical scenario with an appropriate amount of information is:

A 45-year-old asymptomatic male presents for a health maintenance examination. He has a 25 pack-year history of cigarette smoking and he exercises regularly. Surgical history is positive for an open cholecystectomy ten years ago and a right inguinal herniorrhaphy 20 years ago. Physical examination reveals a slightly deformed left clavicle. The lungs have occasional rhonchi on forced exhalation. Routine blood work is within normal limits, and chest radiograph reveals some pleural thickening.

This same scenario is presented below with too much detail:

A 45-year-old male presents for a health maintenance examination. He has no complaints, no health problems, and is on no medications. He has a 25 pack-year history of cigarette smoking, exercises three days a week on a treadmill at home, and drinks a glass of red wine with dinner nightly. Presently he is employed as an engineer, but he worked as an auto mechanic in his early 20’s. Surgical history is positive for an open cholecystectomy ten years ago and a right inguinal herniorrhaphy 20 years ago. Physical examination reveals a slightly deformed left clavicle and surgical scars over the right upper and lower quadrants of the abdomen. The lungs have occasional rhonchi on forced exhalation. The remainder of the physical examination is essentially normal. A complete metabolic profile, complete blood count, and urinalysis are normal. Chest radiograph reveals some pleural thickening.

Please note, however, that in the case of S-Type item sets (multiple questions based on a single clinical scenario), longer case histories are permissible, within reason, as subsequent interrogatories will draw from the original stem.

Finally, an example of this scenario with an insufficient amount of information follows:

A 45-year-old male smoker presents for a health maintenance examination. Presently he is employed as an engineer, but he worked as an auto mechanic in his early 20’s. Physical examination and blood work are within normal limits with the exception of rhonchi and a slightly deformed left clavicle. Chest radiograph reveals some pleural thickening.

Level 1 exam items emphasize basic understanding of scientific mechanisms in health and disease as they are related to the practice of osteopathic medicine. The authors of these items are often basic scientists; however, even these items must be in the context of a clinical scenario. Therefore it is recommended that the author collaborate with a clinician to ensure that the clinical scenario is complete, accurate, and realistic.
**Interrogatories (9.0)**

Consider beginning with the end in mind; that is, start with the interrogatory. What is it that a candidate should demonstrate that he or she knows or knows how to do?

Sample interrogatories for each physician task are listed below. Associated answer choice options are also provided. (These lists are in no way intended to be comprehensive.) Please note that it is best to use only ONE of the option types for each question. That is, don’t use medications, vaccinations, and dietary alterations as answer choices in a single item—use five different types of medications for one item, five different vaccinations for another item, etc., if at all possible.

**Health Promotion & Disease Prevention (9.1)**

- The most likely factor responsible for this outbreak is
  - Options: environmental factors, toxins
- A risk factor that could have contributed to this scenario is
  - Options: risk factors
- The predictive value of this study represents
  - Options: possible significance
- If untreated, this patient is likely to develop
  - Options: sequelae
- This patient’s disorder could have been prevented by
  - Options: prophylactic medications, vaccinations, dietary alterations, lifestyle alterations, dietary supplements
- The most appropriate action to prevent recurrence/disability/morbidity/mortality is
  - Options: medications, vaccinations, dietary alterations, lifestyle alterations, exercise regimens
- The most appropriate action for hospital staff to prevent the spread of this patient's disease is
  - Options: protective pieces of equipment, types of room/staff precautions
- The most appropriate immunization recommendation is
  - Options: vaccinations, ages
- The most appropriate criteria for screening in this situation is
  - Options: ages, testing frequencies, screening tests
- The most appropriate lifestyle change to recommend is
  - Options: dietary alterations, exercise regimens, dietary supplements

**History & Physical Examination (9.2)**

- The most likely diagnosis is
  - Options: disorders, diseases
- Which of the following additional symptoms would you expect to be present?
  - Options: symptoms
- The most likely structural finding in this patient is
  - Options: structural findings
Diagnostic Technologies (9.3)

- Which of the following is most likely to confirm the diagnosis?
  - Options: diagnostic tests
- The most appropriate diagnostic test is
  - Options: diagnostic tests
- The next step in this patient’s work-up is
  - Options: diagnostic tests
- Which of the following tests would have predicted these findings?
  - Options: diagnostic tests
- Which of the following findings is expected?
  - Options: serum levels, microscopic fluid findings, muscle/joint tissue findings, DNA analysis results, pathology results
- The most likely finding on [test name] is
  - Options: test findings
- Based on [test findings], the most likely diagnosis is
  - Options: diagnoses

Management (9.4)

- The most appropriate management for this patient is
  - Options: drugs, OMT techniques, surgical techniques
- The most appropriate initial step in this patient’s management is
  - Options: management steps
- Which of the following should be administered?
  - Options: drugs, vitamins, amino acids, enzymes, hormones
- The initial set-up to perform a muscle energy technique is
  - Options: patient positions

Scientific Understanding of Health & Disease Mechanisms (9.5)

- The most likely etiology/cause is
  - Options: bacteria, toxins, medications, hemodynamic mechanisms, viruses, metabolic defects, pathogens
- The most likely structure affected is
  - Options: nerves, muscles, vessels
- A defect is most likely to be present in which of the following?
  - Options: structures, processes
- Which of the following is defective/deficient/nonfunctioning?
  - Options: enzymes, feedback mechanisms, endocrine structures, dietary elements, vitamins
- Which of the following is involved?
  - Options: enzymes, hormones, cells, neurotransmitters, molecules, spinal segments
- The most likely mechanism/explanation involved is
  - Options: disease mechanisms, pharmacologic mechanisms
- Which of the following was most crucial in the development of this patient’s disorder?
  - Options: factors
- Which of the following is most likely to be increased/decreased?
  - Options: laboratory levels
- The most likely finding on autopsy is
  - Options: autopsy findings
• Given the pedigree, what is the likelihood that the next child (gender) will have the disease?
  o Options: percentages
• The primary action of this prescribed drug in this patient is
  o Options: drug actions
• Somatovisceral findings are most likely present at which of the following spinal levels?
  o Options: individual spinal segments/ranges

**Health Care Delivery Issues (9.6)**

• To which of the following community service agencies should this patient be referred?
  o Options: agencies
• The principle involved in determining the most appropriate action is
  o Options: ethical principles
• The most appropriate response for the physician to give is
  o Options: quotes
• The most ethical course of action is
  o Options: courses of action
• The most appropriate question to ask the patient is
  o Options: questions
• The physician may be found liable due to
  o Options: liability principles
• Which of the following laws/acts/principles has been violated in this case?
  o Options: laws, acts, principles

**Clinical Scenario Patterns (10.0)**

Once it has been decided what the candidate is to demonstrate that he or she knows or knows how to do, the item-writer may find it helpful to consider the use of item patterning. Using the parameters determined earlier and the framed interrogatory, a clinical scenario can be created. Consider selecting one of the following patterns and fill in the blanks with the information that pertains to the scenario:

**Pattern #1**
A [patient description: age, gender] presents with [symptoms, including time frame]. History reveals [historical findings]. Physical examination reveals [findings]. Diagnostic studies reveal [findings].

**NOTE:** In a given scenario, the item-writer may wish to include only some parts of this information. For example, a patient in the emergency department may not have a history available. Results from diagnostic studies may not always be available.

**Pattern #2**
A [patient description: age, gender] presents to the [site of presentation] with [signs, symptoms, findings] in [describe the patient’s condition]. History reveals [pertinent background information].
**Pattern #3**
A [patient description: age, gender] presents with the complaint of [somatic/visceral symptoms]. History reveals [pertinent background information]. Structural examination reveals [structural and palpatory findings].

**Pattern #4**
[Time period] after a [event, trip or meal] a [patient or group description] developed [symptoms].

Other information may be included in clinical scenarios, such as:
- Immunization status/history
- Family history
- Risk factors
- Allergy history
- Initial actions and subsequent findings

**Item Format Basics (11.0)**
Three basic principles to follow in writing items are:

1. The clinical scenario in each item should be in the following order: presentation (complaint-why a patient is seeking care); history (including duration of signs & symptoms); physical findings; results of diagnostic studies; initial treatment; subsequent findings. A stem may contain only some of these components, but it should usually have the first three.

2. An item stem should contain enough information so that a candidate can answer the question before looking at the options. No additional background information should be given in any of the answer choices. Following these guidelines will result in concise answer choices. Sufficient information should be included in the item scenario to rule out each of the incorrect answer choices (distractors) and rule in the answer as the best option. Distractors should always be “attractive to the uninformed”—i.e., reasonable-sounding options.

3. In a multiple-choice item, incorrect responses are often not entirely wrong, and they don’t have to be. This is why “most likely” and “most appropriate” are frequently used in examination items. The answer choices can be placed on a continuum, where one is the most correct, although the others may not be 100% wrong.

**Item Flaws (12.0)**
Common types of faulty items are:

**Pseudocases (12.1)**
These are items where the clinical scenario relates to the question, but the question can be answered without referring to the scenario.
Example: A 52-year-old male presents to the office with a one-week history of flank pain and hematuria. Past medical history is unremarkable. Physical examination reveals a left-sided abdominal mass. CT scan of the abdomen reveals a left renal mass. The greatest risk factor for renal cell carcinoma is

(A) diabetes  
(B) female gender  
(C) hyperlipidemia  
(D) low body mass index  
(E) smoking

Unfocused Items (12.2)
These are items that start with a variation of “Which of the following is correct regarding [topic]?” These items are flawed for two reasons:

1. There is not enough information in the stem to answer the question without looking at the options.
2. The responses are disparate—they can’t be placed on a continuum, so the distractors must be 100% false. This either leads to a very obvious answer or confusion as to whether each option is 100% true or 100% false.

Items With Flaws That Favor Testwise Candidates (12.3)
These are items that contain clues pointing to the correct answer that are unrelated to content. Some examples of these clues typically seen in the distractors are:

- **Grammatical cues (12.3.1)** Some or all of the distractors don’t follow grammatically from the stem, whereas the correct answer does.
- **Logical cues (12.3.2)** A subset of the answer choices covers all possible options, with the remaining answer choices unrelated.
- **Word repeats (12.3.3)** A term or form of a term that is used in the stem matches a term in the correct answer choice.
- **Long answer (12.3.4)** The correct answer is longer or more detailed than the distractors.
- **Absolute terms (12.3.5)** “Always,” “never,” or other absolute terms used in an answer choice nearly always indicate an incorrect answer.
- **Convergence strategy (12.3.6)** In answer choices with multiple elements, the one with the most elements in common with the others is often the correct answer.

Items With Flaws Due To Irrelevant Difficulty (12.4)
These are items that can be confusing due to flawed item format or other factors unrelated to content. Some examples of these factors are:

- **Negative items (12.4.1)** The use of “EXCEPT,” “NOT,” or “LEAST” in the stem creates a situation where the candidate has to switch his or her thinking to the opposite of what is expected.
- **Tricky or unnecessarily complicated stems (12.4.2)** Do not use information that distracts ("red herrings"—information that can confuse or mislead the candidate). Although this does play a role in actual patient care it may not be appropriate for testing situations. On the other hand, superfluous information ("window dressing") is acceptable to the extent that it reflects reality. Not all information has to be directly relevant to the answer. Some might rule out other choices or provide a more detailed patient description.

- **Percentages or statistical information (12.4.3)** In general, it is not appropriate to test on specific percentages or figures. Testing a candidate's understanding of how to apply this information is appropriate, but testing a candidate's knowledge of actual percentages or statistics is not appropriate.

- **Vague frequency terms (12.4.4)** Terms such as “rarely” and “usually” are imprecise, making their intention and meaning unclear when used in answer choices.

- **Regional-specific terms or concepts (12.4.5)** Given that COMLEX-USA is used as a national testing tool, it is important that all terms and concepts included are universally understood.

**Items With Teaching in the Stem (12.5)**
These are items that give the candidate information regarding the topic being tested that extends beyond the specific clinical scenario. This is inappropriate and may even offer an inadvertent clue other items that appears elsewhere in the exam.

**Item Quality Checklist (13.0)**
Each item that is submitted to the NBOME for consideration is subjected to the following series of questions/tests. Sometimes items can be refined without being returned to the author prior to review; other times comments are added to flawed items so that the review committee can refine them; and still other times items are returned to the author for reworking prior to being presented at review. Of course it is best to have items submitted that pass all of the following tests, so that the author’s original intention can remain intact. We ask that all authors develop the habit of applying these tests to each of their items:

- Is the item researched and referenced using a current, standard (universally used/available) source?
- Is the stem presented as a clinical scenario that is realistic and provides sufficient supporting information needed to answer the question?
- Is it easy to identify what knowledge and skill the item is testing?
- Does the item reflect higher-order thinking rather than recall of facts?
- Does the item have at least five answer choices that are:
  - Similar in length and style?
  - Actual entities?
  - Attractive to the uninformed?
  - Unique/distinct from each other?
  - Not mutually exclusive or opposite of another choice?
  - Addressed in the stem (information in the stem should rule in or out each of the choices)?
• Are any laboratory values used in the item compatible with those on the NBOME reference value listing?
• If a visual is used, is it:
  o Essential to answer the question?
  o The best possible quality?
  o The proper size?
• Have the core competencies and the best available medical evidence been considered?

Constructing Matching Items (14.0)
Matching items may be useful when there is a great deal of subject matter to be covered but the number of exam items allocated to that topic area is limited. In this case, the candidates’ knowledge can be assessed by providing a wider scope of plausible choices or options.

Matching items are of two basic types:
• B-Type, in which there are five lettered options, which may each be used once, more than once, or not at all.
• X-Type, in which there are more than five lettered options, which may each be used once, more than once, or not at all.

The steps inherent in creating a set of matching items are:
1. Decide on the central subject, concept or issue.
2. Develop the list of options from which the candidate may choose the correct answers.
3. Write an interrogatory or lead-in. (What is it that the candidate is to do with the list of options?)
4. Construct the item stems (scenarios) to which the options are to be matched.

Many of the basic guidelines outlined for the development of the stem and the option list for an A-Type exam item should be taken into consideration when writing a matching exam item. For example, the stem must be succinct without much extraneous information due to the amount of reading the candidate must do. However, depending on the nature of the exam item, matching sets can permit a wider range of subject matter to be included in the options. For example, depending on the stems, an antiviral and an antibiotic may both be good choices to include in a single matching set.

The interrogatory must be well-worded and to the point, making clear what is expected of the candidate in order to answer the question. Using the information above, a typical matching set is developed below:

Now, suppose the assignment is:

**Patient presentation:** rectal masses and swelling
**Physician Task:** management
The assignment to create a B-Type set is addressed as follows:

- **Content:** Surgical management for rectal masses and swelling
- **Options:** List of possible surgical alternatives
- **Interrogatory:** For each clinical presentation determine the best surgical management.
- **Stems:** Clinical presentations that warrant a surgical procedure

The matching (B-Type) set is written as follows:

<table>
<thead>
<tr>
<th>Interrogatory:</th>
<th>For each numbered item (<strong>patient presentation</strong>) select the one heading (<strong>surgical procedure</strong>) that would best address the patient’s clinical condition. Each lettered heading may be selected once, more than once, or not at all.</th>
</tr>
</thead>
</table>
| Options:      | (A) abdominoperineal resection with end colostomy  
(A) laser vaporization ablation  
(C) left hemicolectomy with primary anastomosis  
(D) sigmoid colon resection with diverting colostomy  
(E) total colectomy with ileostomy |
| Stems (Scenarios): | 1. A 32-year-old female presents with a pelvic abscess due to sigmoid diverticular disease.  
2. A 65-year-old male presents with adenocarcinoma 4 cm from the anal verge.  
3. A 54-year-old female presents with a 3-cm villous adenoma extending from the anus.  
4. An 89-year-old male presents with a 3-cm villous adenoma extending from the anus.  
5. A 75-year-old female presents with a 3-cm well-differentiated adenocarcinoma of the sigmoid colon. |

In the event that an X-Type set is desired, additional choices or options would be created, increasing the choices from which to select the correct answers.

After creating a B-Type or X-Type set, the item-writer should carefully review it and apply the quality checklist provided in this guide (pages 19-20) before submitting the set.

**Constructing Distractors (15.0)**

Experience has shown that effective item-writing requires the use of well-constructed distractors. As has been demonstrated earlier in this guide, test-savvy candidates may take advantage of unwitting clues provided by the item-writer. In order to have a more effective distractor list, the item-writer should:
1. Strive to have distractors of equal or similar length
2. Create distractors that match grammatically with the stem
3. Avoid distractors that are “odd man out” (e.g., four diagnostic procedures and one therapeutic procedure)
4. Not use opposites as choices
5. Not have overlapping numerical figures or values
6. Avoid absolutes (e.g., “always,” “never”)
7. Avoid negatives (including “all of the following EXCEPT” stems)
8. Avoid “all of the above” or “none of the above” as distractors
9. Avoid abbreviations and acronyms where possible

**OPP/OMM Integration (16.0)**

Osteopathic principles are incorporated into all facets of COMLEX-USA examinations. Some items are intended to be “pure” OPP or OMM items, such as:

1. **OPP diagnosis items**—items that involve the interpretation of neural, vascular, and/or lymphatic findings as part of somatic dysfunction findings
2. **OPP treatment items**—items that involve treatment based on neural, vascular, and/or lymphatic findings as part of somatic dysfunction findings
3. **OMM diagnosis items**—items that focus on the diagnosis of somatic dysfunction
4. **OMM treatment items**—items that focus on the treatment of somatic dysfunction

These “pure” OPP/OMM items comprise a subset of the exam items used in the examination. Another subset is that of “OPP/OMM integrated” items, which are utilized throughout the examination.

These items are of two types:

1. Those that integrate somatic dysfunction findings into a non-OPP/OMM case

   **Example:** A 74-year-old male reports the new onset of low back pain. Physical examination reveals an L5 spinous process tender point and tenderness over the sacrum. Lower extremity muscle strength and deep tendon reflexes are normal. Which of the following studies would be most appropriate in the evaluation of this patient’s pain?

2. Those that integrate OPP or OMM distractors into a non-OPP case, where the correct answer does not involve OPP or OMM

   **Example:**
   
   
   (A) cortisone injection of the intraarticular space  
   (B) counterstrain of the medial meniscus tender point  
   (C) lidocaine injection of the gastrocnemius muscle  
   (D) muscle energy treatment of the tibia  
   (E) quadriceps strengthening
For items regarding the diagnosis and treatment of somatic dysfunction, only texts designated on the current OPP/OMM reference listing are acceptable (see Researching and Referencing Content for an Item on pages 13-15). These textbooks contain currently accepted terminology used by the osteopathic profession to describe spinal, cranial and general joint mechanics, as well as the terminology used with various testable OMM techniques.

All OPP/OMM specialists who are given an assignment will also be given the “Testable Somatic Dysfunction” listing. This listing details the content that has been determined to be universally taught in osteopathic schools in the U.S. Content on COMLEX-USA OPP or OMM items should be limited to the topics covered in this list. In addition, terminology used must be consistent with that defined in the current glossary of the Educational Council on Osteopathic Principles (ECOP).

Using Bloom’s Taxonomy (17.0)
Dr. Benjamin Bloom, an educator in the mid-to-late 1900’s, proposed a “taxonomy” to define the levels of learning that an individual can develop and display. Some authorities have proposed that the taxonomy be utilized as another dimension in test design by suggesting that a certain percentage of items be of certain taxonomy levels as one of the dimensions of the examination. From a practical perspective it may be helpful for item-writers to consider using Bloom’s taxonomy to advantage by using the taxonomy to assist in developing exam items that require the candidate to use higher levels of thought other than rote recall. It must be remembered that by raising the level of the thought process required to answer an exam item, the difficulty of the item is not necessarily increased. For example, a minutia question of very high difficulty could be asked that is considered simple recall, yet a more practical application exam item might only prove to be of moderate difficulty.

Therefore, Bloom’s Taxonomy can be used to some advantage as follows:

- Consider making the candidate think harder and use higher-level thought processes
- Challenge candidates to look at the same or similar subject material in different ways

The taxonomy as promulgated by Dr. Bloom is:

**Level 1 - Knowledge & Memory** (pure recall of data)
- Defining
- Recalling
- Identifying

**Level 2-CE - Comprehension** (understanding of meaning)
- Determining
- Explaining
- Inferring

**Level 3 - Application** (use of information)
- Applying
- Choosing
- Organizing
Level 4 - Analysis (understanding of organizational structure)
- Comparing
- Contrasting
- Distinguishing

Level 5 - Synthesis (relation of knowledge from several areas)
- Combining
- Deriving
- Specifying

Level 6 - Pure Reasoning (making judgments about the value of ideas and materials)
- Assessing
- Arguing
- Considering

Here is an example of a recall item:

Which of the following is the definitive test for gestational diabetes?
(A) fasting blood glucose level
(B) glycosylated hemoglobin level
(C) serum acetone and ketone levels
(D) three-hour glucose tolerance test
(E) two-hour postprandial glucose level

This recall item can be converted to an application item by altering the stem and asking the candidate to recognize an abnormal result and make an appropriate alteration in patient management:

A 33-year-old female at 30 weeks’ gestation is given a three-hour glucose tolerance test with the following results:

- Fasting: 90 mg/dL
- 1 Hour: 180 mg/dL
- 2 Hour: 170 mg/dL
- 3 Hour: 120 mg/dL

The most appropriate recommendation for this patient is to
(A) be assured that the results are within normal limits
(B) begin insulin therapy immediately
(C) have the test repeated after delivery
(D) have the test repeated with a 200-g glucose load
(E) make an appointment for diet counseling
Here is an example of an application item:

A 30-year-old male with symptomatic hemorrhoids presents with rectal bleeding. His father was diagnosed with colon cancer several years earlier. The best management approach for this patient is to treat his hemorrhoids and

(A) order a barium enema
(B) order a colonoscopy
(C) order CEA and CA 19-9 levels
(D) order serial occult blood testing for three days
(E) reassure the patient

By revising the stem, the candidate is required to deal with the practical situation of advising a patient by better understanding the data presented along with the consideration of the potential of genetic factors affecting the occurrence of disease. The situation must be carefully analyzed before answering the exam item.

A 30-year-old male is concerned about his strong family history of colon cancer. The earliest age of a first-degree relative to have colon cancer was age 50. The patient is currently asymptomatic. The best course of action in this case is to order

(A) a rectal exam for occult blood and colonoscopy before age 40
(B) a screening barium enema
(C) a screening CA 19-9 level, with a follow-up exam in one year
(D) genetic counseling
(E) immediate colonoscopy

Thus, keeping in mind the basic tenets suggested by Bloom in his taxonomy may help the exam item-writer write better-caliber exam items with greater clinical applicability. This taxonomy is applicable to all types of exam items.

Evidence-Based Medicine (EBM) Principles in Item-Writing (18.0)

Over 20 years ago Fletcher, Guyatt, and Sackett suggested Levels of Evidence for ranking the validity and strength of research. This led them to propose “grades of recommendation” to facilitate the application of medical research to the improvement of healthcare. Since that time, numerous texts, articles, and presentations at both national and international meetings have supported the concept of applying the principles of determining the value of research to the practicality of adopting the best medical practices in the interest of patient care. The term “evidence-based medicine” (EBM) first appeared in the literature in 1992.

From a testing perspective, it is incumbent on the exam item-writer to be familiar with the principles of evidence-based medicine in both referencing and generating an exam item. The exam item author should question the source and strength of the information used as the basis for the exam item, and should also consider another genre of exam items that challenge the examinees’ knowledge of the principles and applicability of the
evidence-based approach to medical practice. Ultimately, the question of best available evidence may come down to professional judgment. However, using the most objective and verifiable information should place the item-writer or reviewer in good stead.

In the practice of modern medicine, future physicians will be called upon to make clinical judgments and alterations in the way they practice by being able to effectively weigh the evidence presented to them through a wide variety of informatics. Item-writers for the NBOME, in moving forward to meet the demands of competency-based assessment, will need to evaluate this skill in candidates being tested for osteopathic medical licensure. This section of the Item-Writing Guide will offer the item-writer some guidance in referencing and creating exam items using these principles to help meet this increasingly important aspect in the practice of osteopathic medicine.

**Definitions (18.1)**

There are currently several systems of defining evidence. The U.S. Preventive Services Task Force recommends the following rubric for ranking evidence about the effectiveness of treatments or screening:

<table>
<thead>
<tr>
<th>Level</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Evidence obtained from at least one properly randomized controlled study</td>
</tr>
<tr>
<td>II-1</td>
<td>Evidence obtained from well-designed controlled trials without randomization</td>
</tr>
<tr>
<td>II-2</td>
<td>Evidence obtained from well-designed cohort or case-controlled analytic studies, preferably from more than one center or research group</td>
</tr>
<tr>
<td>II-3</td>
<td>Evidence obtained from multiple timed series with or without intervention. (Dramatic results from uncontrolled studies may also be considered here.)</td>
</tr>
<tr>
<td>III</td>
<td>Opinions of respected authorities, based on clinical experience, and descriptive studies or reports of expert committees.</td>
</tr>
</tbody>
</table>

The Centre for Evidence-Based Medicine at Oxford University suggests the following classification:

<table>
<thead>
<tr>
<th>Level</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>Systematic review of homogenous randomized controlled trials</td>
</tr>
<tr>
<td>1b</td>
<td>Individual randomized controlled trial</td>
</tr>
<tr>
<td>1c</td>
<td>All-or-none case series</td>
</tr>
<tr>
<td>2a</td>
<td>Systematic review of homogenous cohort studies</td>
</tr>
<tr>
<td>2b</td>
<td>Individual cohort study or low-quality randomized controlled trial</td>
</tr>
<tr>
<td>2c</td>
<td>Outcomes research; ecological studies</td>
</tr>
<tr>
<td>3a</td>
<td>Systematic review of homogenous case-control studies</td>
</tr>
<tr>
<td>3b</td>
<td>Individual case-control study</td>
</tr>
<tr>
<td>4</td>
<td>Case-series (and poor quality cohort and case-control studies)</td>
</tr>
<tr>
<td>5</td>
<td>Expert opinion without explicit critical appraisal, or based on physiology, bench research or “first principles”</td>
</tr>
</tbody>
</table>

*From the Centre for Evidence-Based Medicine, Oxford University (varies slightly for prognosis, diagnosis, and therapy & prevention)
In determining the application of levels of evidence to the practice of medicine, it is important to be able to weigh the validity of the study and to decide on the appropriate balance of risk versus benefit. To assist in making judgments, the U.S. Preventive Services Task Force proposes the following categories of recommendations for the practitioner to consider as he or she weighs the evidence presented in a study. The classification is as follows:

<table>
<thead>
<tr>
<th>Level</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Good scientific evidence suggests that the benefit of the clinical service or study substantially outweighs the potential risks. Clinicians should discuss this service or study with eligible patients.</td>
</tr>
<tr>
<td>B</td>
<td>At least fair scientific evidence suggests that the benefit of the clinical service or study outweighs the potential risks. Clinicians should discuss this service or study with eligible patients.</td>
</tr>
<tr>
<td>C</td>
<td>At least fair scientific evidence suggests that there are benefits of the clinical service or study, but the balance between benefits and risks is too close for making general recommendations. Clinicians need not offer this service or study except for individual considerations.</td>
</tr>
<tr>
<td>D</td>
<td>At least fair scientific evidence suggests that the risks of the clinical service or study outweigh the potential benefits. Clinicians should not routinely recommend this service or study to asymptomatic patients.</td>
</tr>
<tr>
<td>E</td>
<td>Scientific evidence is lacking, the service or study is of poor quality, or it offers conflicting results such that the risk benefit cannot be adequately assessed. Clinicians should help patients understand the uncertainty surrounding such a service or study, and advise patients of the potential for harm, if that is appropriate.</td>
</tr>
</tbody>
</table>

**General Principles (18.2)**

Whenever a study or article is reviewed, it is incumbent on the reviewer to seek out signs of poor study construction or design, insufficient or inappropriate sampling, and bias of many different dimensions.

In considering the use of evidence-based medicine to determine the best medical practices, it may be helpful to acknowledge that medicine is part science and part art. In that regard, the following may need to be considered:

1. Conducting randomized studies may, in some cases, be unethical, and may have to be reduced to studies of observation without controls or randomizations.
2. Some areas and populations in medicine have been understudied and may not have large numbers of studies or representatives in a cohort. The levels of evidence available in these populations may, in fact, be sparse.
3. The “gold standard” studies may be too expensive to conduct in the classical scientific sense.
4. Studies published in the medical literature may not be representative of all of the studies performed in a particular domain or on certain topic.
5. Some studies may have to be abandoned as results prove either too beneficial or too detrimental for the study to continue.
It is also helpful to develop a vocabulary to assist in reviewing and analyzing the data that is presented in a study. A more complete glossary of useful terms is available on the Centre for Evidence-Based Medicine (CEBM) Web site. Some commonly used terms include:

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNT</td>
<td>Number needed to treat</td>
</tr>
<tr>
<td>NNH</td>
<td>Number needed to harm</td>
</tr>
<tr>
<td>RRR</td>
<td>Relative risk reduction</td>
</tr>
<tr>
<td>ARR</td>
<td>Absolute risk reduction</td>
</tr>
<tr>
<td>RR</td>
<td>Relative risk</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>Ability to identify the population with a condition that have a positive result</td>
</tr>
<tr>
<td>Specificity</td>
<td>Ability to identify the population without the condition that have a negative result</td>
</tr>
<tr>
<td>Meta-analysis</td>
<td>Scientific combination of psychometric and statistical results from many institutions</td>
</tr>
<tr>
<td>PPV</td>
<td>Positive predictive value</td>
</tr>
<tr>
<td>NPV</td>
<td>Negative predictive value</td>
</tr>
</tbody>
</table>

Knowing the advantages and disadvantages of study designs can be helpful in analyzing studies and references, and it is also useful in creating exam items by comparing the merits and take-home message of any study or group of studies aimed at a particular population or sub-group of a population. The following is a list of commonly used studies that offer advantages and disadvantages that can be explored by an item-writer:

1. Case-control study
2. Cross-sectional survey
3. Cohort study
4. Randomized controlled study
5. Crossover design
6. Anecdotal or observational study

The key for the physician in using evidence-based medicine rests in formulating a clinical question to address improved patient care and minimized medical errors.

A few approaches to formulating and answering the question may be obtained by crossing the level of evidence with any or all of the following:

1. Therapy, prevention, etiology and/or harm
2. Prognosis
3. Diagnosis
4. Differential diagnosis, symptom prevalence
5. Economic and decision analysis

Another approach to defining an answerable clinical question includes the following:

1. Patient and/or problem
2. Intervention
3. Comparison of interventions
4. Outcome
When putting these principles to work, if the reviewer is to review a study or write exam items, the framework for formulating questions as outlined below may prove helpful:

1. Are the results of the study valid?
2. Was the study taken with a representative spectrum of patients like those that would be seen in the practice of medicine applicable to the patient population being cared for?
3. What type of study was conducted?
4. What is the best course, diagnostic tool, drug, treatment combination, etc?
5. Where is the information to be found (i.e., what methods were used to collect evidence)?
6. Were the index test and the reference standard carried out on all patients?
7. Was there an independent blind comparison conducted?
8. What were the results?
9. What was the measure used?
10. What does the interpretation of the data actually mean to the population under the physician’s care?
11. Were the results sufficient to support the use of the methods employed, and can the methods and results likely be replicated?
12. Were the conclusions fairly drawn, and is little or no bias reflected in the conclusion as determined by the presented evidence?
13. What level of evidence is identifiable?
14. What recommendation can be made based on the evidence presented in the study?

Writing the Exam Item (18.3)

Taking the information presented, let us attempt to write a case series that deals with analyzing and applying the clinical question to practice. The series contains elements of comprehension, analysis, and application.

<table>
<thead>
<tr>
<th></th>
<th>Standard Reference +</th>
<th>Reference</th>
<th>Standard Reference -</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Index Test +</strong></td>
<td>240</td>
<td>150</td>
<td></td>
<td>390</td>
</tr>
<tr>
<td><strong>Index Test -</strong></td>
<td>10</td>
<td>600</td>
<td></td>
<td>610</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>250</td>
<td>750</td>
<td></td>
<td>1000</td>
</tr>
</tbody>
</table>

The study took three years to complete. There were no deaths or injuries during the study, and permission to conduct the study was obtained from the patient or the guardian depending on their clinical condition at the time of entry into the study with full disclosure of the nature of the study.
1. The accuracy of the study would be best determined by calculating which one of the following parameters?
   (A) negative predictive value (NPV)
   (B) number needed to harm (NNH)
   (C) number needed to treat (NNT)
   (D) positive predictive value (PPV)
   (E) sensitivity

2. The study was ethical to conduct because
   (A) it harmed no one
   (B) it was a non-invasive study
   (C) no deaths occurred during the study
   (D) permission was obtained with full disclosure
   (E) the level of evidence gleaned from the study was high

3. This study would be classified as a/an
   (A) anecdotal report study
   (B) case-controlled study
   (C) cohort study
   (D) cross-sectional study
   (E) randomized control study

[Answers: (E), (D), (B)]

Other areas that could have been explored in this scenario include, but are not limited to:
1. Would you adopt use of the test based on the results?
2. Would you change the manner in which you diagnose patients with dementia?
3. How would you change the medical systems you employ based on this study?

In the sample provided above, it should also be observed that Bloom’s taxonomy is used to advantage in this case by demonstrating the incorporation of comprehension, analysis, and application into the item series, thus raising the level of the thought processes used by the candidate.

Writing exam items in the arena of evidence-based medicine (EBM) can be both challenging and rewarding. This section has been presented to heighten awareness of the scientific side of medical practice that needs to be considered when reviewing the medical literature and other sources of information such as the Internet. The principles described in this section may also help the item-writer to be more critical of item content and to write better, more innovative types of exam items.

References:
www.cebm.net
www.minervation.com/cebm/docs/levels.html
Item-Writing Strategies for Medical Educators (NBOME)
Core Competencies (19.0)
The NBOME has recently adopted a document outlining seven core competencies to be demonstrated by all candidates applying for medical licensure. The goal of the COMLEX-USA examination series is to measure all of these competencies in the candidates who take our examinations. Some of these are measured through the performance examinations, which go through a separate development process from that for which this guide is intended. However, many of the competencies are addressed through the computerized exams (CE); therefore, when writing items please keep in mind that we are committed to testing the following seven competencies:

1. Osteopathic philosophy and manipulative medicine
2. Medical knowledge
3. Patient care
4. Interpersonal and communication skills
5. Professionalism
6. Practice-based learning and improvement
7. Systems-based practice

Writing on these competencies requires an outcomes-based perspective, which is often not easy to do, but is well worth the effort. The reader is referred to the "The Seven Osteopathic Competencies: Implications for Medical Licensure Testing and Osteopathic Medical Practice" available and downloadable from the NBOME Web site, www.nbome.org.

Writing Multimedia and Advanced Type Exam Items (20.0)
Writing multimedia exam items is a difficult and challenging task. It requires that the item-writer use graphic materials, video, or audio in a way that allows the candidate to interact with the program, and in a way that is essential to the item. The graphics also must be of sufficient clarity to be identifiable by the candidate. In addition to appropriate content, the item-writer must ensure that the audio-visual materials meet certain specifications in order to be delivered in the CBT format.

Computer-based testing (CBT) permits a wide array of exam item types that would not otherwise be possible in a paper-and-pencil format. However, there are limitations with computer-based testing, such as memory consumption, exam item functionality, contrast and brightness, and screen definition of the presented material. (Disabled or impaired candidates who cannot hear or see also present challenges in testing what is determined to be required knowledge and/or skill demonstration).
Media Types (20.1)
The following list provides examples of the multimedia types that may be considered by exam item-writers for submission:

1. Video clips
   a. Congenital malformations
   b. Range of motion testing
   c. Gait
   d. Neurologic disorders
   e. Patient positioning for treatment
   f. Pediatric milestones
   g. Posture
   h. Psychiatric disorders
   i. Somatic dysfunctions
   j. Surgical or office procedures

2. Audio clips
   k. Bowel sounds
   l. Breath sounds
   m. Doppler studies and sounds
   n. Heart sounds
   o. Phonocardiograms

3. Interactive audio-visual program
4. Animation
5. Tabulated image or exhibits (rather than text)
   p. Serial image cuts
6. Activated or action graphs
7. Pointer-indicated selection (hot spot)

In writing items that utilize “still” graphics, the following graphic material can be selected for importation into a exam item. It is important to recall that the graphic used should be essential to the item’s scenario and not simply a tack-on for embellishment.

1. Clear photos of gross pathology or clinical situations
2. CT and MRI studies (with and without contrast)
3. Diagrams/Tables
4. Doppler studies
5. ECG strips
6. Endoscopic images
7. Epidemiologic patterns
8. Gram stains
9. Graphs
10. Gross anatomic or surgical specimens
11. Laparoscopic images
12. Nuclear scans (static)
13. Pathology slides
14. Radiographs
15. Rhythm strips
16. Skin lesions, burns, or wounds
Following are two examples of acceptable “still” graphics:
**Guidelines of Media Use (20.2)**

1. All submissions need to be original (not published in any other source, including the Internet).
2. A graphic or audio/video clip should be essential to the exam item and not simply enhancement of the text.
3. When writing exam items with extensive graphics or videos, keep in mind that more time may be required by the candidate to view the graphics or videos, thus text may need to be minimized.

**Specifications of Submitted Media (20.3)**

1. **Specifications of Submitted Still Images**
   a. Out-of-focus visuals cannot be adjusted or used; they should not be submitted.
   b. Quality is compromised when the target source of a digital image is not an object (i.e., taking a digital image of a printed source such as a negative, a publication, or a print results in lower quality). Therefore, it is recommended to only use original images—those that are first-generation.
   c. The minimum acceptable resolution is 150 pixels/inch with a real image size of 4” x 4”. ECG’s or other tracings must be actual size.
   d. Images can be submitted in any of the following formats: jpg, gif, bmp, pdf, psd, pct, pict, wpg, tiff, png, raw.
   e. Images must not be embedded in a Word file. They should be submitted separately. (This is to ensure that the visual size does not go below 4” x 4” and that nothing is compromised by making the visual fit into the document.)
   f. Submitted digital photos should be in color, as compared to grayscale for most radiological studies.

2. **Specifications of Submitted Videos**
   a. Video clips should average 10-30 seconds in length.
   b. Videos can be submitted in any of the following formats as long as they are high quality: .avi, .mpg, mp4, mov.
   c. The minimum video size is 640 x 480 pixels (standard) or 720 x 480 pixels (widescreen).
   d. All video clips must be well lit; shooting against a bright background or casting of heavy shadows must be minimized or eliminated.
   e. The recommended frame rate is 30 frames per second (fps).
   f. Items that have been converted into “flash” programs are acceptable, but text for the items should be submitted separately, with the primary file clearly designated.
   g. The setting should be an examination room (or resemble an examination room), with no clutter or unnecessary items in the background.

3. **Specifications of Submitted Audio Clips**
   a. The preferred file format is .wav, but .wma, .ra, .mp3, and .mov will be accepted if they are of high quality.
   b. The volume should be consistent for the duration of the clip.
An example of an appropriately built exam item is:

A 50-year-old male is attacked by an irate fellow worker. The patient is assaulted with a knife that is 6 inches in length. The knife entry is just below the xiphoid, and the blade entry was observed to be parallel to the ground. Select from the following anatomical sections the most likely path of the knife blade.

(A) Exhibit 1

(B) Exhibit 2

(C) Exhibit 3

(D) Exhibit 4

(E) Exhibit 5
In this item, each square contains a thumbnail anatomic image from which the candidate is to select the most correct choice. Here (E), the transpyloric plane, is the correct choice. The candidate would be able to view a larger image of the thumbnail by clicking directly in the thumbnail.

Images can be used in the body of the stem or brought into view by activating an exhibit button. Location of the image or images on the screen should be designated, based on size, detail, composition, and format of the item. For example, “include with text and initial screen,” “include as exhibit with button,” or “include thumbnail choices with or without blow-up option.”

Interesting and challenging but clearly discernible cases are the goal for this type of item. In addition, the item-writer should seek to test for more than one area of knowledge or skill. For example, if an interactive item requires that the chest and neck be examined in order to hear the murmur of aortic stenosis, with that option (aortic stenosis) selected as the correct answer from a list of choices, this tests the candidate on all of the following: examination of the chest, location of the sound, recognition of the sound and radiation pattern, and determination of what, if anything, to do about the finding. That is a great deal of information from one exam item.

It is anticipated that, initially, only a small percentage of the new exam item types will be included in the examination. However, developing a bank of items of this type is essential for future testing in CBT format as testing moves toward a more competency-based assessment of candidates.

**Exam Development Process (21.0)**

The annual exam development process begins with an evaluation of the item bank, the outline, and the blueprint. This provides an overview of the topics that must be assigned to writers. The cycle for new items is shown below:
Authors are usually given several months from the time they receive their assignments until the finished items are due to be returned to the NBOME. A general rule of thumb is that the time required to write one item is 30 minutes. Assignments usually consist of 5 to 20 items per author, depending on the individual’s time constraints and interest. CME credit is given for item construction—at a category 1B level—30 minutes per item submitted.

Once items are received in the NBOME office, item format experts and grammarians review them for format and wording issues. Flaws that do not relate to medical content and that can be easily corrected are corrected at this stage. Items with more substantial problems are returned to the author with an explanation of the modification that is required. If the author chooses not to correct the item, at that point it is either removed from the item bank or kept in the bank with an explanation of the type of correction required. The first review committee will then determine whether the uniqueness and quality of the item is worth the effort required to fix it. Obviously, it is best if the original author corrects the problem, preserving the integrity of his or her intention in constructing the item.

A New Item Review meeting is conducted for each Level, at which all new items are reviewed by a multidisciplinary group of experts. Each item is carefully evaluated and refined, if necessary, to reflect the principles of quality items.

One to two months after the New Item Review meeting, an Approved Item Review meeting may be held. This separate multidisciplinary group of experts reviews the same items yet again, using the same process. Multiple reviews are essential to ensure that items are error-free and ready for publication.

After this review, the items to be used for pretesting are selected based on content, and presented to a Preliminary Exam Review committee, which reviews these items, using the same scrutiny as the first two committees have used.

At this point, another proofreading is conducted by a format expert, followed by a proofing by a group of four physicians.

We are now (finally!) ready to publish the exam. (And this process is just for new items—items to be pretested). Once they have been pretested, candidate comments are reviewed to evaluate for ambiguities. Any potentially flawed items are reviewed by Level Coordinators, who see a summary of the candidate comments and the statistics gleaned from pretesting the items. The Level Coordinators then decide if any items should not be used as scorable items in future exams, based on statistics, expert review, and candidate comments.

Once items are approved for use in actual, scored exams, the exam versions are selected. These exam items are then reviewed yet again – by another multidisciplinary group of experts – which can see the statistics obtained from pretesting along with any comments accrued during pretesting. The cycle for scorable exam items is shown on the following page:
**Item-Writing Practice Session (22.0)**
Having learned all of the fundamental item-writing principles and techniques, the potential item-writer is ready to evaluate flawed items, correct them, and construct new items. This final section of the item-writing guide will guide the item-writer through this process.

**Critiquing Written Items (22.1)**
First, we will start with flawed items. Why don’t you see if you can identify the flaws in each of these items? (The flaws are listed at the end.)

1. You visit a retirement village to perform semiannual health maintenance examinations for the residents. You ask each resident questions to assess current psychological status. Which of the following neuropsychiatric findings could be predictive for the development of Alzheimer disease?

   (A) depressive symptoms  
   (B) manic episodes  
   (C) personality disorder  
   (D) psychosis  
   (E) schizophrenia
2. As the Chief of Medicine in a community hospital, you are reviewing a patient death. In this case the peer review committee discovers an error that was the direct cause of the mortality. This information

(A) is available to anyone requesting medical records
(B) is not discoverable
(C) must be reported to the appropriate subspecialty board
(D) must be reported to the patient’s family
(E) must be reported to the state medical society

3. Which of the following conditions is most likely associated with direct hyperbilirubinemia, neonatal jaundice, and early-onset emphysema later in life?

(A) \(\alpha_1\)-antitrypsin deficiency
(B) Crigler-Najjar syndrome
(C) Gilbert syndrome
(D) glucose-6-phosphate dehydrogenase deficiency
(E) hereditary spherocytosis

4. A mother brings her child to the office, complaining that he is being unusually “difficult” lately. Which of the following conditions is associated with her child’s symptoms?

(A) folate deficiency
(B) HIV
(C) iron deficiency
(D) lead poisoning
(E) malnutrition

5. A 75-year-old nursing home resident has developed signs of rickets. You should increase the patient’s daily intake of the vitamin that is suspected to be deficient by

(A) 5%
(B) 10%
(C) 50%
(D) 80%
(E) 100%
6. A 26-year-old female presents with a complaint of spontaneous vaginal bleeding and cramping ten weeks after her last normal menses. Although she has not begun prenatal care, she reports that a home pregnancy test was positive. Examination reveals that her vagina is filled with clotted blood, her cervix is dilated, and products of conception are present in the external os. You recommend dilation and evacuation to limit further blood loss, but the patient immediately refuses because she regards the procedure as "murder." Which of the following ethical principles applies to this situation?

(A) beneficence  
(B) fetal viability  
(C) informed consent  
(D) respect for autonomy  
(E) the fetus as a person

**Flaws:**
1. (a) This is a pseudocase where the clinical scenario relates to the question but the candidate does not need the clinical scenario to answer the question  
   (b) There are inconsistent answer choices (some are symptoms, others diagnoses)

2. (a) The information in the stem is not necessarily universal  
   (b) Choices A and B are different than the others

3. (a) There is no clinical scenario in the stem  
   (b) Answer choices are inconsistent

4. There is not enough information in the stem to answer the question without looking at the options

5. (a) This tests percentages  
   (b) The information is not commonly known

6. (a) Answer choices are unclear  
   (b) More than one answer choice could be correct

**Building a New Exam Item (22.2)**
The ultimate goal of this guide is to enable the new item-writer to write a good exam item and to help a good exam item-writer write an even better one. As mentioned earlier, exam item writing is both science and art. The item-writer must take the experience and expertise of his or her particular discipline and translate it into a scorable exam item that can help the medical boards make an appropriate decision for licensure. The ability of the COMLEX-USA to attain that end begins with the exam item-writing assignment, and is maintained through careful quality assurance and continual product improvement.
As an item-writer, you are about to embark on the most essential element of the COMLEX-USA examination sequence - the new exam item. This is what the whole process is about—the construction of a good exam item.

Utilizing your experience and expertise, create an exam item for your discipline that would be appropriate for a Level 1, 2 or 3 candidate, by doing the following:

1. Review the COMLEX-USA blueprint specifications (pages 2-3) to select a topic of interest from each dimension according to the design of the test. (Please be aware that all item-writers are asked prior to receiving assignments what dimension 1 topics they are willing to use in writing items. The writer then receives an assignment that pairs pertinent topics with dimension 2 physician tasks, based on NBOME needs.)

2. Consider the seven osteopathic core competencies (summarized on page 31 and detailed through the NBOME Web site: www.nbome.org) for framing the exam item.

3. Create an item-writing assignment similar to the grid (pages 5-6).

4. Create item details and consider additional essential features such as age and gender (page 6).

5. Write an interrogatory that clearly asks what the candidate is expected to demonstrate that he or she knows or knows how to do (pages 14-16). You may elect to start with the scenario and build the interrogatory later.

6. Develop the stem of the exam item using one of the suggested exam item patterns (pages 16-17).

7. Review the stem to be sure that sufficient information has been included (to make the distractors plausible and an answer possible), and that extraneous and irrelevant or confounding information has been minimized.

8. Write a list of 5 (or more for X-Type and extended A-Type) parallel options, keeping in mind the cardinal rules that make for a series of good distractors as described (pages 21-22).

9. Verify that the candidate must use a higher cognitive level, in accordance with Bloom’s Taxonomy (pages 23-24).

10. Check grammar and format for clarity and consistency of tenses.

11. Determine if an appropriate graphic or image might improve the exam item or test more areas of knowledge or skill, adjusting the text accordingly so that the graphic or image is essential to the item.

12. Indicate the reference (pages 10-12).
13. Determine whether the best level of available medical evidence has been used (pages 25-30).

14. Ask if the exam item could lend itself for use as an "advanced item type" (pages 31-36).

15. Critically review and revise the created exam item using the checklist (pages 19-20).

16. Verify that the exam item is still at the level for which the item was to be written.

17. Consider what percentage of candidates will answer the item correctly.

Use the index at the back of this guide to locate specific areas that may be helpful in the construction of the new exam item.

As you complete this exercise, make notes of the issues that arise and make notes as to how you might alter the approach to the creation of the next exam item so that it will be better than the last.

**Resources (23.0)**

Contact us at any time using the following resources:

<table>
<thead>
<tr>
<th><a href="http://www.nbome.org">www.nbome.org</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive testing department email address: <a href="mailto:testing@nbome.org">testing@nbome.org</a></td>
</tr>
<tr>
<td>NBOME phone number: 773-714-0622</td>
</tr>
<tr>
<td>NBOME’s free online Item-Writer Training Course. E-mail requests for materials <a href="mailto:iwriter@nbome.org">iwriter@nbome.org</a></td>
</tr>
</tbody>
</table>